Department for Transport

Provision of market research for value of travel time savings and reliability

Phase 2 Report

Issue | 14 August 2015

In Partnership with:

ITS Accent

Job number 236606-10

Ove Arup & Partners Ltd
13 Fitzroy Street
London
W1T 4BQ
United Kingdom
www.arup.com

ARUP

The following people have contributed to the production of this report, listed by firm.

Arup

Stefan Sanders
Efie Drivyla
Jeremy Bourdon

Institute for Transport Studies (ITS), University of Leeds (and affiliates)

Richard Batley
Manuel Ojeda Cabral
Phani Kumar Chintakayala
Charisma Choudhury
Andrew Daly
Thijs Dekker
Anthony Fowkes
Stephane Hess
Daniel Johnson
James Laird
Peter Mackie
Mark Wardman
John Bates (John Bates Services; working for ITS)
Michiel Bliemer (University of Sydney; working for ITS)
Maria Börjesson (Royal Institute of Technology, Sweden; working for ITS)
John Parkin (University of West of England; working for ITS)
Tom Worsley (Visiting Fellow, ITS)

Accent

Chris Heywood
Miranda Mayes
Rachel Risely
Nancy Curzon
Rob Sheldon

Contents

Acknowledgements

We would like to acknowledge the contributions of various parties which have helped to facilitate the study, as follows.

- Staff from the Department's Transport Analysis and Strategic Modelling (TASM) Division, in particular the project officer Adam Spencer, and colleagues Alice Crossley, Margot Shatz and Marcus Spray, who guided the study on a day-to-day basis.
- Members of the Department's wider Project Board, who reviewed the progress and direction of the study at key milestones.
- Stakeholders from various public and private sector organisations in the transport sector, who attended workshops at the outset of the study and at the end of the survey design and testing phase.
- Operators of railway stations, 'other PT' services, motorway service stations, petrol stations and other areas where interviewers worked.
- Members of the Analytical Challenge Team who reviewed our Main Report and offered comments namely Professor Glenn Lyons (University of West of England, UK), Professor Jonas Eliasson (Royal Institute of Technology, Sweden) and Professor Kay Axhausen (Swiss Federal Institute of Technology).

1 Introduction

1.1 Background to the commissioning of this study

This document represents the Final Report of the study **'Provision of market research for value of time savings and reliability'** undertaken by the Arup/ITS Leeds/Accent consortium for the Department for Transport (the Department).

In the context of transport appraisal, one of the most important concepts is that conventionally referred to as the 'value of time'. This does *not* refer to the value that might be placed on time spent in travel, but should be seen as shorthand for the 'value of changes in travel time', relative to a reference case when investment takes place. These changes may be positive or negative but historically have been referred to as 'savings'. In this report we have chosen to refer to the 'value of travel time' (VTT) to convey this concept.

Travel time savings are usually the largest single component of the monetised benefits of projects. Furthermore, time-related benefits such as reliability and relief of overcrowding on public transport (PT) are conventionally valued through multipliers on the value of travel time.

There have been three waves of national VTT studies in Britain. First, a series of research studies during the 1960s, the results of which were adopted and synthesised by the Department. Second, the MVA/ITS Leeds/TSU Oxford study of the 1980s leading to revised values of travel time in 1987. Third, the AHCG study of 1994 which was further analysed by ITS Leeds and new guidance provided in WebTAG in 2003[1]. These values have subsequently been revised and updated by the Department to reflect changes in incomes and travel patterns, as documented in WebTAG A3.1[2].

However, the evidence base for the values used in appraisal is now twenty years old. During that time, incomes, prices, demography and the mix of travel by purpose and trip length have all changed. But possibly more significant is that the world has changed in other ways – the internet revolution, the quality and comfort of vehicles, working practices and, perhaps most fundamentally, the ways in which people perceive time spent travelling. It is not really possible to accommodate such phenomena simply by updating historical values.

Against this background, the Department has, since 2009, taken steps to review the theoretical, methodological and evidential basis of WebTAG A1.3, and thereby respond to the emerging critiques such as those noted above. Among the key actions have been the Department's commissioning of the following scoping and review studies:

[1] Mackie, P.J., Wardman, M.R., Fowkes, A.S., Whelan, G.A., Nellthorp, J., and Bates, J.J. (2003) 'Value of Travel Time Savings in the UK. Report to Department for Transport'. Available at: http://eprints.whiterose.ac.uk/2079/2/Value_of_travel_time_savings_in_the_UK_protected.pdf

[2] https://www.gov.uk/government/publications/webtag-tag-unit-a1-3-user-and-provider-impacts-november-2014

- Values of travel time savings: updating the values for non-work travel (ITS Leeds, John Bates and DTU, 2010)[3]: This study scoped out the research activities that would be required to update the values for non-work travel time savings, and issued recommendations about which elements should be taken forward.
- Values of travel time savings for business travellers (ITS Leeds, John Bates and KTH, 2013)[4]: This study reviewed the feasibility and theoretical accuracy of different methods for estimating VTT for business travellers, and evidence from the UK and overseas concerning the values emanating from the different methods. A particular stimulus for this study was the critique surrounding the productive use of travel time.
- Values of travel time savings: understanding the uncertainty around the non-work values (ITS Leeds and John Bates, 2013)[5]: This study estimated a range around the values based on the level of statistical confidence.
- Values of travel time savings: analysis of non-work values since 1994 (ITS Leeds, Arup and URS, 2013)[6]: This study explored the existence of relationships between changes in values of travel time savings and changes in socio-economic and technological factors, using a meta dataset of more than a thousand values of travel time savings spanning almost 50 years.
- Peer review of proposals for updated values of travel time savings (ITS Leeds, John Bates, Arup and URS, 2013)[7]: This study peer reviewed the methodology used by the Department to update appraisal values of travel time savings, and audited the updating calculations.

Arising from this body of work, the Department's conclusions with respect to the valuation of travel time savings for both non-work and business were documented in the 2013 report from the TASM Division[8]. A key conclusion was the Department's decision to commission new market research to estimate the value of travel time savings and reliability. The contract for this market research was subsequently tendered through the T-TEAR Lot 2 Framework and, following a competition, was awarded to the Arup/ITS Leeds/Accent consortium in June 2014.

[3] https://www.gov.uk/government/publications/values-of-travel-time-savings-updating-the-values-for-non-work-travel

[4] https://www.gov.uk/government/publications/values-of-travel-time-savings-for-business-travellers

[5] https://www.gov.uk/government/publications/values-of-travel-time-savings-understanding-the-uncertainty-around-the-non-work-values

[6] https://www.gov.uk/government/publications/values-of-travel-time-savings-analysis-of-non-work-values-since-1994

[7] https://www.gov.uk/government/publications/peer-review-of-proposals-for-updated-values-of-travel-time-savings

[8] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/253485/technical-research-next-steps-appraisal.pdf

1.2 Study aims and objectives

The Department set the following **aims** for the study:

- Provide recommended, up-to-date national average values of in-vehicle travel time savings, covering business and non-work travel, and based on primary research using modern, innovative methods.
- Investigate the factors which cause variation in the values (e.g. by mode, purpose, income, trip distance or duration, productive use of travel time etc.) and use this to inform recommended segmentation of the values.
- Improve our understanding of the uncertainties around the values, including estimating confidence intervals around the recommended values.
- Consistently estimate values for other trip characteristics for which values are derived from the values of in-vehicle time savings.

The overall **objectives** of the study were to:

- Provide recommended, up-to-date national average values for in-vehicle travel time savings.
- Improve understanding of what drives the values of travel time savings and the uncertainty around the values.
- Consistently estimate values for trip characteristics of related factors, e.g. reliability and crowding.

1.3 Delivery of the study

With reference to **Table 1.1**, the study has been conducted in two phases. An inception meeting with the Department was held on 3rd June 2014, and was immediately followed by a workshop with stakeholders potentially affected by revisions to VTT guidance.

Table 1.1: Study phasing and timelines

Date	Activity
June 2014	Inception
June-September 2014	**PHASE 1:** Development and testing of survey approach
September-October 2014	Review of Phase 1 and decision to proceed
October 2014-March 2015	**PHASE 2:** Field survey Estimation of behavioural values Conversion to national average appraisal values
March 2015	Submission of draft Final Report
April 2015	Review and audit Finalisation of Final Report

Phase 1 of the study, which was undertaken from June to September 2014, involved the development and testing of methods for undertaking the requisite market research. This phase culminated in an extensive pilot survey conducted in two waves, and the estimation of behavioural values on this dataset.

Having reviewed the Phase 1 report, and convened a further workshop with stakeholders, the Department took the decision to proceed to Phase 2, which was undertaken from October 2014 to March 2015.

Using the methods developed in Phase 1, Phase 2 involved a substantial field survey, the estimation of behavioural values of travel time savings on this dataset, and the conversion of behavioural values into national average appraisal values. At key milestones, study deliverables and outputs have been scrutinised at various levels, as follows:

- The study team has adhered to rigorous internal quality-assurance procedures throughout.
- An important component of the Arup/ITS Leeds/Accent bid was an ‘Analytical Challenge Team’ (ACT). The ACT comprised leading academics and practitioners from the VTT field, and was positioned at ‘arm’s length’ from the core study team. The ACT has independently reviewed all study reports.
- All study reports have also been reviewed by the Department’s Project Board, comprising representatives of key divisions potentially affected by revisions to VTT guidance.
- Finally, through a separate contract, the Department commissioned the SYSTRA/Imperial College London/Technical University of Denmark consortium to undertake an independent peer review and audit of the data collection and modelling work. The findings of this audit are reported in a self-standing deliverable.

1.4 Scope of the study

The scope of the study can be summarised as follows.

1.4.1 Trip purpose and mode

In pursuit of the aforementioned aims and objectives, we employed an analysis framework based upon the primary dimensions of **trip purpose** and **mode** (see **Table 1.2**). The Department commissioned analysis of the three ‘required’ modes of car, bus and rail, plus two ‘optional modes, namely walk and cycle and ‘other PT’; the air access mode was not commissioned.

At the outset, it was acknowledged by both the Department and the study team that, given the dearth of previous VTT research on walk and cycle, Stated Preference (SP) analysis of this mode would be somewhat exploratory. As it transpired, the process of developing and testing the survey approach for walk and cycle had to be repeated several times before ‘proof of concept’ could be established. Consequently, it was agreed with the Department that walk and cycle would be researched on a deferred schedule, and reported separately from the mechanised modes. The remainder of this report will not therefore cover walk and cycle in any detail.

With regards to the mechanised **modes**, it should be clarified that our definition of bus covers urban and non-urban local services, but not inter-urban coach. A number of minor modes such as inter-city coach, taxi/hire car and motorcycle were also excluded from consideration.

Table 1.2: Summary of survey design

			Trip Purpose				Experiments	Covariates
			Commuting	Non-Work	Employees' Business	Employers' Business		
Mode	Required	Car	SP	SP	SP	SP	i) Time ii) Time & Reliability iii) Time & Quality (e.g. crowding, congestion and other types of time)	a) Income b) Distance/Duration c) Productive Time d) Trip Type
		Bus	SP	SP	N/A	N/A		
		Rail	SP & RP	SP & RP	SP & RP	SP		
	Optional	Walk & Cycle	SP	SP	N/A	N/A		
		'Other PT'	SP	SP	SP	SP		
		Air Access	N/C	N/C	N/C	N/C		

N/A = Deemed to be not applicable on the grounds that trip rates are relatively low
N/C = Not commissioned by the Department
SP = Stated Preference
RP = Revealed Preference

With regards to **trip purposes**, we segmented non-work by commuting and 'other' non-work, on the grounds that these definitions are well-established in practitioner guidance, and provide an appropriate level of detail for most requirements. For business, we segmented by employee and employer. It was decided by the Department that professional drivers should be excluded from the scope, since it was unlikely that the Cost Saving Approach (CSA) would be improved upon for this category. We followed the National Travel Survey (NTS) categories in defining professional drivers to be those whose sole or main occupation is driving a vehicle. Service engineers and others who use a vehicle as a tool of their trade were deemed to be within scope.

Since some purpose/mode combinations (e.g. employees' business by walk and cycle) account for a modest number of trips in NTS (outside London, at least), it was decided by the Department and the study team that these combinations were out of scope. Also, non-residents of the UK were excluded from the survey work.

We acknowledge that these exclusions leave gaps in the coverage of our behavioural valuations of travel time savings and reliability. However, for appraisal purposes, we recommend that these gaps should be filled by 'transferring' valuations from other purpose/mode combinations.

In order to prepare the ground for the experiments, qualitative work was undertaken with focus groups and this is described in **Chapter 2**. We pass immediately to the main features of the experimental design, as laid down in the Department's ITT.

1.4.2 Experiments

For each purpose-mode combination within scope, multiple variants of SP experiment were developed, involving different **experimental** trade-offs: time/money (SP1), time/money/reliability (SP2), and time/money/quality (SP3). These different experiments are described in **Chapter 3**.

Whilst the focus of our data collection was SP, we have sought to validate our SP-based estimates of travel-time savings and related factors against Revealed Preference (RP). For reasons of cost-effectiveness and practicality, our RP data collection was focussed on rail; a core reason being the relative lack of meaningful time vs. money trade-offs in real world car driving situations in the UK, with the exception of a few river crossings and one toll road. A more detailed explanation of the SP and RP design is given in **Chapter 2**.

In general terms, our SP experiments offered an abstract choice using 'unlabelled' alternatives (i.e. A vs. B). Where possible, we 'pivoted' attribute levels around travellers' current trips. The current trip was identified either through interception of the traveller in the course of a trip, or through telephone interview where the respondent was asked to think back to a recent trip; further discussion of the survey approach is provided in **Chapter 2**.

The attraction of the pivot approach – which is in our view current best practice in SP analysis of VTT – is that the SP experiment presented to any given respondent is grounded in realism. We made some exceptions to this approach, for attributes and/or choice contexts where we believe it is more appropriate to employ 'fixed' attribute levels.

Such attributes included:

- Headway; where regular timetabled services constrain headway to fixed increments (5, 15, 30 mins, etc.).
- Crowding; where a pivot approach would have been unduly complicated.

Such choice contexts included:

- Rail operator choice; where we focussed on the same O-D trips as the RP analysis.

1.4.3 Covariates

In the course of conducting pilot surveys for both SP and RP, we collected background data relating to the traveller and trip, which could potentially be employed as **covariates** in valuations of travel time savings and reliability. Such background data included:

- **Income:** To encourage as many people as possible to provide this information, we emphasised that this data would be confidential and used for analysis purposes only. We asked respondents to select an income band rather than record a specific amount, and we also asked questions about both household and personal income, as well as an income proxy (monthly accommodation costs).
- **Distance/duration:** We employed a Google maps tool to identify the origin and destination of the trip of interest, and to automatically derive the distance. For corroborative purposes, we also asked the respondent the trip distance and times of departure and arrival.
- **Productive time:** We considered the ability to undertake both work and non-work activities whilst travelling and the ‘productivity’ of travel time devoted to these activities.
- **Trip type:** We made provisions in our survey design such that our results could be segmented by industry-standard categories for different modes, for example urban/inter-urban/rural for car, and short/long/London and South East (LSE) for rail).
- **Group size and composition:** We specified group size and composition as a segment to be identified in the surveys, and to be carried through to the appraisal values by having an influence on the average value for each distance/income band/etc. segment.
- **Driver vs. passenger:** Following from group size and composition, our car surveys distinguished between the driver and the passenger.
- **Trip frequency:** Our surveys asked the respondent how often they made the trip of interest; familiarity with the trip is especially pertinent to valuations of reliability.
- **Time of day:** Further to the discussion under distance/duration above, we asked the respondent for the start and end time of the trip of interest.

In the case of **business travel**, we collected additional background data, especially concerning the company’s policy towards travel, and the extent to which such policies influenced travel planning.

1.5 Spatial scope

The study was focussed on England but with some coverage in Scotland and cross border flows into Wales.

1.6 Exclusion from scope

With reference to earlier discussion of **Table 1.2** above, we note that the Department excluded the following interests from the scope of this study:

- Access trips to airports
- Business travel by bus and by main mode walk and cycle
- Inter-urban bus/coach
- Large scale meta-analysis of VTT and Value of Travel Time Reliability (VTTR)
- The 'Hensher' approach to estimating the value of business travel time savings

1.7 Additions to scope

In the course of Phase 1 of the study, the following variations on the initial scope were agreed with the Department.

- In order to accommodate the 3-game format (i.e. SP1, SP2 and SP3), the Department commissioned a degree of additional data collection. However, this variation did not affect the pilot survey conducted in Phase 1.
- 30 cognitive depth interviews were commissioned to allow for detailed testing of the pilot questionnaires, specifically the SP options.
- The walk and cycle and 'other PT' options in the proposal were taken up, although the former are reported in a separate deliverable.

1.8 Purpose of this report

The requirements of this Final Report are:

- To describe the methods employed in undertaking all research tasks.
- To present detailed results of the modelling.
- To issue practical guidance on valuations to be used for modelling and appraisal purposes.
- To issue recommendations.

This report is accompanied by all data and calculations needed to reproduce the recommended values, as well as all data collection tools.

In addressing these objectives, the layout of the report is as follows:

Chapter 2 describes the research methods which have been devised and implemented. This covers the design and testing including focus groups, cognitive testing and pilots and all aspects of the SP and RP data collection.

Chapter 3 describes the market research findings for the main field surveys covering the general public SP, employers' business SP and RP questionnaires.

Chapter 4 describes the modelling of the 'core' SP data covering the commute, employees' business and other non-work purposes for each mode, and reports estimates of behavioural values of travel time and related factors.

Chapter 5 describes a series of 'auxiliary' models covering a range of purpose/mode combinations, including RP and SP operator choice, SP mode choice for concessions and non-concessions, and employers' business SP.

Recognising that valuations of business travel time savings are an important part of the brief, **Chapter 6** reconciles evidence on these valuations from different perspectives – SP vs. RP, and employee vs. employer.

Chapter 7 describes the processed involved in converting behavioural values to appraisal values, and reports national average estimates of appraisal values of travel time and related factors.

Chapter 8 summarises the research undertaken, reports key findings and issues recommendations.

A number of appendices are also provided, which offer additional detail on specific technical aspects of the work.

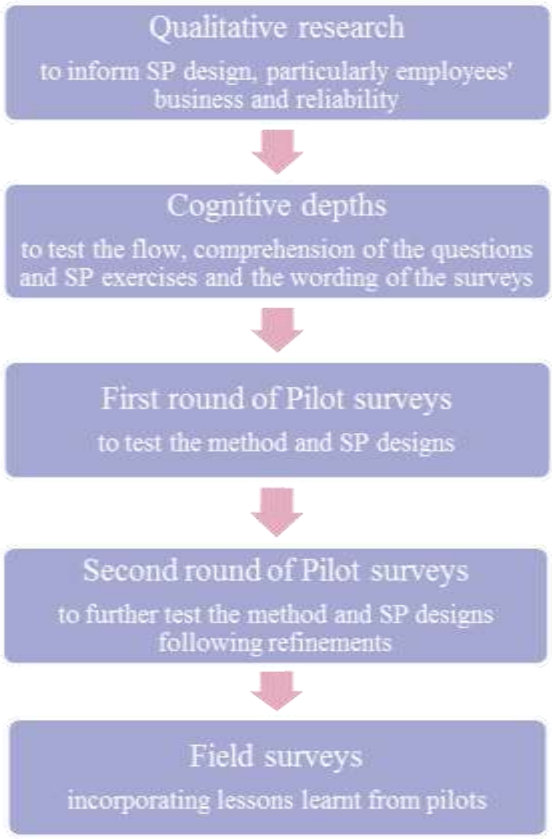
2 Research Method

2.1 Overview

In the ITT for this study, the Department stipulated that: *"Data collection and analysis should make use of up-to-date, innovative techniques, including the extension of willingness-to-pay techniques to valuing business travel time savings"*. Furthermore, the ITT noted – quite reasonably – that: *"The quality of collected data and evidence will be crucial to the robustness of the resulting values. Regardless of the technique to be used, the data collection approach should be thoroughly piloted and tested, e.g. through piloting and cognitive testing of surveys and survey delivery methods"*.

Against this background, this chapter sets out the core methodology and the process for designing and developing the data collection and analysis methods. As noted in **Chapter 1**, these methods were focussed around Stated Preference (SP), but complemented by Revealed Preference (RP) as a validation device. These methods were designed and developed in a systematic fashion, as summarised in **Figure 2.1**. Whilst the ITT directed us to focus on the development and application of *empirical* methodology for measuring VTT, it is worth remarking that, from a *theoretical* perspective, we subscribe to the standard microeconomic framework underpinning both non-work and business VTT, as summarised in **Appendix A**. That is to say, we view the empirical methods reported here to be a best-practice (and in some cases, beyond best-practice) implementation of the standard theory.

Figure 2.1: Survey design process



The SP surveys covered the following segments:

- Commuting
- Other non-work travel
- Employees’ business travel
- Employers’ business travel

and three SP experiment types:

- Time vs. cost (SP1)
- Time vs. cost vs. reliability (SP2)
- Time vs. cost vs. quality (SP3)

To aid the subsequent discussion of the design process, **Figure 2.2** synthesises the context and content of the different experiments which were eventually developed.

Figure 2.2: Summary of SP formats by game and mode

Game and mode	Description of SP format
SP1	SP1 used a generic format across all modes, presenting respondents with a choice between two options described only on the basis of travel time and travel cost, where one option is cheaper, but the other option is faster.
SP2	SP2 also presented respondents with a binary choice, still focussing on travel cost and travel time but where, for travel time, five different typical trip outcomes were presented for each alternative as a representation of travel time variability.
SP3	SP3 used somewhat different presentations across modes while nevertheless retaining a binary choice context with one exception (rail operator choice).
SP3 car	For car, the two options were described in terms of travel cost for each trip and the amount of time that each trip spends in three types of driving conditions (free-flow, light traffic, heavy traffic).
SP3 rail	For rail, two different experiments were used, with a split between respondents that is discussed later in this chapter. 1. For the first group, we presented a choice that is very similar to SP1, with the difference that for each alternative, we additionally define the level of crowding applying to the trip. 2. For the second group, we presented a choice between up to three operators, described in terms of travel time, fare and headway.
SP3 bus	For bus, two different experiments were also used. 1. For the first group, we presented a crowding game analogous to the rail game, albeit with different crowding definitions. 2. For the second group, we presented a choice between two bus routes described in terms of free-flow time, slowed down time, dwell time, headway and fare.
SP3 ‘other PT’	For ‘other PT’, two different experiments were used once again. 1. For the first group, we presented a crowding game analogous to the bus game. 2. For the second group, we presented a mode choice game (‘other PT’ against either bus or rail) using time, headway and cost as attributes.

Somewhat different games were presented for concessionary travel, and these are discussed in more detail later in this chapter.

By contrast, the RP surveys covered a restricted part of the brief, specifically:

- Rail commuting
- Rail other non-work travel
- Rail employees' business travel

Section 2.12.4 provides a detailed rationale for the narrower focus of the RP.

The rest of this chapter is structured around three main elements:

- **Section 2.12 to 2.11** cover the methodology utilised for the main stage SP and RP surveys.
- **Section 2.12** covers the key SP and RP design issues.
- **Section 2.13** discusses the development and testing of the methodology.

2.2 Market research method

There were three distinct surveys:

- General public Stated Preference (SP) research
- Revealed Preference (RP) research
- Employers' business SP research

An overview of the data collection method for each is shown below.

2.3 General public SP research method

The core research method for the SP survey was an intercept approach with on-line or telephone interviews supplemented by telephone recruitment.

For the general public SP research, the proportions of intercept and phone recruitment were:

- Circa 80% via intercept approach
- Circa 20% by phone

Taking the NTS as the benchmark for representativeness, 80% of all trips recorded in the NTS are less than 10 miles. However, most studies have found a strong relationship between VTT and distance, and as we use distance weighting (see **Chapter 7**) it is important to be able to estimate VTT for longer trips accurately. Using the NTS definition of 'long distance' (i.e. greater than 50 miles), only 2% of trips in the NTS fall into this category.

A telephone sampling approach predominantly samples short distance trips. On the other hand an intercept sampling approach, as has been conventional in most VTT studies, favours longer distance trips, since these have a higher probability of being intercepted.

Based on these considerations, the proposed approach was two-pronged: predominantly using intercepts to ensure an adequate sample of the longer distance movements, and more generally business trips, but using telephone sampling to strengthen the sample in the shorter distances.

Car, rail and bus users were recruited through both methods. ‘Other PT’ users were intercept recruited.

80%:20% was a judgement based split, intended to focus the majority of the dataset on the intercept survey (which we anticipated would be biased towards longer trips and more frequent travellers), whilst devoting adequate dataset to the ‘corrective’ device of the telephone survey (which would capture a wider range of trips/travellers). This plan was conceived on the basis that any residual bias in trips/travellers in the sample would be corrected at the implementation stage (as discussed in **Chapter 7**).

The intercept approach was also chosen since interviewers can be located where the target respondents are (e.g. at bus stops, rail stations at motorway service areas). This was particularly important to be able to recruit adequate samples of specific groups of target respondents who would otherwise be extremely difficult to recruit through other sampling approaches, namely:

- Those making specific ‘other PT’ and bus trips on corridors where there was a rail alternative (required for the operator choice SP exercise).
- Those making trips on specific rail routes (to provide comparisons with the RP sample).
- Long distance car and rail travellers.
- Employees’ business travellers.

A household sampling approach was also rejected because of the overall survey time constraints. Whilst some practitioners advocate face-to-face or telephone interview approaches for SP-based WTP surveys, on-line surveys are considered acceptable provided:

- the complexity (in terms of the language used, length of descriptions, number of tasks and number of exercises) is minimised;
- qualitative research is undertaken to determine what language is most appropriate; and
- cognitive testing is undertaken to confirm that respondents understand the language and, more generally, understand the tasks being presented.

All of these measures were undertaken in the course of the present study, and we are confident therefore that the on-line survey method was an appropriate one.

2.4 Revealed Preference (RP) method

Whilst SP methods have advanced considerably since the last national UK study of VTT in 2003, there continues to be a debate as to whether SP can elicit credible valuations of travel time savings and reliability. It therefore makes sense to validate SP against robust RP data, where:

- there are real-world situations where spending money saves time;
- travellers are familiar with opportunities available; and
- time vs. cost trade-offs offer real choices.

In discussion with the Department, it was decided to focus RP analysis on rail; see **Section 2.12.4** for a detailed rationale.

The RP surveys were recruited entirely through an intercept approach (see **Section 2.6.4** to follow) and comprised users of the following stations who were travelling to London:

- Birmingham New Street
- Birmingham Moor Street
- Birmingham Snow Hill
- Stoke
- Stafford
- Rugby
- Peterborough

These locations were chosen as they were considered to offer the best range of real world choice between cost and time for rail travellers.

At each station, interviewers stood on the relevant London-bound station platforms, which were determined in advance to ensure split between operators. Recruitment included screening to ensure participants were:

- familiar with the times and costs of alternative operators; and
- making a rail trip originating at the surveyed station and ending at the London terminus station.

2.5 Employers' business SP method

The surveys of employers focussed on 'briefcase' business travel undertaken by their employees. Briefcase travel was defined as *"trips made by office-based staff travelling to conduct meetings and similar business activities but not to provide trade services"*.

The surveys were undertaken by telephone, and the target respondent was the person within the company who was responsible for making decisions about how employees travel for business purposes. In larger companies there are often many such people, provoking the concern that responses may be dependent on the specific person interviewed. However, we can take reassurance from the fact that definitive travel policies were more prevalent in larger companies; in our sample, 77% of companies with over 250 employees had formal travel policies, as compared with 23% for companies with less than 20 employees.

2.5.1 Recruitment

The telephone sample used was supplied by Sample Answers and used LBM Direct Marketing and Experian Business Files, which in turn are based upon Thomson Directories and Companies House.

The business sample was loaded into Accent's Telephone Unit software which randomly allocates a number to each logged-in interviewer.

The target respondent was *"the person within the company who was responsible for making decisions about how employees travel for business purposes, for example when travelling to meet clients, customers or suppliers or when* *travelling between different offices within their organisation"*. In smaller companies this could be the owner, managing director, finance director, operations manager, procurement manager or HR manager. In larger companies this person could also be a travel manager or fleet manager.

The SP options were based on a hypothetical trip that an employee made by either car, train or 'other PT'. The target was 400 interviews, with 133 for each of the modes.

Respondents were sent the SP options for the three SP exercises. These were customised based on answers to various questions within the questionnaire, and then e-mailed or posted.

2.5.2 Quotas

In addition to the mode used by the employee for the hypothetical trip (133 car, 133 rail and 133 'other PT'), there were quotas on company size, industry grouping and region as shown in **Table 2.1**, **Table 2.2** and **Table 2.3** respectively.

Table 2.1: Company size (employees)

Number of employees	Number of companies
1-19	66
20-49	66
50-249	133
250+	133

Table 2.2: Broad industry grouping

<table>
<tr><th>Industry</th><th>Min:max</th></tr>
<tr><td>Agriculture, Forestry and Fishing</td><td rowspan="2">5-12%</td></tr>
<tr><td>Mining and Quarrying; Electricity, Gas and Air Conditioning Supply; Water Supply; Sewerage, Waste Management and Remediation Activities</td></tr>
<tr><td>Manufacturing</td><td rowspan="2">15-20%</td></tr>
<tr><td>Construction</td></tr>
<tr><td>Wholesale and Retail Trade; Repair of Motor Vehicles and Motorcycles</td><td rowspan="2">3-10%</td></tr>
<tr><td>Transportation and Storage</td></tr>
<tr><td>Accommodation and Food Service Activities</td><td rowspan="2">7-13%</td></tr>
<tr><td>Information and Communication</td></tr>
<tr><td>Financial and Insurance Activities</td><td>7-13%</td></tr>
<tr><td>Real Estate Activities</td><td>7-13%</td></tr>
<tr><td>Professional, Scientific and Technical Activities</td><td>7-13%</td></tr>
<tr><td>Administrative and Support Service Activities</td><td rowspan="4">11-18%</td></tr>
<tr><td>Education</td></tr>
<tr><td>Human Health and Social Work Activities</td></tr>
<tr><td>Arts, Entertainment and Recreation</td></tr>
<tr><td>Other Service Activities</td><td>6-12%</td></tr>
</table>

Table 2.3: Region

Industry	Min:max
North East	1-5%
North West	8-14%
Yorkshire and the Humber	6-10%
East Midlands	6-10%
West Midlands	6-10%
East of England	9-15%
London	16-22%
South East	16-22%
South West	8-14%

Company size quotas were determined in discussion with the Department, with a view to ensuring adequate coverage of all industry groupings, whilst also focussing some survey effort on larger companies undertaking ‘briefcase’ travel.

2.5.3 Monitoring

The telephone fieldwork quality was monitored through listening in. The telephone unit management software produces detailed call outcome reports.

10,340 recruitment calls were made and the main call outcomes were:

- Recruited 5%
- Call again/not available during survey period 5%
- No reply/answerphone 57%
- Refusal 16%
- Number not recognised/fax etc. 8%
- Not in scope[9] 9%

2.5.4 Recruitment issues and mitigations

It proved very difficult to find ‘other PT’ users for a number of reasons:

- Many ‘other PT’ trips were not the main mode, but the access and egress mode for a rail trip.
- ‘Other PT’ trips were limited to some cities, but the sample was designed to be representative by region.
- ‘Other PT’ trips often appeared to be ‘under the radar’ as far as some employers were concerned, since the trip was typically short, less likely to be subject to company travel policy, not paid for in advance, and/or covered by an employee’s season ticket or travel card. We will return to this issue when reconciling business values of time in **Chapter 6.**

[9] For example, no business trips were made by employees.

Due to the above issues, it was agreed with the Department to drop the 'other PT' quota, and to increase the car and rail quotas to compensate.

Towards the end of the fieldwork period, we used an external agency to support the recruitment of rail respondents (36 interviews came from this source) and increased the incentive to £20.

2.6 General public intercept recruitment method

The intercept CAPI[10] survey was administered face-to-face using Android tablets. Interviewers approached a random sample of adults (typically 1 in 3) and asked scoping questions to check whether the respondent was in-scope and matched required quotas. If in-scope, the respondent was invited to undertake a follow-up survey on-line or by phone. The interviewer collected their contact details (name and telephone number for follow-up telephone interview and name and e-mail address for follow-up on-line survey). The intercept interview data was uploaded to Accent's servers during or after each fieldwork shift. Those providing e-mail addresses were sent an e-mail with a unique web-link to the survey at the end of the shift. The names and phone numbers of those preferring to undertake the interview by phone were loaded into the telephone unit sample on a daily basis.

All intercept fieldwork took place on weekdays with fieldwork shifts either 07:00-13:00 or 13:00-19:00.

Further details of the recruitment method including a copy of the recruitment questionnaire, the email text, and specific intercept methods for each mode are included in **Appendix E.**

2.6.1 Sampling locations

Figure 2.3, shows the intercept locations, which were designed to cover car, rail, bus and 'other PT' users across the country. The survey locations were selected to reflect:

- Coverage of the key trip purposes.
- A reasonable geographical spread across England, some coverage in Scotland as well as some cross-border flows into Wales.
- A reasonable spread of the key market segmentations relevant to each mode[11].
- Specific locations where travellers have a real opportunity to choose between different trips with different times, costs, reliability and/or quality features.

[10] Computer Aided Personal Interview

[11] Ensuring that, for example, rail included flows such as London long, non-London long, South East outer, South East inner, car included inter-urban, urban and rural, and bus included London, Metropolitan/PTE, freestanding large urban areas and market towns/rural hinterland.

Figure 2.3: Maps of the sampling locations for rail, car, bus and 'other PT'

Rail SP locations

Car locations

Bus locations

'Other PT' locations

Rail RP locations

Appendix B lists the locations by mode and provides contextual information on each of the locations. In total, 881 intercept shifts were undertaken as follows:

- Bus 73
- Car 239
- 'Other PT' 109
- Rail SP 229
- Rail RP 231

The numbers of intercept recruits by mode and purpose are shown in **Table 2.4**.

Table 2.4: Intercept recruitment interviews by mode and purpose

	Employees' business	Commute	Other non-work	Concessionary
Bus	N/A	1,401	1,295	193
Rail RP	3,963	610	2,282	N/A
Rail SP	2,932	2,571	2,413	N/A
'Other PT'	909	2,167	1,086	N/A
Car	2,574	2,025	1,865	N/A
Total	**10,378**	**8,774**	**8,941**	**193**

Sampling took place at the following types of locations for the different modes:

- Car: Motorway and A-road service stations and petrol stations in towns or in local high streets nearby.
- Bus: at bus stops.
- Rail: at rail stations.
- 'Other PT': London Underground: at stations / Metro and tram: at stops.

Bus concessionary passholders were only sampled in Sheffield (SYPTE), Leeds (WYPTE), Bristol and Brighton, and on bus routes with a parallel rail route so that a bus vs. rail SP exercise could be undertaken.

For 'other PT', most respondents (except those sampled in London Underground central locations) were sampled on routes towards the centre (along a rail or bus route) so that a bus or rail vs. 'other PT' SP exercise could be undertaken.

2.6.2 Monitoring

The intercept survey fieldwork was monitored through spot checks and supervisor accompaniments.

Refusals, out-of-scopes and drop-outs (i.e. those who did not complete the interview for one reason or another) were recorded, and the number of recruits was monitored so that the response rate could be measured.

The overall proportion of intercept interviews that were completed was 71%, with 15% out of scope, 11% refusals and 2% dropping out.

The proportions by mode are shown in **Table 2.5**.

Table 2.5: Intercepts: proportions of interviews, refusals, out-of-scopes and drop-outs

	Total %	Bus %	Rail RP %	Rail SP %	Car %	'Other PT' %
Interviews	71	72	70	79	62	76
Out-of-scopes	15	6	21	13	19	7
Refusals	11	19	7	6	16	14
Drop-outs	2	3	2	2	3	2
Sample size	**39,475**	**3,757**	**9,849**	**9,993**	**10,403**	**5,462**

There were much higher proportions of out-of-scopes for the rail RP and car SP intercept interviews than for the other modes. This was because of the additional screening for RP (i.e. the rail traveller had to be aware of alternative operator costs and times) and car (the traveller had to be travelling on a specific route). Refusals were highest for users of bus and car. Refusals were higher for bus, as a larger proportion of bus travellers compared to other modes did not have internet access and were unwilling to give their phone numbers.

The average numbers of interviews, refusals, out-of-scopes and drop-outs per shift are shown in **Table 2.6**.

Table 2.6: Average number of interviews, refusals, out-of-scopes and drop-outs per intercept shift

	Total %	Bus %	Rail RP %	Rail SP %	Car %	'Other PT' %
Interviews	32	37	30	34	27	38
Out-of-scopes	7	3	9	6	8	4
Refusals	5	10	3	3	7	7
Drop-outs	1	2	1	1	1	1
Base (shifts)	**881.5**	**73**	**231**	**229**	**239.5**	**109**

The numbers of refusals are within the range of what one would expect in surveys of this nature, and we do not believe that this has had significant implications for the findings.

The recruitment mode and purpose was checked against completed interview mode and purpose on a daily basis to assess whether shift targets should be adjusted.

The number of intercept interviews per shift was low for some modes and particularly so for the employees' business and commuting quotas. To meet the targets the following measures were taken in the course of the fieldwork:

- Undertaking circa 100 additional intercept shifts.
- Extending the length of many shifts from six hours to seven or eight hours.
- Increasing the frequency of reminders to the on-line recruits.

- Increasing the number of reminders and the size of the incentives for car and employees' business recruits (to £20). This affected 3% of questionnaires. Those offered £20 incentives were more likely to be making employees' business trips and using car or rail than the rest of the sample[12].

2.6.3 SP recruitment issues and mitigations

There were some reallocations of shifts if, for example, services were not running (as was the case at Beeston Centre and Clifton South in Nottingham and Wolverhampton St. Georges for the 'other PT' sample).

Because many 'other PT' locations such as Hucknall and Bulwell (which were selected so that there was a rail alternative mode) were outside the city centre and relatively quiet outside the peak, there were also very low levels of employee business travel. It was agreed to change some 'other PT' shifts to more central locations, and especially to London, where employees' business travel was more prevalent.

One complication was that the trip purpose collected at the intercept interview did not always match the purpose subsequently given when the respondent completed the survey. This was particularly an issue for those who were recruited as employees' business, but recorded another purpose when completing the main questionnaire. This was not a problem for the subsequent modelling but for the fieldwork, because it meant that additional respondents had to be recruited in order to meet the targets.

To address the discrepancy between recruits and interviews for employees' business, a reminder of the purpose mentioned at recruitment was added at the beginning of the main questionnaire during the fieldwork period. This was then made bold face and repeated as a further mitigation.

Nonetheless, there was a large discrepancy between recruits and interviews for employees' business.

For the main SP questionnaire, this could have been partially caused by the NTS-based purpose questions, which asked for:

- the origin location;
- the nature of the origin location; and
- the reason for being at that location if it was not the home or workplace.

A similar set of questions was also asked for the destination. This was potentially confusing, with workplace potentially being interpreted as a client location. Some respondents coded an employee business purpose as 'other', rather than using the two codes which were designed to identify employees' business within the questionnaire.

To help address these issues the following changes/additions were made during the fieldwork:

[12] There is of course the possibility that the increased incentive could have introduced bias, but the small size of the affected sub-sample made it impossible for us to test for this (and indeed suggests that any such bias would have been minor).

- Changed 'workplace' at Q2 and Q7 with 'normal place of work'.
- Added the following prompt after Q7 and Q8 (if employees' business codes not selected):

  "You indicated that you made an employer's business journey when you were initially contacted but your answers do not describe such a trip.
  If you were making an employer's business trip can you please press the back arrow and change your previous response(s)."

However, this change from the purpose mentioned at recruitment to the purpose recorded in the survey also occurred for the RP questionnaire, wherein the purpose question was more straightforward. It would therefore appear that this problem was introduced – at least to some extent – at the recruitment stage. All recruiters were instructed to show the screen when asking the purpose question, as this included the following definitions:

- Commuting (to from work/place of education)
- Travelling in the course of your employer's business
- Non-work

Following the completion of the intercept survey, but before the end of the fieldwork period, we addressed a shortfall in employees' business and commuting respondents by re-contacting those who undertook the pilot. We asked these individuals if they had made any car or rail employees' business or car commuting trips in the preceding two weeks and, if they had, they were invited to participate in the field survey. This mitigation was agreed with the Department.

2.6.4 RP recruitment issues and mitigations

Analysis of initial RP responses showed a relatively high level of 'non-trading', which was a situation we were trying to avoid. To cite an example, a number of travellers on Virgin Trains reported that Virgin Trains had both a lower cost and faster time than the alternative of Chiltern Trains. Whilst respondents' perceptions may, in this regard, have been erroneous, it meant that there was no time vs. cost trade-off which could be modelled.

Recruiters were asked to check that the respondent understood the relevant screening questions (e.g. *"Could you have saved money getting to London at this time of day by using London Midland?"* and *"If you had chosen to get to London quicker by using Virgin would it have cost you more at this time of day?"*) and to code 'don't knows' as 'no' rather than 'yes'.

Further interim analysis then showed a small reduction in the proportion of 'non-traders'.

A small follow-up survey was undertaken with those who were identified as non-traders. There were 44 responses. When showed the costs they entered for the train they caught and competing operator(s), 18 said that their responses were correct. When these respondents were asked why their answer contradicted that given at recruitment (e.g. *"Why did your answers in the questionnaire show a lower price for travelling by Virgin Trains when you had indicated it was higher at recruitment?"*) their responses were:

- I misunderstood the question at recruitment 4
- I did not say that at recruitment 2
- I found out later that Virgin Trains/East Coast[13] ticket was cheaper 5
- I made a mistake when entering the values 3
- Other 4

The remaining 26 who said their responses were wrong were invited to correct those responses. For 17, their revised values still showed a lack of trading.

Even with careful design, it is difficult to eliminate non-trading altogether. As we will see when analysing the RP data (**Section 5.3**), instances of non-trading must however be 'cleaned' from the modelling dataset, and the measures outlined above should therefore be seen as an attempt to minimise such data loss.

2.7 General public telephone recruitment method

For the general public telephone sample, RDD[14] sample was purchased that geographically represented the population of England as shown in the 2011 Census by region. The objective was to collect a broadly geographically representative sample based on regions. The overall target quotas, specified by mode and trip purpose and reflecting the overall all sample targets, are shown in **Table 2.7.**

Table 2.7: Target quotas

Mode	Employees' business	Commute	Other non-work
Car	200	200	200
Bus	N/A	100	100
Rail	200	200	200

Adult respondents were contacted and screened using a recruitment questionnaire and, if in-scope, invited to participate in the research either on-line or through an interviewer administered telephone interview. The former were sent a web-link to the customised survey using the same e-mail invite as for the intercept survey.

For those who undertook the whole interview by phone, the SP options for the three exercises were sent. These were customised based on answers to various questions within the questionnaire. Therefore, only after these questions were

[13] Virgin Trains for the Birmingham to London route and East Coast for the Peterborough to London route

[14] Random Digit Dialling. RDD sample is created by selecting a known existing number, and randomising the last couple of digits to generate a new telephone number that may or may not exist, and may or may not be a residential number. This is checked using a pulsing machine to dial the resulting phone numbers, and tell at the exchange whether the number is valid or not.

The list of phone numbers that remains is checked against a listed business sample. If a telephone number matches that of a business, then it is discarded. Generally speaking, what remains is a list of working residential phone numbers. The main advantage of RDD is that all households in a given geographical area are given equal opportunity to participate in the research. This is particularly important in areas such as London where the proportion of ex-directory numbers is very high.

asked (about a third of the way through the interview), could the customised SP options be generated. These were then either:

- E-mailed to the respondent so that they could continue with the interview, looking at the options on a computer (65% of interviews).[15]
- E-mailed to the respondent and an appointment made to complete the interview at a later time. The respondent could then refer to the options on a computer screen or print them out (33% of interviews).
- Printed out and posted to the respondent and an appointment made to complete the interview at a later date and time. The respondent could then refer to the hard copies of the options (2% of interviews).

The telephone fieldwork was undertaken between 14:00 and 21:00 Monday to Friday, between 10:00 and 18:00 Saturday, and between 11:00 and 19:00 Sunday, to help ensure that those in employment could be recruited.

The average interview length was 29 minutes.

The number of telephone recruits by mode and purpose is shown in **Table 2.8** below.

Table 2.8: Telephone recruitment interviews by mode and purpose

Experiment type	Employees' business	Commute	Other non-work	Total
Car SP	309	284	290	883
Bus SP	N/A	42	53	95
Rail SP	149	25	74	248
Total	**458**	**351**	**417**	**1226**

31,960 recruitment calls were made and the main call outcomes were:

- Recruited 6%
- No reply/answerphone 50%
- Refusal 31%
- Number not recognised/fax/business etc. 8%
- Not in scope 4%

2.7.1 Recruitment issues and mitigations

The telephone recruitment method worked well. At the beginning of the fieldwork period, it became apparent that bus concessionary passholders should be excluded from telephone recruitment, since only passholders travelling on routes where there was a non-zero cost rail alternative were in-scope.

[15] Some respondents asked to have the SP options read out to them over the phone as they did not wish to wait for the email or for them to be posted to them.

2.8 Incentives

All participants were offered a £10 incentive (an Amazon or Boots voucher or a donation to a charity) on completion of the main questionnaire. Towards the end of the fieldwork period, some were offered a £20 incentive to help meet certain quotas. In total, 3% of the general public sample and 25% of the employers sample received a £20 incentive. For employers these were more likely to be rail and from larger companies as these were the quota groups that were being targeted.

2.9 Dates

Fieldwork took place between 24th October and 15th December 2014. The latter date was a 'hard' deadline agreed with the Department *a priori*, so as to avoid conducting survey work during the Christmas and New Year period, when travel behaviour might be atypical.

2.10 Final questionnaires

Specimens of the main stage questionnaires are included in **Appendix E** for the following elements:

- General public intercept recruitment
- General public telephone recruitment
- Employers' business telephone recruitment
- General public SP on-line/CATI main questionnaire
- General public RP on-line main questionnaire
- Employers' business main SP questionnaire

2.11 Overall sample targets

With reference to the earlier discussion in **Chapter 0,** the originally commissioned SP sample sizes for car, rail and 'other PT' were 1,500 each. Given the decision to administer the 3-game rather than 2-game format (see **Section 2.12.1** for a detailed commentary), the car and rail SP sample sizes were subsequently doubled, to ensure that these samples would deliver robust valuations across the segments and covariates of interest.

Following the commissioning of this additional data collection, the target samples for each survey were as given in the following tables: **Table 2.9** for the general public SP; **Table 2.10** for the general public RP; and **Table 2.11** for the employer SP.

Table 2.9: Target general public Stated Preference interviews by mode and purpose

Mode	Employees' business	Commute	Other non-work	Total
Car	1,000	1,000	1,000	3,000
Bus	N/A	500	500	1,000
Rail	1,000	1,000	1,000	3,000
'Other PT'	500	500	500	1,500
Totals	**2,500**	**3,000**	**3,000**	**8,500**

Table 2.10: Target general public Revealed Preference interviews by purpose

Mode	Employees' business	Commute and other non-work	Total
Rail	1,250	1,250	2,500

Table 2.11: Target employers Stated Preference interviews by mode*

Mode	Car	Rail	'Other PT'	Total
Rail	133	133	133	400

*mode of a typical employee's trip

2.12 Key design issues

2.12.1 2 vs 3 game format

With a few exceptions (namely walk and cycle and bus concessions); respondents received all three games (i.e. SP1, SP2 and SP3). In order to mitigate for 'order effects'[16], SP1 was presented initially, whilst the order of SP2 and SP3 was randomised.

This format emerged over the course of Phase 1 of the study, having deliberated whether it would be preferable to present all three games to each respondent, or only two of the three (specifically SP1, plus SP2 or SP3). Whilst the two-game format would arguably moderate cognitive burden, the decision to present three games was influenced by the following considerations:

- First, there was the issue of comparability. The two-game format would collect values of reliability or quality from a given respondent, but not both. This would have limited our ability to compare VTT in these two contexts (i.e. in the presence of reliability or quality attributes). It would also have prevented us from ascertaining whether any differences in valuations of reliability and quality were caused by underlying differences in those valuations, or by differences in the two groups of respondents supplying those valuations.
- Second, in order to estimate meaningful and robust values, it was judged that it would be advantageous to maximise the volume of data from games with more than two parameters (i.e. SP2 and SP3). However, the two-game format

[16] The ordering of the SP games 1, 2 and 3 in the questionnaire.

would deliver a dataset comprising around 50% SP1 observations, and 25% each of SP2 and SP3 observations.

- Third, the *a priori* expectation (which was subsequently confirmed in the results) was that respondent behaviour in SP1, which is the most abstract, would be most affected by design effects across the three games. This was a further reason for maximising the number of SP2 and SP3 observations, relative to SP1.

The downside of the 3-game format was the additional cognitive burden imposed on the respondent. In order to mitigate this, the number of repetitions of the games presented to any given respondent was moderated. The implication of this recommendation was that, as mentioned in the previous section, the Department opted to commission a degree of additional data collection. This allowed the 3-game format to be implemented whilst not compromising the overall size of the dataset.

2.12.2 Stated Preference (SP) experiments for car, bus, rail and ‘other PT’

This section describes the principal experimental games which were developed for car, bus, rail and ‘other PT’.

SP1: Time vs. cost

Time vs. cost represented the ‘base’ experiment; this was the focus of the 2003 VTT work, and the additional interests of the present study (e.g. reliability, crowding etc.) can be rationalised as incremental additions to the ‘base’. The critical requirement of the ‘base’ experiment was to deliver a robust average value of travel time saving across the sample and suitably weighted across the population. Having estimated an average value, the intention was to explore key variations in values associated with the traveller and the trip.

Against this background, it was important to ensure that the base experiment would be ‘unpolluted’ by modal (including socio-economic-demographic features of typical mode users) and quality attributes (such as the prevalence of crowding/congestion). The base experiment therefore offered a binary choice between unlabelled alternatives, where the alternatives were represented in terms of time and cost differences only (and the preamble instructed the respondent to imagine ‘free-flow’ conditions for car and ‘uncrowded’ conditions for PT).

Nevertheless, we acknowledge that it is impossible to know in advance how respondents interpret travel time in an SP1-style experiment, i.e. whether they regard it as free-flow time, congested time, etc. This was a motivation for our extensive work on combining data across games and studying the differences in valuations. We will return to this issue when interpreting our estimates of VTT in **Chapter 7**.

Whilst the base experiment omitted explicit consideration of other variables which could be relevant to the respondent’s current trip (e.g. interchange), background questions in the survey elicited such information, including whether the respondent was on the outward/return leg when intercepted, and how many interchanges were involved.

The base experiment was structured around a single leg of that trip, and the respondent was advised that A and B were identical in terms of other trip features including interchange. Values for travel time and costs (i.e. SP1) were 'pivoted' around reference values for the actual trip, with percentage changes obtained from the statistical design (**Figure 2.4**).

Figure 2.4: Time vs. cost experiment (car example)

	Option A	Option B
One way fuel cost	£33.30	£35.00
One way travel time by car	4 hours 23 minutes	3 hours 30 minutes
	Option A	Option B

We deliberately did not state the respondent's reference time/cost in the SP presentation, so as to avoid (actively) inducing reference effects.

Pivots were **not** applied to the reliability (SP2) and quality (SP3) variables, on the grounds that there was no obvious reference point for the pivots. The detailed statistical design for the experiment (which is discussed in more detail in **Section 2.12.3**) has been informed by the following considerations:

- The number of choice tasks issued to each respondent.
- The likely ranges of travel time and cost, given the reference trip.
- Prior valuations of travel time savings.
- The levels employed for changes in travel time and cost from the reference trip, taking into account both the size and sign of time savings.
- Any practical/operational constraints on the presented combinations of attributes (e.g. smaller percentage changes are generally used on longer trips).

Whilst the time vs. cost experiment outlined above was employed in a consistent fashion across most modes within scope, the following exception should be noted. In the case of concessionary bus travel, the passenger travels at zero cost. For this reason, non-concessionary bus passengers were assigned the time vs. cost game (which we refer to as SP1a in the case of bus), but concessionary bus passengers were assigned a mode choice game instead (SP1b). The latter was framed around the practical situation – common in many areas of England – where concessionary passengers can travel by bus for free, or by rail for a nominal fare (**Figure 2.5**).

Figure 2.5: Mode choice experiment for bus concessionary

	BUS	RAIL
One way fare	Free	£1.50
One way in vehicle travel time	1 hour 6 minutes	21 minutes
Frequency	Every 30 minutes	Every 60 minutes
	Option A	Option B

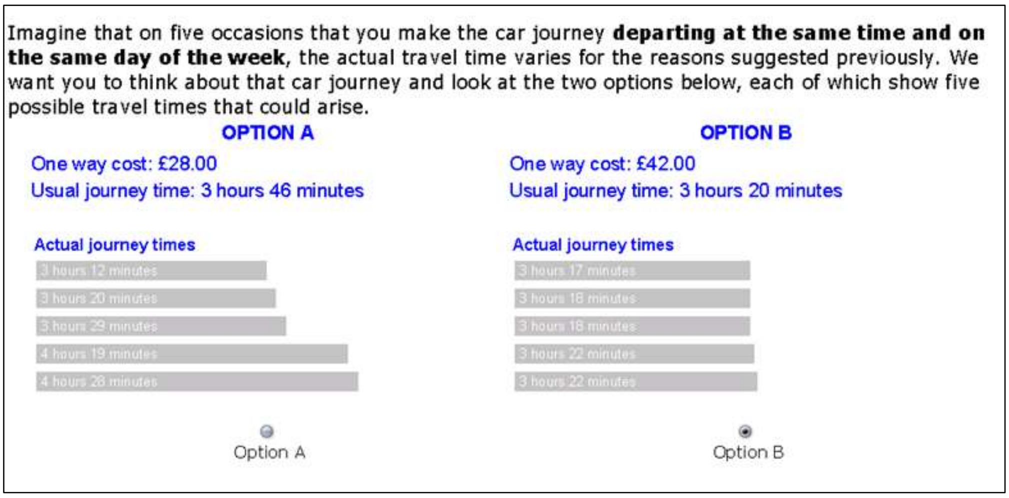
SP2: Time vs. cost vs. reliability

A joint value of travel time (VTT) / value of travel time reliability (VTTR) experiment calls for the enhancement of the VTT experiment to include variability in travel time. Like the base VTT experiment, the objective was to ensure that estimates of VTT would be unpolluted by other attributes (such as headway) which could, in practice, interact with VTTR. Reliability is a challenging concept for SP, and it requires good design to deal with this effect independently, never mind in combination with other effects.

There continues to be debate in the literature concerning the most appropriate SP presentation of reliability, and this was therefore a topic of our qualitative research (discussed in **Section 2.13** to follow). Having tested a number of alternative presentations, the qualitative research identified a preference for a variant on the 'Hollander'[17] presentation (**Figure 2.6**), with two qualifications:

- For car, reliability was specified in terms of the mean and standard deviation of travel time, thereby permitting elicitation of the 'reliability ratio'.
- For PT, reliability was specified in terms of the timetabled travel time, mean lateness relative to the timetable, and the standard deviation of lateness, thereby permitting elicitation of the 'lateness multiplier' and potentially also the PT interpretation of the 'reliability ratio' (i.e. in terms of lateness rather than travel time).

Figure 2.6: Time vs. cost vs. reliability experiment (car example)



Informed by findings from the qualitative research, the preamble advised respondents that unreliability presented in the SP was associated with unpredictable (e.g. breakdowns and accidents on road, or staff shortages on rail, etc.) rather than predictable (e.g. trips taking longer in rush hour, or fast/slow trains, etc.) variations in trip time. The preamble also advised that the five trips presented in the SP departed at the same time and on the same day of the week, thereby eliminating rescheduling of the trip as mitigation for unreliability.

We acknowledge that, according to the microeconomic theory underpinning VTTR, a relevant consideration is the probability that each of the five outcomes

[17] Hollander, Y. (2006) 'Direct versus indirect models for the effects of unreliability'. Transportation Research Part A: Policy and Practice, 40 (9), pp699-711.

under options A and B will arise. However, to simplify an experiment which our respondents already found complex, and consistent with previous (successful) applications of the Hollander presentation in the UK, we deliberately omitted outcome probabilities from the presentation. We accept that this is a pragmatic restriction on the generality of the analysis.

SP3: Time vs. cost vs. quality (i.e. crowding/traffic conditions)

Quality variables might be seen as those that reflect intrinsically different types of time, which can then have different valuations, or else specific travel conditions that lead to different travel time values because in-vehicle time (or indeed any other form of time) is essentially different in nature. In principle, if the 'base' experiment (SP1) adequately represents the value of reliable and uncongested/uncrowded time savings, then the value of quality attributes (SP3) can be conceptualised as multiples of VTT (i.e. multiples of 'base' time vs. cost valuations).

The ITT for this study did not prescribe the specific quality attributes that should be covered. Several such attributes could in principle be examined, but given the finite data sample at our disposal and the complexity of potential inter-relationships between quality attributes, we were constrained as to how many dimensions could sensibly be analysed. Having consulted participants at a Technical Workshop on 3rd June 2014, and discussed with the Department, the following key quality attributes were agreed for the mechanised modes:

- Car: time spent in different traffic conditions (**Figure 2.7**).
- Rail: two quality games were developed, the first looking at time spent in different crowding conditions (**Figure 2.8**), and the second looking at frequency (**Figure 2.9**).
- Bus: two quality games were again developed, the first covering time spent in different traffic conditions (including stopping) and frequency, and the second covering time spent in different crowding conditions (**Figure 2.10**).
- 'Other PT': two games were developed; the first looking at time spent in different crowding conditions was designed in the same manner as for bus (**Figure 2.11**); and the second looking at frequency (**Figure 2.12**).

For each of the above quality attributes, the SP presentations sought to comply with industry best practice for analysing the relevant attribute. It should however be acknowledged that the scope of the present study was ambitious – covering values of travel time savings, reliability, and several dimensions of quality – and this necessitated compromise in some instances. For example, in the case of car traffic, the six types of time employed in the M6 Toll study were consolidated into three types for the current study. Similarly, in the case of rail, the current study employed a simplified version of MVA's (2008) crowding presentation[18].

The second game for 'other PT' (SP3b) considered crowding, and was designed in exactly the same manner as bus.

[18] MVA Consultancy (2008) 'Valuation of Overcrowding on Rail Services'. Prepared for Department for Transport.

Figure 2.7: Time vs. cost vs. quality experiment for car

	Option A	Option B
One way travel cost	£37.60	£42.00
Traffic conditions	1 hour 45 minutes in heavy traffic 11 minutes in light traffic 2 hours 53 minutes in free flowing traffic	2 hours 11 minutes in heavy traffic 2 hours 36 minutes in light traffic 57 minutes in free flowing traffic

Option A   Option B

Figure 2.8: Time vs. cost vs. quality experiment for rail (SP3a)

	East Coast	First Capital Connect
One way train fare	£27.50	£13.50
One way train time	45 minutes	1 hour 23 minutes
Frequency	Every 15 minutes	Every 60 minutes

Figure 2.9: Time vs. cost vs. quality experiment for rail

	Option A	Option B
One way travel time	3 hours 54 minutes	3 hours 18 minutes
One way travel fare	£18.00	£24.00
Crowding level when you boarded	 Seated, 100% of seats occupied, eight people stood around each door	 Standing, 100% of seats occupied, one person stood around door

Option A   Option B

Figure 2.10: Time vs. cost vs. quality experiment for bus (SP3a)

	Option A	Option B
Time travelling at normal speed Time slowing down or in congested traffic Time bus spends at bus stops	16 minutes 5 minutes 7 minutes	8 minutes 5 minutes 7 minutes
One way fare	£2.20	£1.25
Frequency	Every 30 minutes	Every 60 minutes

Option A   Option B

Figure 2.11: Time vs. cost vs. quality experiment for bus (SP3b)

	Option A	Option B
One way travel time	16 minutes	26 minutes
One way travel fare	£0.95	£1.45
Crowding level when you boarded	standing but with discomfort	standing but reasonably comfortable
	Option A	Option B

Figure 2.12: Time vs. cost vs. quality experiment for 'other PT' (SP3a)

	METRO	BUS
One way travel fare	£1.85	£1.50
One way travel time	33 minutes	30 minutes
Frequency	Every 5 minutes	Every 30 minutes
	Option A	Option B

2.12.3 SP design approach

This section gives an overview of the process used to generate the experimental designs which determine the specific combinations of attributes that a respondent was faced with in the stated choice scenarios. We start by giving an overview of the design process used before explaining the procedure used to develop different designs for different respondents. Finally, we explain the inputs to different designs and how the design outputs were then used to calculate the values presented to a given respondent.

Experimental design approach

The field of experimental design has seen substantial developments over recent years, with academic work almost completely moving away from the use of orthogonal design techniques which have been almost standard for the previous two decades. A key shortcoming of orthogonal techniques is that they use no prior information about the likely sensitivities of respondents to the individual attributes. However, in almost all contexts, we have strong *a priori* expectations about the sign of coefficients and generally also the rough relative sensitivities.

Making use of this prior information means that this new generation of designs, known as 'efficient' designs (Rose and Bliemer, 2014a)[19], present respondents with more meaningful trade-offs that increase the information content in the data. In practice, this leads to substantially lower standard errors of the estimated parameters, as compared to parameters estimated using orthogonal designs. The other benefit of efficient designs is that they can be generated in such a way as to avoid scenarios with dominated alternatives; these scenarios would produce no

[19] Rose, J.M. and Bliemer, M.C.J. (2014a) 'Stated choice experimental design theory: The who, the what and the why', in Hess, S. and Daly, A. (ed.) 'Handbook of Choice Modelling', Edward Elgar, Cheltenham, pp152-177.

information of use to an analyst and in the worst case could lead to a respondent losing interest.

Efficient designs are optimised on the basis of prior coefficient values selected by the analyst. This means that the asymptotic variance covariance matrix (and thus the standard errors) are minimised in models estimated on the resulting data if the coefficient values estimated from the real data are equal to those (i.e. the priors) used in optimising the design. If the coefficients that are estimated from the data are different from those used in optimising the design, then these values will still be recovered, but the standard errors will be higher. However, they are generally still much lower than for orthogonal designs. In fact, orthogonal designs are only ever equivalent to *efficient* designs in the case where all parameter values are zero.

The designs produced for this study follow the state-of-the-art in the field, making use of Bayesian D-efficient designs. D-efficiency is a specific error measure used in optimising the designs, as discussed at length by Rose and Bliemer (2014b)[20]. D-efficient designs result in reliable and unbiased parameter estimates when used wisely.

Clearly, the approach requires the analyst to optimise the designs for an assumed choice model, and in this case we optimised for an additive multinomial logit model. Every choice model, whether additive or multiplicative, aims to map part-worths and utilities to probabilities. Efficient designs ensure that choice tasks make sense and that attribute levels are such that sufficient trade-offs are made (e.g. between time and cost). Since the behaviour of the respondent is independent of the model estimated, we would argue that a choice task that sufficiently trades off attributes will be efficient in both an additive or multiplicative choice model.

McFadden (1973)[21] proved that the maximum likelihood estimates of the multinomial logit model are consistent, which means asymptotically unbiased. This assumes that the underlying assumptions of the logit model hold, namely that respondents exhibit compensatory behaviour and make trade-offs in each choice task, and that the attribute levels are orthogonal (uncorrelated) to the error terms. Strictly dominant alternatives in a choice task would result in non-compensatory choice behaviour and should therefore be avoided in experimental designs. In this study, we have achieved this by removing all choice tasks with strictly dominant alternatives based on a regret measure, see Bliemer *et al* (2014)[22].

Fosgerau (2014)[23] has demonstrated that biased parameters in mixed logit models can result when correlating the range of the attribute levels to the parameter priors of individuals. In our D-efficient designs, we have not imposed any direct correlation between the parameter priors and attribute levels; the attribute levels in D-efficient designs are chosen such that they minimise the volume of the

---

[20] Rose, J.M., and Bliemer, M.C.J. (2014b) 'Survey artefacts in stated choice experiments. Proceedings of the International Conference on Transport Survey Methods', Leura, Australia.

[21] McFadden, D. (1973) 'Conditional logit analysis of qualitative choice behaviour'. In Zarembka, P. (ed) 'Frontiers in Econometrics'. Academic Press, New York.

[22] Bliemer, M.C.J., Rose, J.M. and Chorus, C. (2014) 'Dominance in stated choice surveys and its impact on scale in discrete choice models'. Proceedings of the International Conference on Transport Survey Methods, Leura, Australia.

[23] Fosgerau, M. (2014) 'Manipulating a stated choice experiment'. Forthcoming.

asymptotic variance-covariance matrix and as such optimise the reliability of the parameter estimates without biasing the results.

Furthermore, it is known that scale may be different across choice tasks due to large differences in choice task complexity, and this can introduce the complication of needing to address this during estimation; see Rose and Bliemer (2014a). Random and orthogonal designs in particular suffer from this, while D-efficient designs typically result in choice tasks with more similar choice task complexity (as measured by entropy), such that scale issues are usually small.

Practitioners have often been concerned about the impact that the assumptions about priors may have on the final design. And indeed, D-efficient designs based on fixed local priors are known to be less robust against mis-specification; see for example Walker *et al* (2015)[24]. This issue can be easily resolved by using Bayesian priors, as we have done in the present study. Bayesian D-efficient designs provide a good balance between efficiency and robustness. We adopted a Bayesian procedure in our design process, such that our designs are robust against a fairly wide range of mis-specifications.

By making use of Bayesian designs, we have allowed for uncertainty in the priors, thus optimising the designs for a broader range of values for the coefficients. In particular, we worked with wide regions, using normally-distributed priors, with a standard deviation that is 50% of the mean value. The mean values in turn were based on an extensive review of values obtained in past studies.

Specific details for individual designs are given below, in terms of any additional constraints imposed on the combinations that were presented to respondents. Following the generation of designs, an extensive testing process was carried out in which choices were simulated with these designs for a wide range of coefficient values and where models were then estimated to establish that these values could be retrieved from the data. This approach ensured that the design process used did not in any way bias the coefficient values in a given direction.

Despite these provisions, we cannot rule out the possibility of bias due to mis-specification of the model, since respondents may not behave according to the model assumptions. However, we are confident that we do not introduce additional biases by choosing a D-efficient design approach. Our designs have pre-specified levels that do not depend on the choices the respondents make (they depend on distance classes for example, but not on respondent behaviour), thereby avoiding bias induced by the design.

Alternatives to D-efficient designs are random designs and orthogonal designs, or indeed manual designs. The first two approaches typically suffer from choice tasks with dominant alternatives, which are inconsistent with the logit model assumptions of compensatory behaviour, and as such may lead to biased coefficients. Using D-efficient designs with priors we are able to automatically remove such problematic choice tasks. Furthermore, orthogonal designs in unlabelled experiments are actually D-efficient designs under the assumption that the coefficients are all equal to zero. An attraction of the third approach, as used in the previous UK and Scandinavian studies, is that manual designs can avoid

[24] Walker, J.L., Wang, Y., Thorhauge, M., and Ben-Akiva, M. (2015) 'D-efficient or deficient? A robustness analysis of SP experimental designs in a VOT estimation context'. Proceedings of the Annual Meeting of the Transportation Research Board, Washington DC, USA.

meaningless choice tasks. But by doing so, manual designs also make use of prior beliefs on coefficient signs and values. The difference with efficient designs is that prior beliefs are used in a more structured and automatic way.

Using priors that are closer to the true values, we can always improve upon an orthogonal design. However, priors in D-efficient designs need to be chosen wisely; in order to prevent any deterioration due to misspecification, we have used conservative priors. This means that if from other studies a (positive) coefficient was *b*, then we used a prior between 0 and *b*. A simple conservative rule of thumb is to choose 0.5**b*. This in practice ensures that the D-efficient design will be more efficient than an orthogonal design. Finally, in order to make the design more robust against misspecification, we have adopted a Bayesian D-efficient design approach in which the prior is assumed to be randomly distributed around the conservative point estimate.

Different designs across purposes and reference trips

While it is clear that different designs are needed for different games (e.g. separate design for time vs. cost and for time vs. cost vs. reliability trade-offs), it is important to recognise that an efficient design is optimised for the specific values of attributes and priors used in the design. This had two separate dimensions in the present context, relating to trip purpose and core trip characteristics.

First, as substantial differences in values of time (and other valuations) were expected to exist between business and non-business travellers, separate designs were produced for these two purpose segments for those games where business travellers were a sampling target. In essence, this means that the trade-offs presented to business travellers were geared towards their likely higher willingness-to-pay, thereby giving us more robust estimates in the analysis.

Second, the surveys presented respondents with trips framed around a recent trip they had made – especially in relation to current travel time and cost. An optimal design for a given reference trip can be expected to be a function of these characteristics, and simply using a generic design (in terms of percentage variations to be applied to trip characteristics) across all types of trips would lead to a major loss of efficiency.

For this reason, separate designs were produced for a set of representative trips. Each respondent was then given a design based on the trip closest to their reference trip (in terms of the smallest percentage difference between the reference values for the design and that respondent’s values for time and cost), with percentage variations (or pivots) applied to the specific reference trip for that person, where these pivots were obtained from the design.

The number of reference trips used varied by purpose, with the lowest number for bus (2) and the highest number for rail (20). Each game presented a respondent with five separate choice scenarios. The actual designs made use of a number of rows that was larger than the number of tasks assigned to a single respondent, to ensure sufficient richness in the variations in the data. As an example, for car SP1, the range of boundary value of times across the data ranged from £0.15/hour to £372/hr. This was of course partly a result of some very cheap and very expensive reference trips, but even when looking at the reference trips used in the design process, the range went from £0.45/hr to £90/hr. The overall design was then split into a number of distinct blocks at the design stage, minimising correlation between attributes and blocks, and each block was used as closely as possible a uniform number of times across the sample of respondents. The number of rows for designs was set to 25 after extensive testing. In total, 315 designs were produced for this study.

The subsequent discussion details the inputs used in all designs and explains how the design outputs were used to compute the values presented to respondents. We also look at any additional constraints imposed on the designs, where it should be noted that, by default, the design approach already avoided scenarios in which one alternative is dominated, e.g. there is no possibility in the simple time vs. money trade-offs that one option is both faster and cheaper than the other.

Car games

- No additional constraints were imposed on SP1 given the aforementioned avoidance of dominance.
- For SP2, and for reasons of realism, we excluded cases where either the shortest travel time or the highest travel time was combined with the highest level for travel time variability.
- For SP3, we imposed additional constraints which guaranteed that the implied distances for the two trips differed by no more than 25%, again for reasons of realism (assuming that light traffic speed would be 80% of free-flow, and heavy traffic speed would be 60% of free-flow). The design allowed for both increases and decreases around reference values, and the design process sought to achieve attribute level balance, i.e. guaranteeing that increases were as likely as decreases. We gave some flexibility in this process, i.e. allowing the share of increases and decreases to be different from 50-50 to allow the constraints on the total sum to be met.
- For all games, we defined an adjusted free-flow reference time (AFF) on the basis of the current free-flow time (CFF) and current total time (CTT) as:

  $$AFF = CFF + \max(0,-(CFF+\min(\Delta FF)*CTT)))$$

  Here, min (ΔFF) was the smallest (i.e. most negative) additive percentage shift used for free-flow time across both alternatives and across all five tasks for the respondents. This means that the adjusted free-flow was shifted upwards for all tasks and all alternatives in those cases where any of the alternatives in any of the tasks would have required censoring. We then used the definition:

  $$FF = AFF + \Delta FF * CTT$$

  to compute the values to be presented.

Rail games

- No additional constraints were imposed on any of the designs.

Bus games

- No additional constraints were imposed on SP1 or SP2.
- For SP3a, we imposed additional constraints which guaranteed that the implied distances for the two trips differed by no more than a third, again for reasons of realism (assuming that slowed down time speed would be 50% of free-flow).

- No additional constraints were imposed on SP3b.

'Other PT' games

- No additional constraints were imposed on any of the designs.

2.12.4 RP design approach

Whilst SP methods have advanced considerably since the last national UK study of VTT in 2003, there continues to be debate as to whether SP can elicit credible valuations of travel time savings and reliability. This challenge is especially vocal in the area of business travel, given that respondents might not act as agents for their employer's best interests. It therefore makes sense to validate SP against robust RP data.

We believe that the rail market offers not only the best such possibilities but also the opportunity to collect large amounts of data cost-effectively. We considered the scope for conducting a similar exercise for car, but there are very limited opportunities in the UK for doing this; our view is that an RP exercise for car would be difficult and costly, with no guarantee of generating usable results. For these reasons, it was agreed with the Department that the present study would concentrate RP analysis on rail.

The narrower focus of the RP was dictated by two further considerations. First, the employers' business survey was structured around typical (not actual) trips conducted by their employees, and not readily amenable to RP. Second, the possibility of examining reliability and crowding through RP was explored with the Department, but deemed to be out of scope.

Corresponding to the rail SP3a experiment (see **Section 2.12.2**), we designed an RP survey around rail operator choices in carefully selected locations. That is to say, the RP survey collected time, cost and headway data for the current trip (i.e. the chosen operator), as well as for the alternative trip (i.e. the alternative operator).

Following consultation with the Department and relevant stakeholders, we selected the following O-D trips as the basis for the RP experiment:

- Birmingham to London – this provides a rich trade-off context given three operators and a range of different tickets across the day. We reasoned that this O-D would yield large samples, with sensible choices for business travellers but few commuters.
- West Coast flows from Stoke to London – this provides a different set of trade-offs, particularly attractive in the leisure market. We reasoned that this O-D would provide some longer and shorter distances than Birmingham, especially if intermediate flows were considered (e.g. Rugby).
- Peterborough to London – this has small time differences but a large number of fares and crowding/reliability variations. We reasoned that Peterborough would be primarily of use for the commuting market, but could also supplement leisure travellers if surveys were undertaken sequentially.

Additional background analysis using ticket sales data

Ticket sales data can provide insights into whether given O-Ds actually have patterns of demand that are consistent with rail travellers choosing between operators. To facilitate such analysis, the Department provided LENNON ticket sales data for three specific flows. From this data, the volume shares of different tickets were examined in order to determine the possible extent of operator choice.

- *Stoke to London*
  - Tickets are either for London Midland, Virgin Trains or the inter-available Any Permitted. The latter are likely to use Virgin Trains given that this service is markedly quicker than London Midland. Similarly, business travellers who use full fare (inter-available) tickets will most likely use Virgin on speed grounds. However, business travellers increasingly take advantage of reduced fare tickets. Looking at the tickets that are operator specific (%Any=0), Virgin Trains were found to capture the highest but not entirely dominant share for first class, where business travel would be expected to be prevalent, and there was close competition in the reduced market, where leisure travel would be expected to be prevalent (in combination with some business).
  - This is not a commuter market. It was concluded that business and particularly leisure travellers are confronted with real time-cost trade-offs and no single operator dominates.
- *Rugby to London*
  - The Rugby to London flow, compared to the Stoke to London flow, additionally include services provided by London Midland and Virgin Trains originating in the West Midlands. The ticket shares were found to be broadly similar to the Stoke to London flows and again seemed to provide a fertile environment for RP sampling, with the more useful trade-offs perhaps existing for leisure travel.
- *Birmingham to London*
  - This flow involves competition between Chiltern Trains, London Midland and Virgin Trains, which offer a range of time-cost trade-offs. The Any Permitted ticket is most likely to involve use of Virgin Trains, but not entirely, since Chiltern can be quicker to some London destinations. London Midland captures a significant enough proportion of the first class market, which tends to be business-oriented, to indicate that meaningful RP analysis of this ticket type would be possible. The season ticket market is small, but Chiltern Trains can compete with the Any Permitted ticket which is likely to be Virgin. The standard full market is not particularly large and is dominated by the Any Permitted ticket.
  - In the standard reduced market, which is by far the largest, there seemed to be strong competition across operators. In the standard advance purchase market, where inter-available tickets are not present, there was again evidence of strong competition. These markets will contain some business travellers.

- The evidence indicated that there was strong competition between operators, and that this context would provide a firm basis for the development of robust RP models.

2.13 Developing and testing the approach

The rest of this chapter discusses the methods which were employed to develop the SP and RP surveys during Phase 1 of the study, as follows:

- Qualitative research principally to help inform the SP design for employers' business and reliability (**Section 2.13.1**).
- Cognitive depths to test the flow, comprehension of the questions and SP exercises and the wording of the surveys (**Section 2.13.2**).
- The piloting approach (**Section 2.13.3**).

2.13.1 Qualitative research

Sixteen pre-tasked focus groups and ten depth interviews with large employers were conducted between 25 June and 15 July 2014 in order to address the following specific areas of the ITT:

- How do companies determine, monitor and enforce travel policies, and to what extent do employees follow these?
- How is travel time used for productive ends?
- How can travel time reliability best be presented?
- How do we offer realistic variations in car costs, and indeed how do we treat the issue of occupancy which is so often neglected in SP studies of motorists?

Figure 2.13: Structure and location of focus groups

Employer Groups

Business size:	medium	medium	small	small
Mode:	rail	car	car	rail
Location:	Edinburgh	London	London	Manchester

Employee Groups

Position:	higher managerial	higher managerial	lower managerial	lower managerial	lower managerial	trades people
Mode:	car	rail	car	car	rail	car/van
Location:	Manchester	London	Cardiff	Edinburgh	Ipswich	Birmingham

Self Employed Groups

Type:	white collar	blue collar
Location:	Manchester	London

Non Work Groups

Age:	younger	older	younger	older
SEG:	ABC1	ABC1	C2DE	C2DE
Mode:	car	car	car	rail
Location:	Cardiff	Plymouth	Birmingham	London

The structure and locations of the groups are shown in **Figure 2.13** above.

All group participants were asked to complete a pre-task in order to aid their understanding of the concept of VTT. The pre-task comprised a simple workbook designed to get participants thinking about how time was used and how much they valued it. For work and non-work trips, participants were required to provide examples of when they may have bought time savings in the past. Employers and employees were asked to provide any business-related examples of where the organisation would be willing to pay for savings when travelling.

The business depth interviews were undertaken with those from large organisations with over 250 employees, and were structured to obtain the views of those from a range of business sectors, specifically:

- Four public sector
- Six private sector

The findings from the qualitative research were used to inform the SP and RP survey designs. The following sections set out the main findings and recommendations from the qualitative research.

How different segments value travel time

An overview of how the four different segments typically value savings in travel time is given in **Figure 2.14.**

Figure 2.14: Overview of how different segments value time savings

Non-work (i.e. commuting plus other non-work travel)	These respondents were found to be highly cost-conscious, but the need to simplify their busy lifestyles could create an appetite for buying time savings. Such decisions were affected by the respondent's income, as well as by features of the trip such as travel time.
Self-employed business	These respondents craved a much better work/life balance. However their ability to time-shift, combined with their fear of passing on costs and subsequently losing work, meant that they had a limited appetite for buying time savings.
Employers' business (SME and larger)	Employers claimed that they would always look at efficiencies including travel-related time savings, even though it was not necessarily their top priority. The principal consideration was employee productivity.
Employees' business	Unlike employers, employees value time savings for both personal and business benefits. Some employees were prepared to buy time savings in order to enhance their personal lives. They were not convinced that employers would pay for travel time savings simply to improve staff well-being.

Insights on the SP experiments

Referring back to the SP presentations from **Section 2.12.2**, the testing of different formats revealed the following findings:

- In order for respondents to make a realistic choice, they needed to be provided with relevant contextual information. This reinforced the importance of grounding the SP preamble in terms of the respondent's personal experience.

- Modal choice SPs can compromise the purity of VTT estimates. It was felt that mode based experiments included too many variables; for example, innate mode preference and complex ticketing considerations. This finding reinforced the decision to employ mode choice games only in specific parts of the brief (e.g. concessionary travel).
- Car travellers found the notion of higher costs for faster routes difficult to comprehend, unless tolls were introduced. This was one of the reasons for our decision to present SP2 and SP3 to every respondent, in case the data quality for SP1 was affected by hypothetical bias in any way.
- Some of the terminology was considered to be mode-specific. For example, *"one-way travel cost"* and *"single leg journey"* were considered to be vocabulary usually reserved for train journeys. It was recommended that the most appropriate terminology for each mode should be employed in the field surveys.
- On balance, it was found that pictorial as opposed to textual versions of the SP presentation for reliability worked best. Various pictorial presentations were also tested, and in this regard the use of the five horizontal bars (to represent five occurrences of the same trip) and a common colour throughout was found to work best.
- For the bus quality experiment, the use of time in different conditions (e.g. free-flow, slowing down in congested traffic and at bus stops, etc.) was well understood and it was recommended that this format should be taken forward to the field surveys.
- For employers' business, the very wide range of potential employee trips meant that the most practical approach was for the SP experiment to be framed around a 'typical' employee trip.

2.13.2 Cognitive depths[25]

The final stage of methodological development was to subject the SP surveys to cognitive testing. This section sets out the findings from 26 cognitive depth interviews undertaken in the course of the pilot survey (**Table 2.12**). There were two stages of cognitive depths: the first 10 were undertaken in advance of the pilot and the second 16 during the pilot.

Table 2.12: Cognitive depths structure: Stages 1 and 2

Questionnaire type	Stage 1 (i.e. pre-pilot)	Stage 2 (i.e. during pilot)
Main general public SP questionnaire covering car, bus and rail	10	10
Employers' business SP questionnaire	-	6

[25] Cognitive depth is a type of interviewing technique which focuses on respondents' thought processes in answering survey questions and uses specialised techniques such as thinking aloud, probing, observation and paraphrasing.

The main findings from the cognitive depths are summarised below. Many of the recommendations relate to detailed changes to the questions and these were used to amend the wording of the pilot questionnaires for Wave 2 of the pilot survey.

General public SP questionnaire

For both waves of the pilot, the questionnaire worked well in terms of the flow and routing. Respondents on the whole said that they enjoyed completing the questionnaire, and they found the SP exercises 'easy' or 'fairly easy' to understand. Some respondents experienced problems with the Google maps tool used to identify the origin and destination of the journey (and calculate the distance travelled), and remedial actions were taken to resolve these problems for Wave 2 of the pilot survey.

Another comment was that the SP preambles lacked explanation of why the research was being conducted. Whilst this omission was to some extent deliberate, all SP preambles were reviewed and refined before Wave 2.

In general, respondents found the background travel questions clear and did not cite difficulties in answering key questions about the trip of interest, such as the travel time and cost.

Specific findings on the three SP exercises are summarised below.

- *SP1: Time vs. cost*

In the first stage of cognitive testing, the SP preamble was felt to be 'wordy'. Respondents liked the use of bullet points. Typically, but not always, respondents had to read the introduction several times in order to fully digest its meaning. In the second stage of cognitive testing, a revised introduction to the SP was felt to work well, although some participants requested a title to the introductory screen. Titles were therefore added to all the SP introductory screens for Wave 2 of the pilot.

In both stages of cognitive testing, the SP exercise itself was felt to be simple and clear. The options were felt to be realistic.

- *SP2: Time vs. cost vs. reliability*

In the first stage, the SP preamble was felt by some participants to be overly long, although others struggled to understand the concept of reliability from the preamble. Some thought the introduction to be too 'abstract'. It was also suggested that consistent language regarding travel time should be used (e.g. in Wave 1, the preamble said *"one way travel time"*, but the example SP exercise said *"journey time"*). The preamble was amended accordingly for Wave 2.

In the second stage of the cognitive depths (following changes), the revised preamble was felt to work well overall although, in a similar vein to SP1, there were requests for a title to be added to the introductory screen. Some respondents suggested that the introduction should be shown over more than one screen, with the example SP shown on a separate screen. These changes were made for Wave 2 of the pilot survey.

Relative to SP1, this was considered to be a more complex SP exercise. Overall, most respondents were able to consider the scenarios presented, although it was evident that some did not give equal attention to all of the information. For example, some focussed more on the usual travel times rather than the actual travel times.

In the first stage of cognitive testing, some respondents felt that text shown in a grey box was not distinct enough for people with visual impairments. In the second stage, the concept of being early or late for car was questioned. These elements were refined for the Wave 2 pilot.

- *SP3: Time vs. cost vs. crowding/traffic conditions*

In the first stage, the SP preamble was considered to be too long and too detailed by some participants. In general, however, this preamble was considered easier to understand than for SP2. The preamble was amended to be clearer for the Wave 2 pilot. As for SP1 and SP2, a heading was suggested in both stages of testing, and this was added for Wave 2.

Overall, respondents understood the SP exercise and were able to make informed choices. Indeed, for some respondents, SP3 was felt to be more realistic than SP1 as it included a 'comfort' factor.

Employers' business SP questionnaire

The concept of 'briefcase' travel was understood by respondents, although some respondents struggled to recount how many staff at each grade worked in their respective organisations.

- *SP1: Time vs. cost*

In the case of employers who were being asked to cite a typical trip undertaken by one of their employees, it was suggested that the respondent could be reminded of the typical trip before each SP. The preamble was considered rather wordy by some, although most expressed no difficulty in understanding the SP.

- *SP2: Time vs. cost vs. reliability*

The concept of reliability was found to be difficult to understand by some and there was a suggestion that presenting options A and B in separate boxes would make it clearer. For rail, some thought that the early arrival times were unrealistic, although others found the exercise straightforward and realistic. For car there was a request for more background information.

- *SP3 Car: Time vs. cost vs. traffic conditions*

This exercise included more information which was perceived to improve clarity, although some respondents queried whether each option was the 'same' trip.

- *SP3b Rail/Underground: Time vs. cost vs. crowding*

Respondents queried whether they should assume that a seat had been booked for the rail trip. They also remarked upon the relatively complexity of the experiment, and that levels of crowding were in some cases unrealistically high. For Wave 2, we responded to these points by adding contextual information to the SP preambles and questions.

2.13.3 Piloting approach

This section sets out the method and findings for two waves of pilots for the following surveys:

- General public Stated Preference (SP) research
- Revealed Preference (RP) research
- Employers' business SP research

The method for the pilot surveys was intended to be as faithful as possible to that which would be subsequently rolled out for the field survey. As it transpired, one exception to this was the general public telephone survey, where the first birthday rule (i.e. the adult member of the household with the next birthday) used in the pilot survey to randomly select a respondent was subsequently dispensed with for the field survey.

Pilot target sample sizes

For each of two waves of the pilot survey, the target sample size was 550 SP with general public and 90 SP with employers' business (**Table 2.13**). In the case of employers' business, there were also quotas on company size, industry grouping and region. In addition, there was a target of 100 RP for each wave of the pilot, comprising 50 employees' business and 50 commute or other non-work.

Table 2.13: Summary of target pilot sample sizes for field surveys

		Trip Purpose			
		Employees' business	**Employers' business**	**Commute**	**Other non-work**
Mode	**Car**	50 SP*	30 SP	50 SP	50 SP
	Bus	N/A	N/A	50 SP	50 SP
	Rail	50 SP*/50RP	30 SP	50 SP/25RP	50 SP/25RP
	'Other PT'	50 SP	30 SP	50 SP	50 SP

Note: * of which 5 self-employed

Intercept locations

The intercept locations agreed with the Department for the pilots were a subset of the proposed field main survey locations as shown in **Appendix B**. In summary, 67 shifts were booked in the first wave and 74 in the second wave of the pilot, as shown in **Table 2.14**:

Table 2.14: Pilot shift numbers by mode and wave

Mode	Wave 1	Wave 2
Bus	9	12
Car	18	19
'Other PT'	15	15
Rail	25	28

At some non-motorway service stations, where permission could not be secured, car drivers were recruited on-street near to car parks.

Incentives

All fieldwork with the general public was incentivised to encourage participation. £10 incentives (Boots or Amazon voucher or a donation to charity) were provided to everyone that completed a survey, whether on-line or by telephone.

For the employer survey, £10 incentives were introduced towards the end of the first pilot and retained for the second pilot.

2.13.4 Pilots

This section describes the pilot testing of the SP and RP questionnaires described in the previous sections. The pilot surveys were administered in two waves each lasting around two weeks, specifically:

- Wave 1: 11-22 August 2014
- Wave 2: 10-22 September 2014

Overall, the intercept and telephone recruitment method for both pilots worked as planned.

2.13.5 Wave 1 recruitment

For the Wave 1 intercept survey, 2,588 respondents were recruited. The numbers of intercept recruits by mode and purpose are shown in **Table 2.15**.

Table 2.15: Wave 1 intercept recruits by mode and purpose (row percentages)

		Trip Purpose			
		Employees' business	**Commute**	**Other non-work**	**Total**
Mode	**Car**	91 (18%)	169 (34%)	232 (47%)	**492**
	Bus	N/A	167 (46%)	194 (54%)	**361**
	Rail	145 (14%)	344 (34%)	533 (52%)	**1,022**
	'Other PT'	35 (5%)	373 (52%)	305 (43%)	**713**
	Totals	**271**	**1,053**	**1,264**	**2,588**

It should be noted that the some of the rail respondents (i.e. those on routes where there was operator choice) were recruited for the RP survey.

78% of respondents gave usable e-mail addresses and 11% gave usable telephone numbers.

Those who gave usable email addresses were sent emails, although 11% were 'out-of-office' or bounced back as the e-mail address was not recognised. 36% of those sent e-mails entered the survey and 32% completed it. From the 303 'telephone' recruits, 48 interviews were undertaken.

The number of SP interviews achieved was 676 against a target of 550. **Table 2.16** shows the numbers by mode and purpose.

The overall response rate (i.e. proportion of interviews to recruits) was 27%, slightly ahead of the 25% expected, although it should be noted that the employees' business response rate was much lower than the other two segments.

Table 2.16: Pilot SP interviews by mode and purpose

		Trip Purpose		
		Employees' business	**Commute**	**Other non-work**
Mode	**Car**	15 (target=50)	40 (target=50)	73 (target=50)
	Bus	N/A	54 (target=50)	57 (target=50)
	Rail	29 (target=50)	76 (target=50)	119 (target=50)
	'Other PT'	13 (target=50)	111 (target=50)	89 (target=50)
	Totals	**57**	**281**	**338**

The response rate by mode was:

- Car 26%
- Bus 31%
- Rail 28%
- 'Other PT' 30%

The response rate by purpose was:

- Employees' business 21%
- Commute 30%
- Other non-work 34%

Although quotas were not achieved for certain categories, this does not impact on the Wave 1 findings as part of the purpose of the pilot was to understand the different responses by mode and purpose.

2.13.6 Wave 1 intercept recruitment issues

Overall, the intercept approach worked well. The main issues were:

- Difficulty in recruiting for employees' business, particularly for 'other PT', mainly because of the August pilot survey period and low throughput of business travellers on 'other PT'.
- Difficulty in gaining permissions to interview at non-motorway petrol stations. If permission was refused, recruitment interviews were undertaken on-street instead.

In order to help achieve quotas, the following changes were implemented for the Wave 2 pilot intercept survey:

- Shifts were rescheduled from those where recruitment numbers were high (i.e. 'other PT') to ones were recruitment numbers were low (i.e. car).
- There was an increased focus at recruitment on targeting travellers on employees' business.

For the telephone recruitment approach, the first birthday rule proved to be restrictive. Since this restriction turned away potential recruits, it was recommended that the first birthday rule should be dropped for Wave 2. Because of the use of quotas, this had no impact on the nature of the sample, but reduced the number of calls that were made.

Of the 676 SP interviews, 606 (90%) were on-line and 70 (10%) were undertaken by telephone.

The average questionnaire length for the on-line research was consistent with expectations: 24 minutes for the on-line survey. For the telephone research the average questionnaire length was 30 minutes. This was longer than the 25 minutes planned. The additional questionnaire length was absorbed for the main field survey.

Given the 'pivot' approach to the SP design, travel costs were required as an input to the customised SP. For the telephone research, however, probing for estimates was found to be difficult. For the main field survey, the cost questions were revised to make the collection of this information easier.

2.13.7 SP Wave 1 pilot survey with employers

Overall, 74 interviews were undertaken. The target of 90 interviews was not achieved. It proved very difficult to complete interviews and the main reasons for this were:

- The fieldwork was undertaken in August when many staff were on leave.
- Difficulties in finding in-scope respondents.

Because of difficulties recruiting in-scope respondents, some modifications to the approach were made on Friday 22nd August (i.e. during Wave 1):

- Because of the large number of respondents who reported that there was no company travel policy, interviewers stopped mentioning this at the beginning of the interview as respondents could be in-scope even if there was no formal policy. This amendment to the survey was made for Wave 2.
- Initially, participation in the employer survey was not incentivised, but it was found that the £10 incentive helped to reduce refusals; this change was made for Wave 2.

Six respondents withdrew once they had started the questionnaire, as they found that they were unable answer questions on employee travel.

2.13.8 Employer survey key findings

Overall, the employer survey questionnaire flowed well. The average questionnaire length was 25 minutes.

One problem with the questionnaire was that even when respondents were the relevant decision makers in terms of company travel, they often had little idea about the details of employee trips, and therefore found this aspect of the questionnaire difficult to answer. There were occasions where employers really did not know the costs for rail or London Underground that their employees paid, and as travel costs were required for the SP experiment, the SP could not therefore be undertaken. Some interviewers suggested that it would be helpful to have illustrative fares for rail (based on distance) and for London Underground.

To mitigate this problem, the Wave 2 recruitment questionnaire asked respondents whether they would be comfortable speaking about trip specifics (such as costs and times) for the mode selected. If not, the mode could be changed. In addition, interviewers were provided with illustrative fares.

2.13.9 RP Wave 1 pilot survey

55 rather than 100 interviews were undertaken, mainly because too many of the Birmingham-London segment were sent invitations to complete the SP survey rather than the RP (hence over-sampling for rail SP). This problem was rectified for Wave 2. Overall, the questionnaire worked well. The average questionnaire length was 14 minutes.

2.13.10 Wave 2 pilot recruitment

For the Wave 2 intercept survey, 2,815 respondents were recruited. The numbers of intercept recruits by mode and purpose are shown in **Table 2.17**.

Table 2.17: Wave 2 intercept recruits by mode and purpose (row percentages)

		Trip Purpose			
		Employees' business	Commute	Other non-work	Total
Mode	**Car**	26%	31%	42%	**661**
	Bus	N/A	58%	42%	**382**
	Rail	35%	28%	37%	**1092**
	'Other PT'	17%	53%	30%	**680**
	Totals	**24%**	**39%**	**37%**	**2815**

It should be noted that the some of the rail respondents (i.e. those on routes where there was operator choice) were recruited for the RP survey.

The overall recruitment target of 2,200 was met. 89% of recruited respondents gave usable e-mail addresses and 11% gave usable telephone numbers. Those who gave usable e-mail addresses were sent e-mails. 10% were out-of-office, or bounced back as the e-mail address was not recognised. 35% of those sent e-mails entered the survey and 32% completed it. 52% were sent automatic reminders (scheduled 6 days after the initial invite).

From the 497 'telephone' recruits, 96 interviews were undertaken.

The number of SP interviews achieved was 755 against a target of 550. **Table 2.18** shows the numbers by mode and purpose.

The overall response rate (i.e. proportion of interviews to recruits) was ahead of expectation (31%) although, as for Wave 1, the response rate for employees' business was lower than expected.

Table 2.18: Pilot SP interviews by mode and purpose

		Trip Purpose		
		Employees' business	**Commute**	**Other non-work**
Mode	**Car**	47 (target=50)	62 (target=50)	93 (target=50)
	Bus	N/A	57 (target=50)	50 (target=50)
	Rail	51 (target=50)	68 (target=50)	101 (target=50)
	'Other PT'	20 (target=50)	111 (target=50)	95 (target=50)
	Totals	**118**	**298**	**339**

The response rate by mode was:

- Car 29%
- Bus 28%
- Rail 30%
- 'Other PT' 33%

The response rate by purpose was:

- Employees' business 22%
- Commute 35%
- Other non-work 33%

2.13.11 SP Wave 2 pilot survey with general public: car, bus, rail and 'other PT'

Of the 755 SP interviews with the general public using car, bus, rail and 'other PT', 86% were on-line and 14% were undertaken by telephone.

The average questionnaire length was 26 minutes for the on-line research and 31 minutes for the telephone research. Both are longer than the 25 minutes planned but not considered to be problematic.

2.13.12 SP Wave 2 pilot survey with employers

Overall, 90 interviews were undertaken against a target of 90 interviews.

Company size quotas were met except for the largest businesses (26 compared to a target of 30).

The changes made after Wave 1 proved to be effective in terms of recruitment.

In general, the questionnaire flowed well. The average questionnaire length was 31 minutes, rather longer than the 25 minutes planned. The additional questionnaire length was absorbed in the main field survey.

2.13.13 RP Wave 2 pilot survey

112 out of the target of 100 interviews were undertaken. Overall, the questionnaire worked well. The average questionnaire length was 16 minutes.

2.13.14 Modelling of the pilot SP and RP data

Following the pilot surveys, the project timescales allowed a brief opportunity to model the data collected. The objective of this exercise was to seek to confirm the robustness of the SP and RP survey designs, rather than to report valuations of time, reliability and quality at any level of definitiveness.

To this end, we developed separate SP models by mode and game (i.e. SP1, SP2 and SP3) on the raw (i.e. uncleaned) data. The RP dataset was modest in size and did not lend itself to modelling at the pilot stage; our assessment of its robustness was therefore based on more general considerations such as the prevalence of 'non-trading'. Two types of model formulation were estimated on the SP data. First, an additive formulation, estimated in 'utility space', was reported for all SP games. This is a simple model formulation, which has been widely used in previous VTT studies. Second, a multiplicative formulation, estimated in 'log willingness-to-pay (log-WTP) space', was reported for SP1 only. The latter formulation represents current best-practice in VTT studies, and was taken forward to the modelling of the field survey; detailed explanation of this model is provided in **Chapter 4.**

Whilst fuller discussion of the pilot modelling was reported at the end of Phase 1 of the project, for present purposes it will suffice to simply summarise the conclusions documented in that report.

SP findings

- The non-working time study was judged to be fit for purpose in design and analysis terms, and could be expected to produce results of at least the robustness of the 1994 study.
- The relativities between employee/commuting/other non-work and between time/reliability/quality behaved broadly as expected in the pilot survey.
- With some minor exceptions, research on the 'core' modes (car, rail, bus) was judged to have proceeded well, and there was no reason to make substantive changes for Phase 2. One of these exceptions was concessionary bus, where the pilot survey did not deliver an estimate of VTT. We recommended that this problem should be addressed through better targeting of the survey work on locations where there was a genuine choice between bus and train.
- Research on 'other PT' gave some grounds for confidence, whilst also leaving some residual uncertainty. A particular issue was the correct recording of the trip cost where travellers pre-paid using travel cards (e.g. Oyster). For the field survey, we refined the questionnaire so that it automatically calculated a unit trip cost based on the weekly/monthly/annual travel card cost, and we then asked the respondent to check that the calculation was realistic.
- The employees' business SP seemed to work well in its own terms, producing values somewhat above commuting values. We recognised however that challenges awaited us in terms of interpreting the values, and implementing

them within an overall social valuation of employers' business travel time savings.

- Turning to the employers' business SP, there were difficulties in administering Wave 1 of the pilot, but substantive changes were made to Wave 2, and the success of the latter wave of the pilot gave grounds for confidence that employer valuations could be estimated. That said, there was a clear difference in the success of the employer SP on rail and car, relative to 'other PT'.
- The data collection approach was successful in collecting the required quotas across the sampling frame. Inspection of background data concerning the respondents indicated a good range of variation in features of both the trip and respondent.
- The 3-game SP format (i.e. time, reliability and quality) worked well, and there was little evidence of excessive cognitive burden on respondents.

RP findings

- Based on inspection of the pilot data, our understanding of these rail markets, and examination of ticket sales data, we concluded that there was good evidence that the RP component of this study was worth pursuing. However, it was important to ensure that the RP data collected would be as robust as possible and, to this end, the Phase 1 pilots identified a number of improvements to maximise the potential of the RP approach.
- In particular, it was found that around a third of the pilot sample did not face a time vs. cost trade-off; they were using a quicker **and** cheaper operator. For Phase 2 of the study, we therefore introduced additional screening to ensure that the survey recruited both traders and those familiar with the times and costs of alternative operators.
- We anticipated that Birmingham (New St, Moor St and Snow Hill) to London (Euston, Marylebone) would provide the richest data, whilst Stoke, Stafford, Nuneaton and Rugby would also provide a wide range of time-cost trade-offs. We recommended that Peterborough should be used primarily for the commuting market.
- Some amendments to the questionnaire were identified, for purposes of examining familiarity and information acquisition. For example, we recommended that some of the contextual business travel questions should be based around the actual time differences between operators.

3 Market Research Findings

3.1 Overview

This chapter describes the market research findings for the main field survey. This focuses on the findings which are most relevant for the SP and RP analyses, such as time, cost, frequency and variability. In addition, it describes the key characteristics of the samples. Where relevant, the data is compared to the National Travel Survey (NTS), so as to highlight similarities/differences between the SP and RP samples collected here and the 'representative' sample which is the basis of the NTS.

The achieved sample sizes are provided in **Section 3.2** with the results covered in three sections:

- **Section 3.3**: General public SP non-work market research results
- **Section 3.4**: Employees' and employers' business SP market research results
- **Section 3.5**: RP market research results

It should be noted that this chapter reports on the data for the whole sample. The SP and RP analyses reported in subsequent chapters are based on 'cleaned' data (as described in **Section 4.2** and **Section 5.3.1,** respectively).

3.2 Achieved sample sizes

3.2.1 General public SP

With reference to **Table 3.1**, 8,623 SP interviews were undertaken with the general public (target =8,500).

Table 3.1: Total SP interviews (on-line and CATI) by mode and purpose (targets in parentheses)

	Employees' business		Commute		Other non-work		Total	
Car	(1,000)	956	(1,000)	1,032	(1,000)	1,037	(3,000)	3,025
Bus		*	(500)	371	(500)	672	(1,000)	1,043
Rail	(1,000)	1,010	(1,000)	998	(1,000)	1,128	(3,000)	3,136
'Other PT'	(500)	265*	(500)	614	(500)	540	(1,500)	1,419
Totals	**(2,500)**	**2,231**	**(3,000)**	**3,015**	**(3,000)**	**3,377**	**(8,500)**	**8,623**

*Includes 22 bus

84% of the SP interviews were undertaken on-line and 16% by telephone. 45% of SP interviews were completed within a day of recruitment and a further 33% 2 to 7 days after recruitment.

89% of the car, rail and bus SP interviews were intercept recruited and 11% CATI recruited. 'Other PT' interviews were all intercept recruited. The proportion recruited by phone was rather lower than the 20% target, mainly because bus and rail commute and employees' business respondents were relatively scarce. The car sample was 19% telephone recruited (predominantly commute and non-work). It should be noted that any residual bias in trips/travellers in the sample was corrected at the implementation stage (as discussed in **Chapter 7**).

- Of the intercept recruited, 91% completed the questionnaire on-line and 9% undertook the interview by telephone.
- Of the CATI recruited, 90% completed the questionnaire by telephone and 10% completed the questionnaire on-line.

It can be seen that some of the quotas were exceeded. The reasons for this were:

- The quotas were applied based on the recruitment criteria. However, those who changed their purpose from an ‘open’ quota to a ‘closed’ one would still get through.
- Those who partially completed a questionnaire could return to complete the questionnaire and would not be affected by quotas being closed as these were applied at the beginning of the questionnaire.
- There were two surveys (telephone and on-line), which meant that two sets of quotas were required for each mode/purpose segment, and the combined targets were in excess of the actual target to prevent losing respondents recruited through one or the other method. The balance between the two methods was reviewed every day towards the end of the fieldwork.

3.2.2 General public RP

With reference to **Table 3.2**, 2,646 RP interviews were undertaken with the general public (target =2,500). 43% of RP interviews were completed within a day of recruitment and a further 32% 2 to 7 days after recruitment.

Table 3.2: Total RP interviews (on-line and CATI) by purpose

	Employees’ business		Commute/Other non-work		Total	
Rail RP	(1,250)	1,311	(1,250)	1,335	(2,500)	2,646

All RP interviews were intercept recruited. 98% of RP interviews were undertaken on-line, with the remaining 2% by phone.

3.2.3 Employers’ business SP

With reference to **Table 3.3**, the target of 400 employers’ business interviews was achieved, although there was a shortfall on the largest businesses. We do not believe that this shortfall introduced any distinct bias into the analysis, although it is difficult to be absolutely definitive on this matter.

Table 3.3: Employers' business by mode and number of employees

Mode	**target**	**n**
Car	(194-257)*	244
Train	(130-194)*	143
'Other PT'	*	13
Number of employees		
1-19	(67)	74
20-49	(67)	73
50-249	(133)	149
250+	(133)	104
Total	**(400)**	**400**

* 'Other PT' dropped, remaining interviews split between rail and car (agreed revised minimum for rail of 130)

3.2.4 Response rates

For the intercept recruited respondents, the overall response rate was 37%, comprising:

- Car SP 37%
- Bus SP 35%
- Rail SP 37%
- Rail RP 39%
- 'Other PT' SP 33%

93% of recruits provided e-mail addresses for an on-line survey and 7% a phone number for a follow-up telephone survey. The response rate was the same for both approaches.

For the CATI recruited respondents, the response rate was 61% for those who were in-scope and recruited:

- Car SP 62%
- Bus SP 61%
- Rail SP 60%

3.3 General public SP non-work market research results

3.3.1 Introduction

This section provides some of the market research findings from the general public SP commuter[26] and other non-work samples.

Where appropriate, the data is compared to the National Travel Survey (NTS). The NTS data used for comparison covered trips made by persons aged 16+ by all motorised modes during the years 2010-12. It should be noted that the SP sample by its very nature over-sampled longer distance travellers with, for example, over half the car recruitment shifts being inter-urban and over half the rail recruitment shifts being long distance. As noted earlier in **Section 2.3**, it was fully anticipated that the recruitment approach would introduce some degree of bias into the sample with respect to over-sampling long distance trips and more frequent travellers (as compared with NTS), but it is still instructive to identify key areas of divergence.

In drawing such comparisons, it is important to note that:

a) our survey collected data on some variables which are covered by NTS (thereby permitting comparison), alongside many other variables which are not captured by NTS.

b) the intention throughout was to re-weight values for representativeness (against NTS) at the implementation stage; to this end, the Implementation Tool developed in **Chapter 7** has the functionality to re-weight valuations on a trip or distance basis.

3.3.2 Respondent characteristics

This section describes the age, gender and income of the sample. Data on employment status is provided in **Appendix C**.

Age

With reference to **Table 3.4**, the overall SP commuter sample is a good match to the NTS commuter sample. The other non-work SP sample is fairly similar to the NTS sample except:

- there were more respondents in the 17 to 29 age bands than the NTS;
- there were fewer respondents aged over 70 years old than the NTS.

These differences are mainly because the SP sample excluded non fare payers (except for a few bus concessionary passholders recruited in specific locations), and this meant that many retired respondents with free travel were excluded.

[26] Commuters were defined as those whose origin or destination was normal place of work and they were not on business trips.

Table 3.4: Age: commute and other non-work samples compared to the NTS

Age Band	Commute		Other non-work	
	SP %	NTS %	SP %	NTS %
17-20	7	4	15	5
21-29	23	20	20	11
30-39	26	22	14	17
40-49	21	25	15	21
50-59	18	20	14	16
60-69	6	7	15	16
70+	1	1	7	14

The differences in age band by mode and purpose from the SP sample and NTS are shown in **Table 3.5** and **Table 3.6** respectively.

The car samples had a much older age profile than the other modes. The SP sample for commute matches the NTS data reasonably well. For other non-work, the car and rail SP samples again match the NTS data reasonably well, although the ‘other PT’ and bus SP samples have markedly fewer travellers in the 60+ age bands than the NTS sample because those who paid no fare were excluded from the survey[27].

Table 3.5: Age by mode and purpose from the SP sample

Age Band	Car		Train		Bus		‘Other PT’	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
17-20	3	1	8	19	17	25	8	23
21-29	14	10	25	24	28	20	29	30
30-39	23	13	28	13	20	13	30	17
40-49	25	18	19	13	17	12	18	14
50-59	26	23	15	12	14	9	11	9
60-69	8	24	5	15	4	11	3	4
70+	1	11	*	5	*	10	0	1
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

* = less than 0.5%

[27] Except bus concessionaries at specific locations.

Table 3.6: Age by mode and purpose from the NTS sample

Age Band	Car		Train		Bus		'Other PT'	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
17-20	4	4	4	16	8	12	3	9
21-29	18	11	25	23	31	14	32	26
30-39	21	18	33	17	20	11	32	26
40-49	27	22	21	14	20	10	19	12
50-59	22	16	14	11	14	9	9	7
60-69	8	16	3	12	7	19	4	12
70+	1	13	0	7	1	24	0	9

Gender

With reference to **Table 3.7**, the commuter sample was fairly evenly balanced between male and female, whilst the other non-work sample was predominantly female. The NTS indicates that the 57% of commuter trips are by males and 56% of other non-work trips are by females.

Table 3.7: Gender by mode and purpose

Gender	Car		Train		Bus		'Other PT'	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
Male	55	46	54	38	34	30	47	35
Female	44	54	46	62	66	69	52	65
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

Table 3.8 shows gender by mode and purpose from the NTS. The SP sample tended to have higher proportions of female respondents than the NTS, particularly for non-car modes.

Table 3.8: Gender by mode and purpose from the NTS

Gender	Car		Train		Bus		'Other PT'	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
Male	58	45	64	47	46	39	58	46
Female	42	55	36	53	54	61	42	54

Household income

For commute and other non-work, gross annual household incomes, before deductions for tax and National Insurance, are shown in **Table 3.9**; **Table 3.10** shows the same data by mode/purpose combinations.

There was a fairly even income distribution across the income breaks shown to respondents. As would be expected, commuters had much higher incomes than other non-work respondents: 32% had household incomes of £50,000 per year or more, as compared to 16% for the other non-work sample.

In accordance with *a priori* expectations, the rail sample had the highest household incomes and the bus sample had the lowest household incomes:

- 10% of bus commuters had annual household incomes of £50,000 or over, as compared to 27% for car, 33% for 'other PT' and 43% for rail.
- 5% of bus other non-work travellers had annual household incomes of £50,000 or over, as compared to 16% for car and 'other PT' and 22% for rail.

Table 3.9: Gross annual household income by purpose

Income band	Commute %	Other non-work %
Under £10K	7	19
£10-20K	14	19
£20-30K	16	16
£30-40K	14	12
£40-50K	12	9
£50-75K	16	9
£75-100K	9	4
More than £100K	7	3
Don't know	1	3
Refusal	3	6
Not stated	1	1
Sample size	**2,997**	**3,352**

Table 3.10: Gross annual household income by mode and purpose

Income band	Car		Train		Bus		'Other PT'	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
Under £10K	4	10	5	16	18	33	10	22
£10-20K	14	19	9	15	27	22	14	21
£20-30K	17	19	15	17	18	13	16	15
£30-40K	18	14	12	14	10	7	14	12
£40-50K	13	11	12	9	8	4	11	8
£50-75K	15	10	20	11	8	3	16	8
£75-100K	7	4	13	6	1	1	9	5
More than £100K	5	2	10	5	1	1	8	3
Don't know	2	3	1	2	2	4	1	2
Refusal	5	8	2	3	5	10	2	3
Not stated	*	1	1	1	2	1	1	2
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

* = less than 0.5%

Table 3.11 compares the gross annual household income for the SP modes with that from the NTS[28]. The NTS sample also includes employees' business and this may partly explain why there is more representation in the higher income bands, as compared to the SP sample. The distribution of income by mode is similar between the SP and the NTS samples.

Table 3.11: Gross annual household income by mode: SP sample compared to the NTS

Income band	Car		Train		Bus		'Other PT'	
	SP %	NTS %	SP %	NTS %	SP %	NTS %	SP %	NTS %
Under £10K	7	7	11	7	26	24	16	9
£10-20K	17	16	12	11	25	27	18	13
£20-30K	18	15	16	10	16	15	16	11
£30-40K	16	15	13	12	9	11	13	13
£40-50K	12	12	11	12	6	7	10	11
£50-75K	13	20	16	24	6	9	12	18
£75K+	9	13	17	24	2	6	13	25

It should be noted that in our preferred models on which estimates of VTT are based (**Chapter 4**), we have used household income as the preferred income

[28] Trip based.

variable for commuting and other non-work purposes, and personal income for business travel purpose. This position was informed by empirical tests.

Personal income

Respondents' annual personal income, before tax and other deductions (including allowances or support from other household members, state benefits, etc.), is shown by mode and purpose in **Table 3.12**.

Table 3.12: Annual gross personal income by mode and purpose

Income band	Car		Train		Bus		'Other PT'	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
Under £10K	10	24	14	39	34	54	18	48
£10-20K	24	25	14	20	28	20	25	19
£20-30K	24	19	21	16	19	6	21	13
£30-40K	15	9	17	9	8	3	12	6
£40-50K	8	4	10	4	2	1	7	3
£50-75K	6	3	11	3	*	*	7	2
£75-100K	3	1	4	2	1	*	4	1
More than £100K	2	1	4	1	*	*	3	2
Don't know	1	3	1	2	2	3	1	2
Refusal	6	8	2	3	5	10	2	3
Not stated	1	2	1	1	2	1	1	2
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

* = less than 0.5%

3.3.3 Trip characteristics

This section describes the key characteristics of the trip which was the subject of the SP exercises. It covers:

- Distance, time and cost
- Frequency of trip
- Group size
- Activities undertaken during the trip
- Trip time variability

Data on leg of trip, day of week, time of day of the trip and whether a day trip is provided in **Appendix C**.

Distance, time and cost

The mean in-vehicle time for commuting trips was 54 minutes, and the mean in-vehicle time for other non-work travellers was 1 hour 15 minutes.

Door-to-door road distances were calculated automatically from the origin and destination locations marked on the Google maps within the questionnaire. It should be noted that these distances include access and egress modes, and are therefore expected to be longer than the corresponding NTS distances for ‘other PT’, bus and rail. The door-to-door road distances are shown in **Table 3.13** for the SP sample, and the NTS trip based in-vehicle distances are shown in **Table 3.14**.

The NTS data shows markedly shorter distances than the SP sample, particularly for car, train and bus. As noted in the introduction to this section, the SP sample by its very nature over-sampled longer distance travellers with, for example, over half the car recruitment shifts being inter-urban and over half the rail recruitment shifts being long distance. The values will be re-weighted for representativeness to the NTS at the implementation stage (**Chapter 7**).

Table 3.13: Trip distance from the SP sample by mode and purpose

Distance band	Car		Train		Bus		‘Other PT’	
	Comm-ute	Other non-work	Comm-ute	Other non-work	Comm-ute	Other non-work	Comm-ute	Other non-work
	%	%	%	%	%	%	%	%
0 to 5 miles	29	25	15	12	59	53	40	42
6 to 10 miles	17	13	14	9	22	18	26	25
11 to 20 miles	17	13	27	16	10	11	24	16
21 to 30 miles	9	5	7	6	3	3	3	3
31 to 50 miles	11	8	9	8	2	1	2	3
51 to 100 miles	7	10	10	16	2	6	2	5
100 to 150 miles	3	7	9	16	1	2	0	2
150+ miles	7	19	9	19	1	5	3	4
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

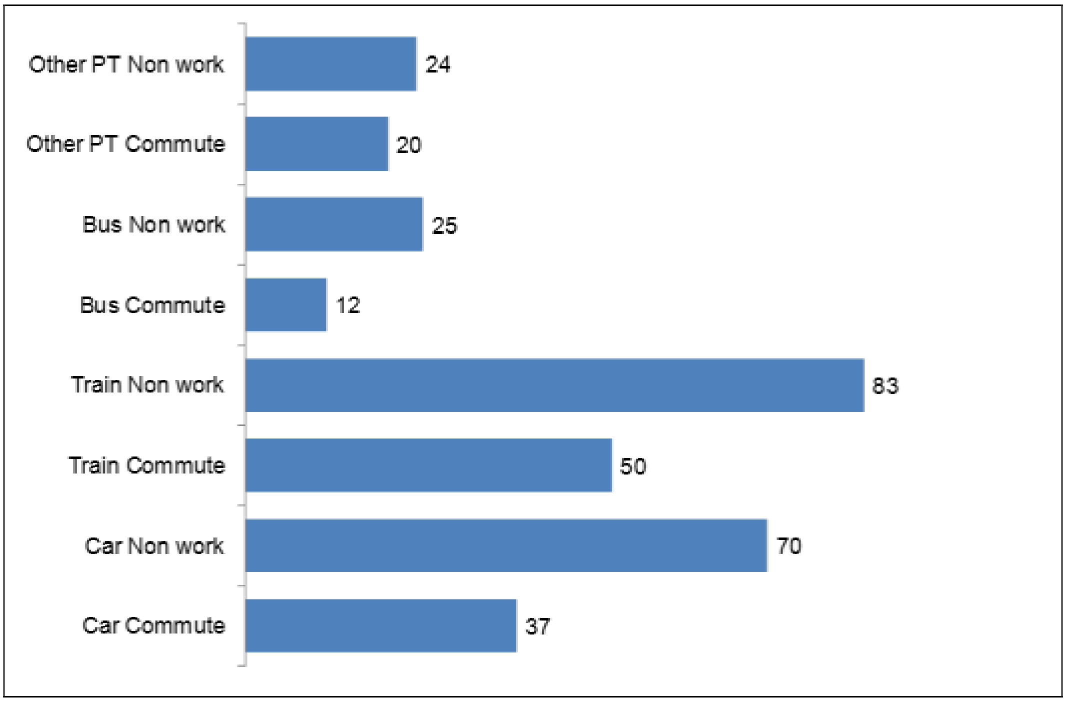
Table 3.14: Trip distance from the NTS sample by mode and purpose

Distance band	Car		Train		Bus		‘Other PT’	
	Comm-ute	Other non-work	Comm-ute	Other non-work	Comm-ute	Other non-work	Comm-ute	Other non-work
	%	%	%	%	%	%	%	%
0 to 5 miles	46	65	9	13	66	77	29	49
6 to 10 miles	24	18	18	18	25	16	40	34
11 to 20 miles	18	10	34	23	8	6	27	16
21 to 30 miles	6	3	15	11	1	1	3	1
31 to 50 miles	4	2	16	15	0	0	0	0
51 to 100 miles	1	1	7	11	0	0	0	0
100 to 150 miles	0	0	1	4	0	0	0	0
150+ miles	0	0	0	5	0	0	0	0
Sample size	**108,070**	**433,868**	**8,720**	**8,116**	**12,841**	**42,094**	**4,750**	**4,507**

The mean door-to-door trip distance for the commuter sample was 35 miles, and for the other non-work sample was 58 miles.

Figure 3.1 shows the mean distance in miles by mode and purpose. The longest distance was for train other non-work, and the shortest was for bus commute.

Figure 3.1: Mean door-to-door distances (miles) by mode and purpose



The in-vehicle times are shown in **Table 3.15** for the SP sample and in **Table 3.16** for the NTS trip based sample. The NTS data shows markedly shorter trip times than the SP sample, particularly for car and train**.**

Table 3.15: In-vehicle times from the SP sample by mode and purpose

Time bands	Car		Train		Bus		'Other PT'	
	Comm -ute	Other non-work	Comm -ute	Other non-work	Comm -ute	Other non-work	Comm -ute	Other non-work
	%	%	%	%	%	%	%	%
0 to 15 minutes	17	17	11	7	24	26	27	29
16 to 30 minutes	26	24	25	18	30	35	36	38
31 to 45 minutes	16	11	18	10	19	16	21	16
46 to 60 minutes	13	8	11	14	16	11	10	8
61 to 90 minutes	12	9	13	13	8	6	5	6
91 to 120 minutes	5	5	12	13	2	2	1	1
121+ minutes	12	27	11	27	1	3	0	2
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

Table 3.16: In-vehicle times from the NTS sample by mode and purpose

Time bands	Car		Train		Bus		'Other PT'	
	Comm-ute	Other non-work	Comm-ute	Other non-work	Comm-ute	Other non-work	Comm-ute	Other non-work
	%	%	%	%	%	%	%	%
0 to 15 minutes	46	65	13	18	20	35	12	25
16 to 30 minutes	35	24	35	29	39	40	35	41
31 to 45 minutes	11	6	25	18	21	14	26	22
46 to 60 minutes	5	2	16	13	12	7	20	9
61 to 90 minutes	2	2	8	10	7	3	6	3
91 to 120 minutes	0	1	2	5	1	1	0	0
121+ minutes	0	1	0	7	0	0	0	0
Sample size	**108,071**	**433,868**	**8,720**	**8,116**	**12,841**	**42,093**	**4,749**	**4,505**

The mean in-vehicle time for commuting trips was 54 minutes, and the mean in-vehicle time for other non-work trips was 1 hour 15 minutes.

For car and rail, the other non-work in-vehicle times were much longer than the commute in-vehicle times:

- Car: 1 hour 4 minutes commute, 1 hour 39 minutes other non-work
- Train: 1 hour 3 minutes commute, 1 hour 35 minutes other non-work

There was little difference between the commute and other non-work in-vehicle times for bus and 'other PT'.

The average in-vehicle times and mean ticket costs for public transport and fuel costs for car are shown in **Table 3.17.**

Table 3.17: Average in-vehicle time and cost by mode and purpose

	Car		Train		Bus		'Other PT'	
	Comm-ute %	Other non-work %	Comm-ute %	Other non-work %	Comm-ute %	Other non-work %	Comm-ute %	Other non-work %
Time (h:mm)	1:04	1:39	1:03	1:35	0:37	0:39	0:31	0:33
Single leg ticket cost* (£.p)	N/A	N/A	15.32	15.43	3.28	2.44	4.42	3.73
Fuel cost (£.p)	6.40	8.79	N/A	N/A	N/A	N/A	N/A	N/A
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

The car sample was asked how they thought about fuel costs for their car. There was strong evidence of cost sensitivity, with 47% saying they chose routes and drove in a fashion so as to keep running costs down, 35% who said they were sensitive to fuel costs at times when they are increasing significantly, and 34% who said they always searched out the lowest cost petrol stations. Only 22% said they only really think about fuel costs for long distance trips, and 22% said that they fill up once or twice a week and did not really consider fuel costs. For the latter group, the responses to the diagnostic question for realism after SP1 were compared to other respondents. This showed almost no difference in the perception of realism of the exercise between the two groupings.

Frequency of trip

As would be expected, the commuter sample made the trip much more frequently than the other non-work sample (**Table 3.18**). However, there are fairly large proportions of infrequent trips made by commuters. Some of these may be genuine infrequent commuters or users of a replacement mode, but some may also be misreporting either purpose or frequency. The bus and 'other PT' samples were the most frequent commuters: 64% bus and 60% 'other PT' made the trip five or more times a week, as compared to 51% for car and 43% for rail.

Table 3.18: Frequency of trip by mode and purpose

Frequency bands	Car		Train		Bus		'Other PT'	
	Commute %	Other non-work %	Commute %	Other non-work %	Commute %	Other non-work %	Commute %	Other non-work %
5+times a week	51	5	43	7	64	17	60	13
3-4 times a week	19	7	19	7	20	17	18	15
1-2 times a week	15	20	15	8	8	31	10	16
1-3 times a month	8	21	10	17	6	15	5	16
Less than once a month	5	34	8	37	1	15	5	23
First time	3	12	5	24	1	5	2	17
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

Group size

Table 3.19: Group size by mode and purpose

Group Size	Car		Train		Bus		'Other PT'	
	Commute %	Other non-work %	Commute %	Other non-work %	Commute %	Other non-work %	Commute %	Other non-work %
Travelled alone	85	49	90	68	92	72	94	71
1 other adult	12	44	7	27	6	24	4	25
2+ other adults	3	7	2	5	2	4	1	4
Children aged 6-17								
1	2	6	1	2	1	6	1	4
2+	1	2	*	1	1	5	*	3
Children aged 5 or under								
1	2	7	*	2	1	8	*	7
2+	*	2	0	*	*	3	*	1
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

* = less than 0.5%

With reference to **Table 3.19**, the car other non-work sample was most likely to travel with others: 51% of car travellers were accompanied, as compared to 32% train, 28% bus and 29% 'other PT'.

Activities undertaken during the trip

Three-quarters of the car sample drove the whole way, 10% shared the driving and 15% were passengers.

Respondents were reminded of their reported one way trip time and asked approximately how much of that time they spend undertaking the following activities (more than one answer could be given):

Work-related activities

- Use laptop/tablet
- Use smart phone/Blackberry/phone
- Other work related to employment

Non-work-related activities

- Driving*
- Talking on phone
- Using smart phone/eBook/tablet/computer
- Reading a book/magazine/newspaper**
- Eating/drinking
- Talking to travelling companions/other travellers
- Listening to music
- Planning things
- Doing nothing/relaxing/looking out of window**
- Other

* car only

** not car

The non-work-related activities were shown first for the other non-work sample.

The amount of time spent on these activities is shown pictorially in **Figure 3.2** and tabulated in greater detail in **Table 3.20**; in the latter table, the three most important activities for each mode/purpose combination are shaded accordingly.

Figure 3.2: Activities undertaken during the trip (average minutes) by purpose[29]

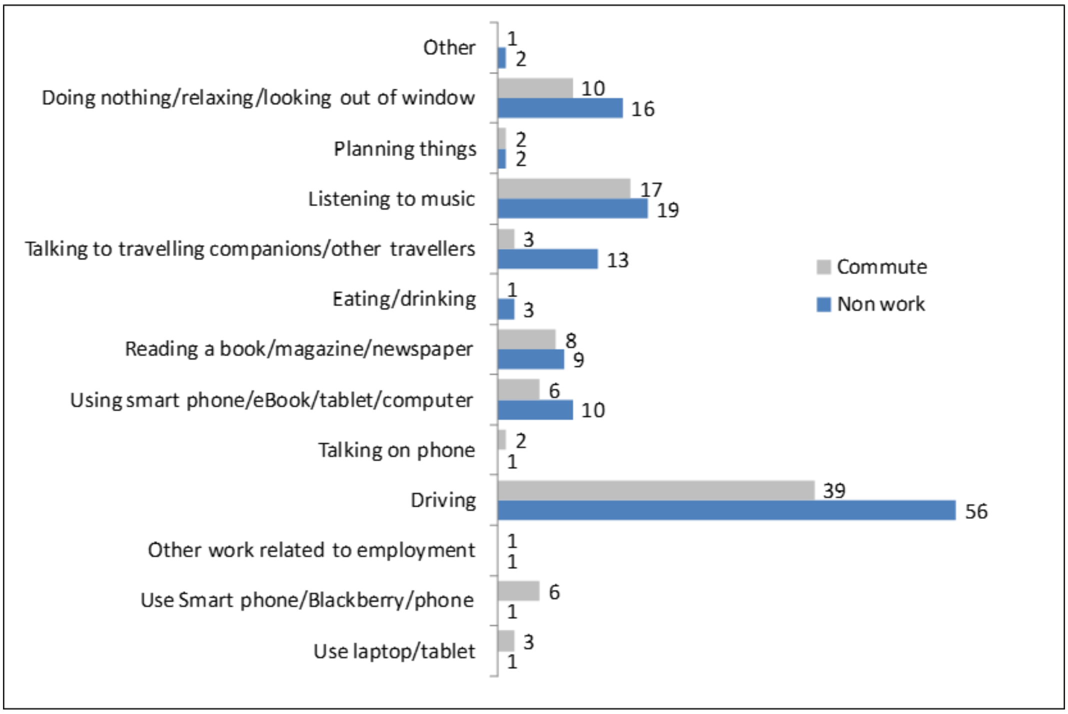


Table 3.20: Average minutes on activities by mode and purpose

Activity	Car		Train		Bus		'Other PT'	
	Com-mute	Other non-work	Com-mute	Other non-work	Com-mute	Other non-work	Com-mute	Other non-work
Work–related								
Use laptop/tablet	*	*	9	3	1	2	3	*
Use smart phone/Blackberry/phone	1	*	10	2	5	2	7	1
Other work related to employment	*	*	3	1	*	2	2	*
Non-work-related								
Driving	39	56	n/a	n/a	n/a	n/a	n/a	n/a
Talking on phone	1	*	1	2	1	2	2	1
Using smart phone/eBook/tablet/ computer	*	1	12	20	8	9	8	8
Reading a book/magazine/newspaper	n/a	n/a	10	16	3	2	7	3
Eating/drinking	1	3	2	5	1	1	1	*
Talking to travelling companions/ other travellers	3	20	3	11	2	7	1	7
Listening to music	27	35	11	15	11	9	9	7
Planning things	2	2	2	4	1	2	1	2
Doing nothing/relaxing/looking out of window	n/a	n/a	11	20	11	14	7	9

[29] The mean in-vehicle time for commuting trips was 54 minutes and the mean in-vehicle time for other non-work travellers was 1 hour 15 minutes.

Activity	Car		Train		Bus		'Other PT'	
	Com-mute	Other non-work	Com-mute	Other non-work	Com-mute	Other non-work	Com-mute	Other non-work
Other	1	2	1	3	*	*	*	*
Total time	**1:15**	**1:59**	**1:15**	**1:42**	**0:44**	**0:52**	**0:48**	**0:38**
Sample size	**1,025**	**1,030**	**993**	**1,113**	**367**	**668**	**611**	**535**

* = less than 0.5%

	1st		2nd		3rd

Variability

Respondents who made the trip at least once a month (91% of the commute and 55% of the other non-work sample) were asked whether the trip time varies from one day to another.

For around three quarters (74% commuters and 72% other non-work) it did vary:

	Commute %	Other non-work %
• Yes, frequently	34	33
• Yes, occasionally	39	39
• No	26	29

This reinforces the case for undertaking SP2 on reliability as part of this study.

The greatest variation in trip times was experienced by the bus commute sample (44% frequently and 39% occasionally) and car other non-work (41% and 40%) and car commute (42% and 36%) samples. The least variability was experienced by the train other non-work sample (**Table 3.21**).

Table 3.21: Whether trip time varies from one day to another by mode and purpose

Frequency	Car		Train		Bus		'Other PT'	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
Yes, frequently	42	41	26	24	44	37	28	23
Yes, occasionally	36	40	41	33	39	39	43	44
No	22	19	33	43	17	24	30	33
Sample size	**947**	**550**	**860**	**444**	**358**	**531**	**566**	**323**

* = less than 0.5%

3.3.4 Car-specific results

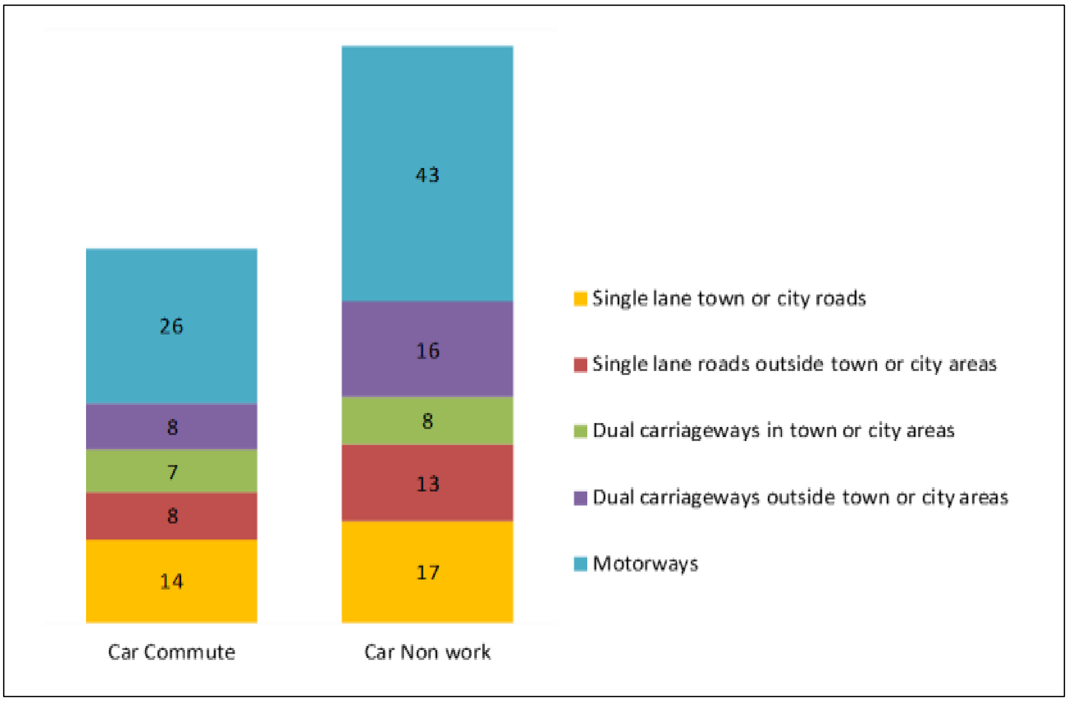
Road types

Car travellers were reminded of how long they said their car trip took, and asked how much of that time in minutes they spent on each of the following road types:

- Single lane town or city roads
- Single lane roads outside town or city areas
- Dual carriageways in town or city areas
- Dual carriageways outside town or city areas
- Motorways

For both commute and other non-work, the longest time was spent on motorways followed by single lane town or city roads. The least time for both purposes was spent on dual carriageways in town or city areas. See **Figure 3.3**.

Figure 3.3: Reported time on different road types (average minutes)



Sample size: 1,025 commute, 1,030 other non-work

Conditions

Car travellers were then asked how much of their trip time in minutes they spent in each of the following road traffic conditions (**Figure 3.4**):

Figure 3.4: Different traffic conditions presented in the SP

Heavy traffic – Your speed is noticeably restricted and frequent gear changes are required.

Light traffic – You can travel close to the speed limit most of the time, but you have to slow down every so often

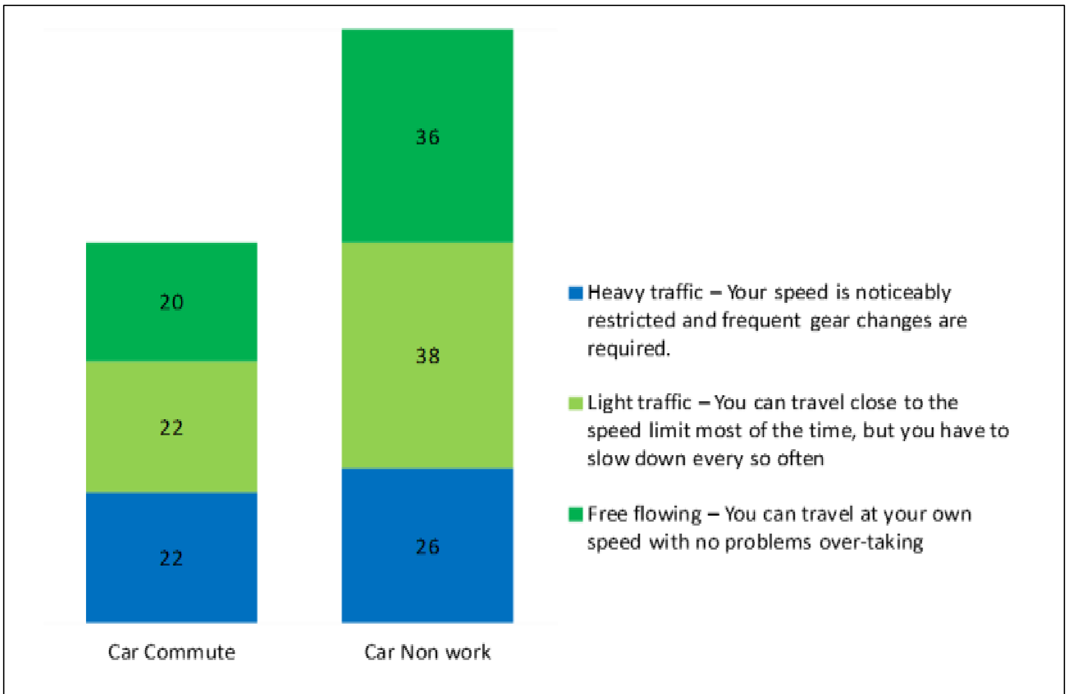
Free flowing – You can travel at your own speed with no problems over-taking

For car commuters there was a fairly even balance between the three traffic levels.

For the car other non-work sample, the longest time was spent in light traffic and free-flowing traffic.

These results are shown in **Figure 3.5**.

Figure 3.5: Reported time in different conditions (average minutes)



Sample size: 1,025 commute, 1,030 other non-work

3.3.5 PT-specific results

This section provides trip information on public transport trips with respect to crowding levels and time spent in different conditions, both of which were variables in SP3. Additional public transport trip information covering access, egress, frequency, wait time, interchange and numbers of modes used is provided in **Appendix C**.

Crowding

Public transport users were asked how crowded the service was when they boarded (**Table 3.22**).

There were relatively low levels of crowding particularly for train and bus, with between 39% and 51% reporting that there were plenty of seats free and they did not have to sit next to anyone.

At the other extreme, 30% of 'other PT' commuters said there were no seats free with some standing or densely packed.

Table 3.22: Crowding level when boarded by mode and purpose

Crowding Levels	Train		Bus		'Other PT'	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
Plenty of seats free and did not have to sit next to anyone	42	39	51	40	30	28
A few seats free but had to sit next to someone. No one standing	20	15	22	12	18	11
Could sit with people travelling with me	7	30	7	30	4	24
Could not sit with people travelling with me	*	1	*	1	1	2
A few seats free but had to sit next to someone. Some people were standing	13	6	10	7	17	15
A few seats free but could not sit with people travelling with me. No one standing	1	1	1	3	*	2
A few seats free but could not sit with people travelling with me. Some people were standing	*	2	1	2	0	4
No seats free - a few others standing	8	3	3	3	16	8
No seats free - densely packed	9	4	4	4	14	7
Sample size	**993**	**1,113**	**367**	**668**	**611**	**535**

* = less than 0.5%

With reference to **Table 3.23**, 27% of commuters and 16% of non-work public transport users said they stood for some or all of their journeys.

Table 3.23: Whether stood for any or part of the journey by mode and purpose

	Train		Bus		'Other PT'	
	Com-mute %	Other non-work %	Com-mute %	Other non-work %	Com-mute %	Other non-work %
Yes, all of it	8	4	5	5	27	19
Yes, about three quarters of it	2	1	2	1	5	4
Yes, about half of it	5	1	3	2	6	4
Yes, about a quarter of it	5	5	4	2	8	7
No, none of it	80	89	86	90	54	65
Sample size	**993**	**1,113**	**367**	**668**	**611**	**535**

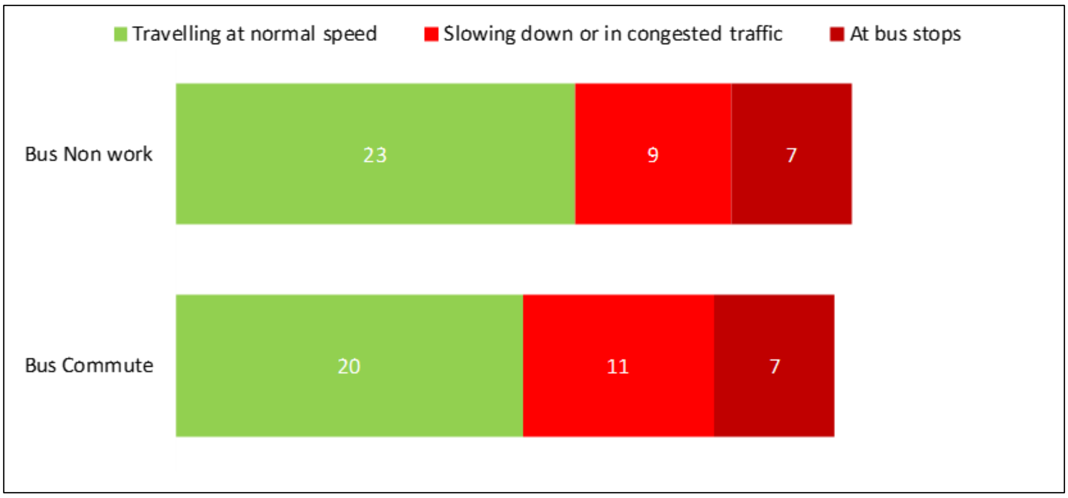
Time in different conditions

With reference to **Figure 3.6**, the bus sample was asked what proportion of their bus journey was spent in the following three conditions:

- Travelling at normal speed
- Slowing down or in congested traffic
- At bus stops.

For bus commuters, almost half the in-vehicle time (18 minutes) was spent slowing down, in congested conditions or at bus stops, as compared to 20 minutes spent at normal speed.

Figure 3.6: Bus users: reported time in different conditions (average minutes)



3.3.6 Summary

Overall, the respondent and trip characteristic findings shown above from the SP commute and other non-work samples are broadly representative when compared to the NTS.

The one area where there were significant disparities is with respect to time and distance, where the SP sample had much longer times and distances than the NTS sample. As mentioned in the commentary, the SP sample by its very nature over-sampled longer distance travellers, but this bias will be re-weighted at the implementation stage (**Chapter 7**).

3.4 Employees' and employers' business SP – market research results

3.4.1 Introduction

This section sets out the market research results for the employers' business survey and the employees' business part of the general public SP survey.

Although different questionnaires were used for employees and employers, there were some common questions on the nature of the company, travel policy and the employee trip.

For employers, the questionnaire focussed upon a 'briefcase' trip made by a typical employee by train or car[30]. Briefcase travel was defined as trips made by office-based staff travelling to conduct meetings and similar business activities but not to provide trade services.

Where there are common questions, the results below compare the two surveys.

3.4.2 Business characteristics

This section sets out key characteristics of the business covering numbers of staff and turnover. Additional data on region, type of organisation and numbers of sites is provided in **Appendix C**.

Number of staff

Table 3.24: Number of people employed at site

People	Employees %	Employers %
1 to 4	22	8
5 to 9	6	8
10 to 19	9	8
20-49	12	21
50-99	10	21
100-199	8	13
200+	29	21
Don't know	3	2
Sample size	**1,486**	**400**

[30] 'Other PT' was dropped as a mode for business travel.

The number of people employed at the location where they were based is shown for both surveys in **Table 3.24**. The employee sample had higher proportions in the smallest and largest bands, relative to the employer sample.

Turnover

Annual turnover for the site and the overall organisation (if multiple sites) was probed. Over half (58% for site, 63% for overall) said they did not know or refused to answer this question. Of those who did answer for the site, the distribution of turnover was:

- Under £100,000 1%
- £100,000 to £499,000 14%
- £500,000 to £999,999 8%
- £1m-£5m 25%
- £5m to £10m 12%
- Over £10m 41%

3.4.3 Employee characteristics

This section sets out the key characteristics of the employees from the general public SP survey, and that of the typical employee selected in the employer survey for the purpose of describing a typical car or rail trip.

It should be noted that the employee sample from the general public survey includes 18% who were self-employed.

Employees were asked which occupational category they belonged to. Two-thirds described themselves as Managers, Directors and Senior Officials or as Professional.

- Managers, Directors and Senior Officials 35%
- Professional 31%
- Sales and Customer Service 9%
- Associate Professional and Technical 6%
- Skilled Trades 6%
- Administrative and Secretarial 4%
- Caring, Leisure and Other Service 2%
- Unskilled manual 2%
- Process, Plant and Machine 1%
- Other 4%

Base: 2,160

With reference to **Table 3.25**, both surveys probed whether the employee’s role within the company was senior, middle-ranking or junior. For the employer survey this pertained to the employee selected for the purpose of describing a typical car or rail trip.

The employer survey had a much higher proportion of senior, and a lower proportion of middle-ranking employees than the general public survey.

Table 3.25: Level of employee

Level	Employees %	Employers %
Senior	38	57
Middle ranking	48	34
Junior	11	9
Other	2	-
Sample size	**2,160**	**400**

The mean typical working hours for employees in the general public survey was 43 hours, slightly higher than the 41 hours for employees considered in the course of the employer survey.

Income

Employees' annual personal income, before tax and other deductions (including allowances or support from other household members, state benefits, etc.), was probed. This is compared to an estimate of the employee's income from the employer survey in **Table 3.26** below. The employer survey includes a high proportion of 'don't knows' and refusals (52% in total).

Table 3.26: Annual gross personal income by purpose

Income bands	Employees %	Employers %
Under £10K	5	1
£10-20K	12	6
£20-30K	18	11
£30-40K	20	9
£40-50K	14	8
£50-75K	15	6
£75-100K	7	5
More than £100K	6	3
Don't know	0	38
Refusal	1	14
Not stated	2	-
Sample size	**2,217**	**400**

* = less than 0.5%

Details of the age and gender of this sample are provided in **Appendix C**.

3.4.4 Trip characteristics

Cost and time

The mean costs for both car and rail were higher for the employer survey than for the general public employee survey. The mean in-vehicle time for both surveys was almost identical for rail at just under two hours. However, for car, the general public survey mean car time was almost an hour longer than for the employer survey. See **Table 3.27**.

Table 3.27: Cost and time

	Employees		Employers	
	Car	Train	Car	Train
Mean cost	£11.31	£72.63	£14.73	£93.84
Mean time	2:34	1:58	1:35	1:57
Sample size	**948**	**1,004**	**244**	**143**

For the SP employee sample, the distributions of in-vehicle time (IVT) for car and rail are shown in **Table 3.28**, whilst the distributions of door-to-door distance are shown in **Table 3.29**; both are compared to NTS.

Table 3.28: Employees' business in-vehicle time (SP compared to NTS)

Time bands	Car		Rail	
	SP %	NTS %	SP %	NTS %
0 to 15 minutes	3	38	2	8
16 to 30 minutes	10	29	5	24
31 to 45 minutes	8	12	6	22
46 to 60 minutes	9	7	11	14
61 to 90 minutes	11	6	15	14
91 to 120 minutes	9	3	25	7
121+ minutes	50	4	35	11
Sample size	**948**	**25,576**	**1,004**	**1,572**

Table 3.29: Employees' business door-to-door distance (SP compared to NTS in-vehicle distance)

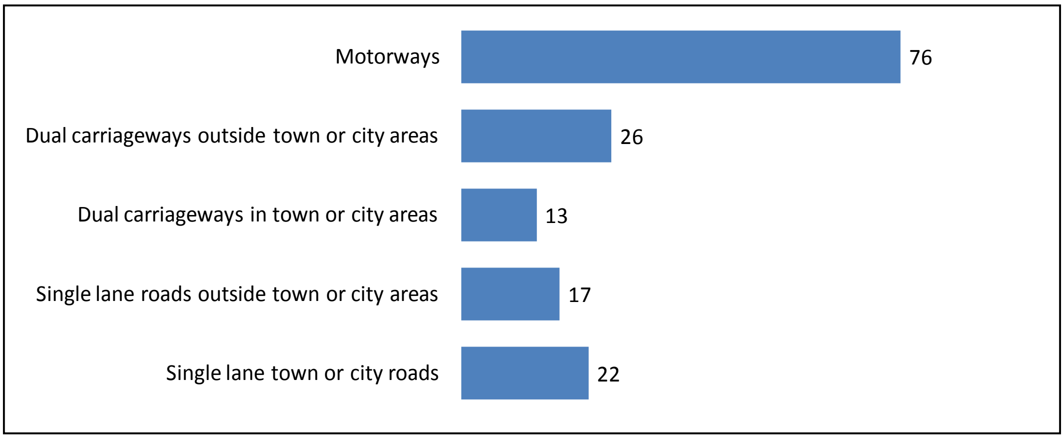
Distance bands	Car		Rail	
	SP %	NTS %	SP %	NTS %
0 to 5 miles	11	38	5	4
6 to 10 miles	5	20	3	13
11 to 20 miles	9	17	8	21
21 to 30 miles	6	7	2	10
31 to 50 miles	11	8	7	18
51 to 100 miles	19	6	16	17
100 to 150 miles	15	2	31	9
150+ miles	23	1	29	8
Sample size	**948**	**25,576**	**1,004**	**1,573**

In common with the commuter and non-work samples, the NTS data shows markedly shorter IVT than the SP sample, particularly for car. The SP sample by its very nature over-sampled longer distance car and rail travellers.

Door-to-door road distances were calculated automatically from the origin and destination locations marked on the Google maps within the questionnaire. It should be noted that these distances include access and egress modes for rail. Again, the NTS data shows markedly shorter distances than the SP sample.

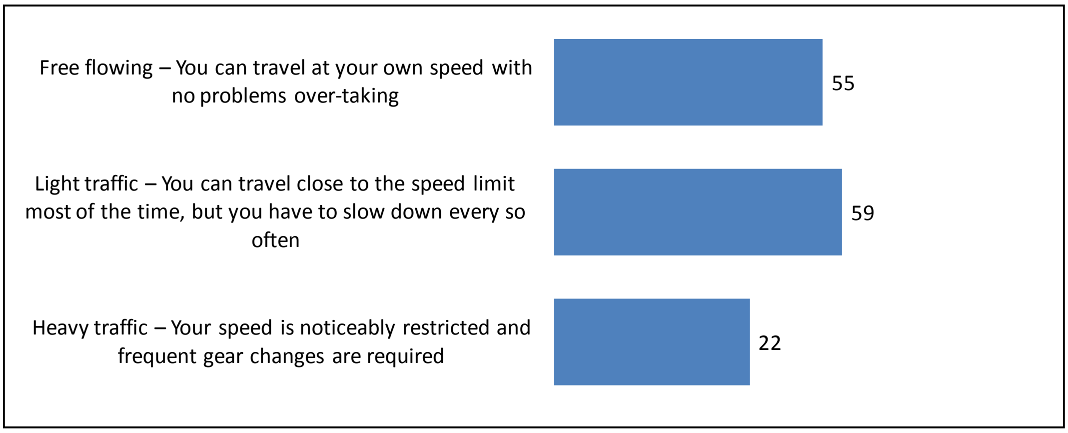
For the general public employees' business car sample, the average time spent on different types of roads is shown in **Figure 3.7** below (these findings should be contextualised against the average travel times given in **Table 3.27**).

Figure 3.7: Time spent on different road types (minutes)



For the same car sample, the average time spent in different traffic conditions is shown in **Figure 3.8** below.

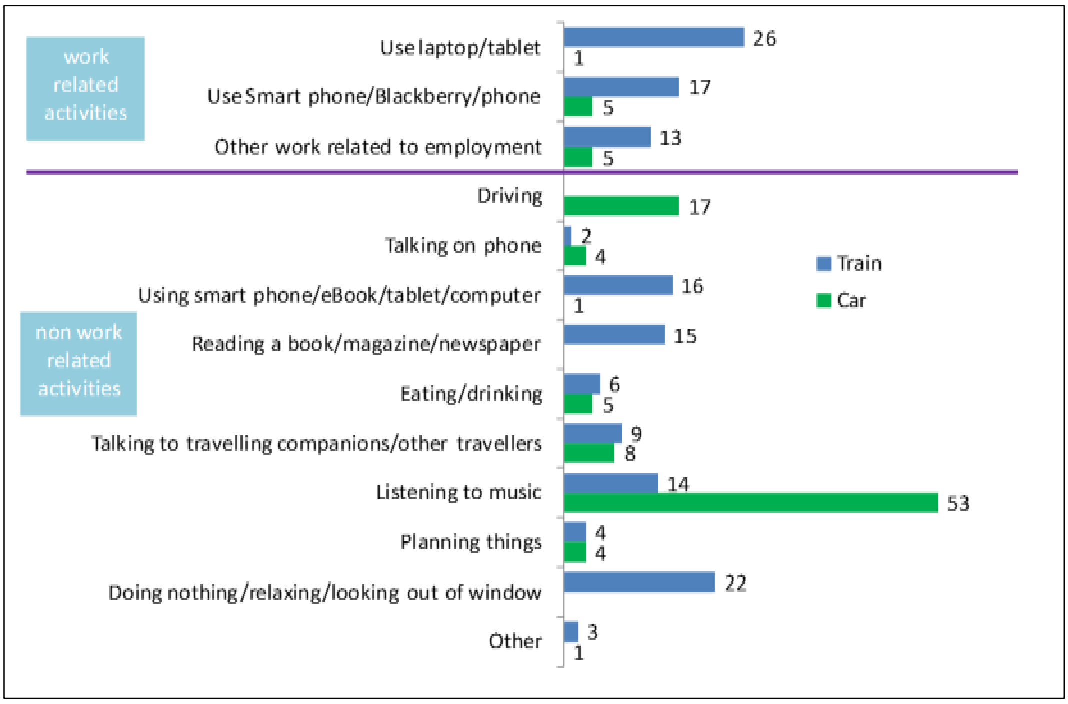
Figure 3.8: Time spent in different traffic conditions (minutes)



Activities

Employees were reminded of their reported one way trip time and asked approximately how much of that time they spend undertaking work and non-work related activities. (Again, these findings should be contextualised against the average travel times given in **Table 3.27**).

Figure 3.9: Activities undertaken by employees during trip (average minutes) by mode



With reference to **Figure 3.9**, the main activities undertaken by mode were:

- Car: Listening to music (53 minutes on average) and driving (17 minutes).
- Train: Using smart phone/eBook/tablet/computer (33 minutes: 17 minutes work related), work related use of laptop/tablet (26 minutes), doing nothing/relaxing/looking out of window (22 minutes), listening to music (14 minutes) and other work related to employment (13 minutes).

The 2009 report on ‘Productive Use of Rail Travel Time and the Valuation of Travel Time Savings for Rail Business Travellers’[31] used different categories and showed the proportions of work related activities undertaken as follows:

- Preparing for a meeting 38%
- Making/receiving calls 43%
- Talking to colleagues/other 12%
- Use of a laptop 23%
- Use of a PDA/Blackberry 25%
- Other work related to employment 36%.

By comparison, in our own SP survey, 35% used a laptop, 56% used a smart phone/Blackberry and 29% did other work related to employment.

There has been a notable increase in the use of electronic devices for work-related activities on-train since 2009. Nonetheless, it is clear that a large proportion of rail travel time is spent on non-work activities.

Data on group size and the leg of the trip is provided in **Appendix C**, whilst deeper analysis of the ‘Hensher’ parameters (covering the proportion of travel time spent working, the productivity of that time, and the proportion of travel time that takes place outside of normal working hours) can be found in **Appendix H**.

3.5 RP market research results

3.5.1 Introduction

This section provides some of the market research findings from the general public RP market research. As was described in **Chapter 2**, the RP sample was comprised of rail passengers travelling from Birmingham New Street, Birmingham Moor Street, Birmingham Snow Hill, Stoke, Stafford, Rugby or Peterborough to London.

Recruitment included screening to ensure that participants were familiar with the times and costs of alternatives operators, and that they were making a rail trip originating at the surveyed station and ending at the London terminus station.

The target number of interviews was 2,500, with half making business trips and half making commuting or other non-work trips. The overall number of interviews was 2,646 with the following numbers by trip purpose:

- employees’ business: 1,311
- commuting: 451
- other non-work: 884

As the RP research was primarily undertaken to validate the SP research, the RP findings are compared with the SP data for relevant sample and trip characteristics. The rail SP data used for comparison covers the same routes,

[31] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/4003/productive-use-of-travel-time.pdf

although it should be noted that, unlike the RP sample, respondents did not have to be aware of the alternative operator’s costs and times; nor did they need to have started their rail trip at the recruitment station.

3.5.2 Respondent characteristics

Age

The median age range for the employees’ business and commuting samples was 40-49, as compared to 21-29 for the other non-work sample.

The other non-work samples had the largest proportion aged over 60 years old (20%, as compared to between 10% for employees’ business and 3% for commuting).

The differences in age band by purpose are shown in **Table 3.30**.

Table 3.30: Age by purpose (RP compared to SP)

Age bands	Employees’ business		Commute		Other non-work	
	RP %	SP %	RP %	SP %	RP %	SP %
17-20	1	1	8	5	8	16
21-29	15	15	25	21	24	19
30-39	25	24	23	34	16	19
40-49	29	26	26	20	14	13
50-59	23	25	14	15	16	7
60-69	7	9	3	4	15	20
70+	1	0	*	0	5	6
Sample size	**1,311**	**167**	**451**	**99**	**884**	**145**

* = less than 0.5%

There is very little difference between the RP and SP samples for employees’ business. For commute, the main difference is that the SP sample had a significantly[32] greater proportion in the 30-39 age range than the RP sample. All other differences are not statistically significant. For other non-work, the significant differences are that the RP sample had a lower proportion in the 17-20 age range and a higher proportion in the 65-59 age range than the SP sample.

Gender

The employees’ business and commuter samples were more likely to be male, particularly for employees’ business, whereas the other non-work sample was more likely to be female (see **Table 3.31**).

[32] at the 95% confidence level

Table 3.31: Gender by purpose (RP compared to SP)

Gender	Employees' business		Commute		Other non-work	
	RP %	SP %	RP %	SP %	RP %	SP %
Male	62	63	56	65	45	36
Female	38	37	44	35	55	64
Sample size	**1,311**	**167**	**451**	**99**	**884**	**145**

As was the case for age, there was very little difference between the RP and SP samples for employees' business. For commute, there was no statistically significant difference between the samples. For other non-work, there was a significantly higher proportion of females in the SP sample than in the RP sample.

Household income

The gross annual household income, before deductions for tax and National Insurance, is shown in **Table 3.32**.

Table 3.32: Gross annual household income by purpose (RP compared to SP)

Income bands	Employees' business		Commute		Other non-work	
	RP %	SP %	RP %	SP %	RP %	SP %
Under £10K	1	1	5	5	10	14
£10-20K	3	2	6	7	13	16
£20-30K	6	8	8	9	17	19
£30-40K	11	8	9	11	14	14
£40-50K	12	14	13	14	11	11
£50-75K	25	28	24	18	15	7
£75-100K	18	19	15	19	8	10
More than £100K	21	17	19	15	8	8
Don't know	0	0	0	0	*	1
Refusal	*	0	0	1	2	0
Not stated	1	2	1	0	1	1
Sample size	**1,311**	**167**	**451**	**99**	**884**	**145**

* = less than 0.5%

The employees' business and commuter samples had much higher household incomes than the other non-work sample: 64% of employers business and 58% of commuters had household incomes of £50,000 per year or more, as compared to 31% for the other non-work sample.

For the three purpose samples, there was little difference between the RP and SP data with respect to household income.

Personal income

As for household income, the employees' business and commuter samples had much higher personal incomes than the other non-work sample (**Table 3.33**).

Table 3.33: Annual gross personal income by purpose (RP compared to SP)

Income bands	Employees' business		Commute		Other non-work	
	RP %	SP %	RP %	SP %	RP %	SP %
Under £10K	3	3	14	10	25	37
£10-20K	7	5	10	8	22	18
£20-30K	15	16	13	19	20	17
£30-40K	18	18	15	18	11	10
£40-50K	16	17	11	15	7	6
£50-75K	20	22	19	17	6	4
£75-100K	11	10	7	6	2	5
More than £100K	9	7	10	5	3	1
Don't know	0	0	0	0	*	1
Refusal	*	0	0	1	2	0
Not stated	2	3	2	0	2	1
Sample size	**1,311**	**167**	**451**	**99**	**884**	**145**

* = less than 0.5%

40% of employees' business and 36% of commuters had personal incomes of £50,000 per year or more, as compared to 11% for the other non-work sample. A quarter of the other non-work sample had annual personal incomes of less than £10,000, as compared to 14% for the commuter sample and 3% for the employees' business sample.

There was little difference between the RP and SP samples for employees' business and commute. For other non-work, there was a significantly larger proportion with personal incomes lower than £10,000 per annum in the SP sample, as compared with the RP sample. This correlates with the higher proportion of retired people and students in the SP sample.

Data on employment status is provided in **Appendix C**.

3.5.3 Trip characteristics

This section describes the key characteristics of the trip for the RP sample. It covers group size and activities undertaken during the trip.

Data on leg of trip, ticket type, time of day of the trip, frequency, access and egress mode is provided in **Appendix C**.

Group size

With reference to **Table 3.34**, 93% of commuters, 81% on employees' business and 63% of the other non-work sample travelled alone.

Table 3.34: Group size by purpose

	Employees' business		Commute		Other non-work	
	%	SP %	RP %	SP %	RP %	SP %
Travelled alone	81	74	93	82	63	58
1 other adult	14	21	4	16	33	37
2+ other adults	5	5	3	2	4	6
Children aged 6-17						
1	0	0	0	5	1	1
2+	0	0	0	1	0	1
Children aged 5 or under						
1	0	0	1	2	2	2
2+	0	0	0	0	1	0
Sample size	**1,311**	**167**	**451**	**99**	**884**	**145**

* = less than 0.5%

For both employees' business and commute, there were significantly larger proportions in the RP samples than in the SP samples that were travelling alone.

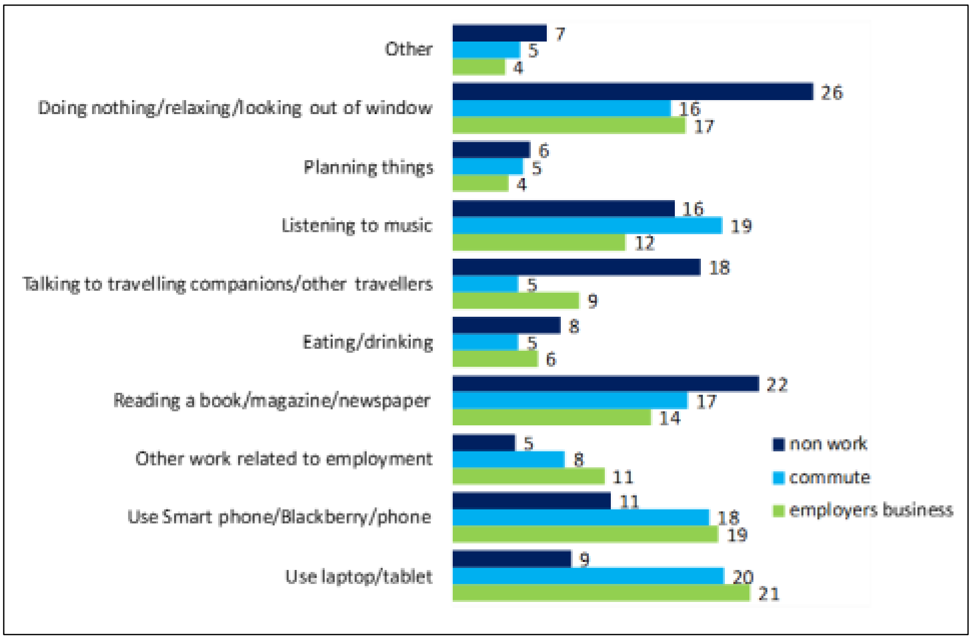
For other non-work, there were no statistically significant differences between the two samples.

Activities undertaken during the trip

Respondents were reminded of their reported one way trip time and asked approximately how much of that time they spend undertaking the work related and non-work activities shown in **Figure 3.10**.

The non-work related activities were shown first for the other non-work sample.

Figure 3.10: Activities undertaken during the trip (average minutes) by purpose



As shown in **Figure 3.10**, the main activities undertaken by purpose were:

- Employees' business: using laptop/tablet (21 minutes), using smart phone/eBook/ tablet/computer (19 minutes), doing nothing/ relaxing/looking out of window (17 minutes), reading a book/magazine/ newspaper (14 minutes), listening to music (12 minutes) and other work related to employment (11 minutes).
- Commuters: using laptop/tablet (20 minutes), listening to music (19 minutes), using smart phone/eBook/tablet/computer (18 minutes), reading a book/magazine/ newspaper (17 minutes) and doing nothing/relaxing/looking out of window (16 minutes).
- Other non-work: doing nothing/relaxing/looking out of window (26 minutes), reading a book/magazine/newspaper (22 minutes), talking to travelling companions/other travellers (18 minutes) and listening to music (16 minutes).

3.5.4 Trip planning

Travellers from Birmingham were asked whether, prior to making their trip, they compared the fares of Virgin Trains, London Midland and/or Chiltern Railways.

Travellers from Stoke, Stafford and Rugby were asked whether, prior to making their trip, they compared the fares of Virgin Trains and London Midland.

Travellers from Peterborough were asked whether, prior to making their trip, they compared the fares of Great Northern and East Coast.

With reference to **Table 3.35**, 43% of those on employees' business and commuting trips compared the fares, whilst 61% of those on other non-work compared the fares.

Table 3.35: Whether travellers compared fares

	Employees' business %	Commute %	Other non-work %
Yes, on the web	38	37	56
Yes, at the booking office	2	3	3
Yes, I consulted friends/colleagues	*	*	1
Yes, other	3	3	1
Yes total	**43**	**43**	**61**
No, I already knew them	16	26	16
No, I'm not bothered about differences in fares	5	3	4
No, too much hassle/effort involved in comparing fares	4	4	4
No, I bought the ticket shortly before travelling	11	12	7
No, other	21	13	9
Sample size	**1,280**	**319**	**878**

* = less than 0.5%

Over a quarter (26%) of commuters and 16% of both employees' business and other non-work travellers said that they did not compare the fares as they already knew them.

Similarly, travellers from Birmingham were asked whether, prior to making their trip, they compared the times of Virgin Trains, London Midland and/or Chiltern Railways; travellers from Stoke, Stafford and Rugby were asked whether they compared the times of Virgin Trains and London Midland; and travellers from Peterborough were asked whether they compared the times of Great Northern and East Coast.

With reference to **Table 3.36**, 52% of those on employees' business, 43% on commuting trips and 61% on other non-work trips did compare times.

Almost a third (31%) of commuters and 18% of those on employees' business and 17% on other non-work trips said that they did not compare times as they already knew them.

Data on the perceived quality of competing operators is provided in **Appendix C**.

Table 3.36: Whether travellers compared train times

	Employees' business %	**Commute %**	**Other non-work %**
Yes, on the web	47	39	56
Yes, at the booking office	2	2	3
Yes, I consulted friends/colleagues	*	0	1
Yes, other	3	2	1
Yes total	**52**	**43**	**61**
No, I already knew them	18	31	17
No, I'm not bothered about differences in times	6	7	6
No, too much hassle/effort involved in comparing times	4	3	4
No, I bought the ticket shortly before travelling	8	8	6
No, other	47	39	56
Sample size	**1,280**	**319**	**878**

* = less than 0.5%

3.5.5 Summary

Overall, the respondent and trip characteristic findings from the RP sample match the SP sample well.

3.6 Analysis of respondents vs. non-respondents

This section compares the age, gender and trip purpose between the samples of respondents and non-respondents from the intercept surveys. The comparison was limited to these aspects as they were the only questions included in the recruitment survey. It should be noted that some potential responders, typically other non-work, could not complete a questionnaire because quotas were closed.

3.6.1 SP sample

For the SP sample, respondents tended to be older than non-respondents: (**Table 3.37**)

- 31% of respondents were aged over 50 years old, as compared to 24% for non-respondents.
- 28% of respondents were aged under 30 years old, as compared to 32% for non-respondents.

Table 3.37: Age

Age bands	Respondent %	Non-respondent %
17-20	9	9
21-29	19	23
30-39	20	22
40-49	21	21
50-59	18	15
60-69	10	7
70+	3	2
Sample size	**8,448**	**21,032**

Respondents were slightly more likely to be female than non-respondents: 50% female for respondents, as compared to 46% for non-respondents (**Table 3.38**).

Table 3.38: Gender

Gender	Respondent %	Non-respondent %
Male	50	54
Female	50	46
Sample size	**8,448**	**21,034**

Respondents were more likely to be making commuting trips and less likely to be making employees' business trips than non-respondents (**Table 3.39**):

- 35% of respondents were on commuting trips, as compared to 29% for non-respondents.
- 26% of respondents were on employees' business trips, as compared to 38% for non-respondents.

Table 3.39: Purpose

Purpose	Respondent %	Non-respondent %
Commuting	35	29
Employees' business	26	38
Non-work	39	33
Sample size	**8,448**	**21,034**

3.6.2 RP sample

As for the SP sample the RP sample respondents tended to be older than non-respondents (**Table 3.40**):

- 30% of respondents were aged over 50 years old, as compared to 26% for non-respondents.
- 25% of respondents were aged under 30 years old, as compared to 30% for non-respondents.

Table 3.40: Age

Age bands	Respondent %	Non-respondent %
17-20	5	7
21-29	20	23
30-39	22	21
40-49	24	23
50-59	19	17
60-69	9	8
70+	2	1
Sample size	**2,639**	**4,153**

There was no difference in gender between respondents and non-respondents (**Table 3.41**).

Table 3.41: Gender

Gender	Respondent %	Non-respondent %
Male	55	55
Female	45	45
Sample size	**2,639**	**4,153**

Respondents were more likely to be making commuting trips and less likely to be making employees' business trips than non-respondents (**Table 3.42**):

- 17% of respondents were on commuting trips, as compared to 9% for non-respondents.

- 50% of respondents were on employees’ business trips, as compared to 55% for non-respondents.

Table 3.42: Purpose

Purpose	Respondent %	Non-respondent %
Commuting	17	9
Employees’ business	50	55
Non-work	33	36
Sample size	**2,639**	**4,153**

3.6.3 Summary

The analysis of respondents compared to non-respondents based on the questions asked in the recruitment questionnaire shows that the final SP and RP samples were reasonably representative of those who were recruited. The main differences were that the recruited sample tended to have an older profile and be less likely to be making employees' business trips. The latter issue was addressed by deliberately over-sampling employees' business respondents.

Therefore, we would expect the results from the recruited sample to be broadly representative of all potential respondents.

3.7 Diagnostic testing

Diagnostic questions were asked after each of the three SP exercises. These asked the participant to indicate whether they agreed or disagreed with the following four statements:

- *“I was able to understand the choices I was faced with”.*
- *“I found the options I was presented with realistic”.*
- *“I was able to make choices as in real life”.*
- *“I found it easy to choose between the options I was presented with”.*

Table 3.43 to **Table 3.54** below summarise the responses to these statements for the three exercises (i.e. SP1, SP2 and SP3) by interview method (i.e. CATI and on-line). It should be noted that these tables show responses from the overall sample, i.e. before the cleaning undertaken prior to modelling (**Section 4.3**). The responses would be even more reassuring after cleaning as there is typically a correlation between illogical SP responses and perceived difficulty, lack of understanding and perceived lack of realism.

The diagnostic scores are good overall:

- For SP1 and SP3, over four-fifths ‘agreed’ or ‘strongly agreed’ that they were able to understand the choices they were faced with (84% and 83% respectively). For SP2, 78% ‘agreed’ or ‘strongly agreed’ that they were able to understand the choices they were faced with.

- Around two-thirds overall ‘agreed’ or ‘strongly agreed’ that the options presented were realistic: 68% for SP3, 66% for SP1 and 64% for SP2.
- Over two-thirds overall ‘agreed’ or ‘strongly agreed’ that they were able to make choices as in real life: 72% for SP1, 71% for SP3 and 68% for SP2.
- Over two-thirds overall ‘agreed’ or ‘strongly agreed’ that they found it easy to choose between the options they were presented with: 79% for SP1, 76% for SP3 and 71% for SP2.

Overall, SP1 had the highest diagnostic scores followed closely by SP3, with SP2 having the lowest scores.

Overall, the CATI administered interviews received higher diagnostic scores for all aspects for the three exercises.

Further analysis by mode (not reported here) showed that bus respondents gave the highest diagnostic scores, followed by train, with car respondents giving the lowest diagnostic scores.

Further analysis by purpose (again not reported here) showed that other non-work respondents gave the highest diagnostic scores, with commute and employees’ business about the same.

3.7.1 SP1

Table 3.43: I was able to understand the choices I was faced with

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	3	0	3
Somewhat disagree	3	2	4
Neither agree nor disagree	10	10	10
Somewhat agree	27	22	28
Strongly agree	57	66	55
Mean	4.31	4.52	4.27
Base	8,486	1,380	7,098

Table 3.44: I found the options I was presented with realistic

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	4	3	4
Somewhat disagree	9	3	10
Neither agree nor disagree	21	15	22
Somewhat agree	35	28	37
Strongly agree	31	51	27
Mean	3.79	4.21	3.71
Base	8,486	1,380	7,098

Table 3.45: I was able to make choices as in real life

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	5	4	6
Somewhat disagree	7	4	7
Neither agree nor disagree	16	12	16
Somewhat agree	33	25	35
Strongly agree	39	56	36
Mean	3.94	4.25	3.88
Base	8,486	1,380	7,098

Table 3.46: I found it easy to choose between the options I was presented with

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	3	1	3
Somewhat disagree	5	3	6
Neither agree nor disagree	13	11	13
Somewhat agree	35	26	37
Strongly agree	44	58	42
Mean	4.13	4.37	4.08
Base	8,486	1,380	7,098

3.7.2 SP2

Table 3.47: I was able to understand the choices I was faced with

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	2	0	3
Somewhat disagree	6	2	6
Neither agree nor disagree	14	12	14
Somewhat agree	34	25	36
Strongly agree	44	61	41
Mean	4.12	4.44	4.06
Base	8,486	1,380	7,098

Table 3.48: I found the options I was presented with realistic

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	4	3	4
Somewhat disagree	9	4	10
Neither agree nor disagree	22	16	24
Somewhat agree	34	26	36
Strongly agree	30	51	26
Mean	3.78	4.19	3.7
Base	8,486	1,380	7,098

Table 3.49: I was able to make choices as in real life

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	5	2	5
Somewhat disagree	8	4	9
Neither agree nor disagree	19	15	20
Somewhat agree	34	25	35
Strongly agree	34	53	31
Mean	3.85	4.23	3.78
Base	8,486	1,380	7,098

Table 3.50: I found it easy to choose between the options I was presented with

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	3	1	3
Somewhat disagree	8	5	9
Neither agree nor disagree	17	14	18
Somewhat agree	36	27	38
Strongly agree	35	54	32
Mean	3.93	4.27	3.87
Base	8,486	1,380	7,098

3.7.3 SP3

Table 3.51: I was able to understand the choices I was faced with

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	2	1	2
Somewhat disagree	3	2	3
Neither agree nor disagree	12	12	11
Somewhat agree	32	23	34
Strongly agree	51	62	49
Mean	4.27	4.43	4.24
Base	8,486	1,380	7,098

Table 3.52: I found the options I was presented with realistic

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	4	3	4
Somewhat disagree	9	4	10
Neither agree nor disagree	20	17	20
Somewhat agree	34	24	36
Strongly agree	34	51	30
Mean	3.84	4.16	3.78
Base	8,486	1,380	7,098

Table 3.53: I was able to make choices as in real life

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	4	3	5
Somewhat disagree	7	4	7
Neither agree nor disagree	17	16	17
Somewhat agree	33	22	36
Strongly agree	38	55	35
Mean	3.94	4.21	3.89
Base	8,486	1,380	7,098

Table 3.54: I found it easy to choose between the options I was presented with

		Interview method	
	Total %	CATI %	On-line %
Strongly disagree	2	2	2
Somewhat disagree	6	4	7
Neither agree nor disagree	15	14	15
Somewhat agree	36	26	38
Strongly agree	40	55	38
Mean	4.06	4.28	4.02
Base	8,486	1,380	7,098

4 Development of 'Core' Behavioural Model Specification

4.1 Context and overview of model development process

In the ITT for this study, the Department stipulated that: *"How values are estimated will depend, at least in part, on the data collection and overall approaches being proposed. However, modelling and value estimation should make use of modern, innovative methods and should be able to:*

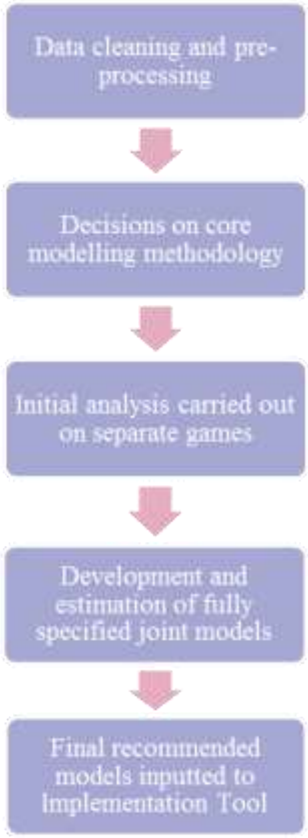
- *produce representative, national average values;*
- *investigate factors causing variation in the values to inform recommended segmentations and relationships / formulae which can be used to derive values for specific segments or contexts (such as income-segmented or High Speed Rail specific values); and*
- *address issues such as taste heterogeneity and non-linearity of preferences."*

Against this background, this chapter sets out the detailed process which was followed in developing a flexible and generic specification for the joint modelling of SP1-3 for each mode. As will become apparent in **Chapter 7**, this specification constituted the 'core' modelling input to the Implementation Tool which was used to generate national average values.

The present chapter omits consideration of mode choice (for both concessionary and non-concessionary travellers) and operator choice (for both SP and RP). These choice contexts were treated as auxiliary modelling strands, and were analysed using specifications more bespoke to each context; the process of developing these specifications is detailed in **Chapter 5** to follow. As will also become apparent in **Chapter 7**, these auxiliary models were not taken forward to the Implementation Tool.

The core behavioural model specification was developed in a systematic fashion, as summarised in **Figure 4.1**. First, we undertook preliminary work to ensure that the data met appropriate quality standards. Second, we developed a generic modelling specification in line with (and in some cases beyond) current best practice in the field of discrete choice modelling, before applying the model separately for each mode and SP game (i.e. SP1, SP2 and SP3). Third, having identified the set of covariates applicable to each mode and game, we implemented a more flexible version of the model capable of modelling SP1-3 jointly. Fourth, for each mode, we introduced additional elements of functionality, and identified the final specification to be taken forward to the Implementation Tool.

Figure 4.1: Modelling process



4.2 Glossary of modelling terms

To assist readers of this chapter, it will be useful to introduce and define a number of technical terms, associated with the statistical features of the choice models which are reported in this chapter (**Table 4.1**). In the course of this chapter, we will also introduce and define a number of parameters of the choice models; these will be introduced as the chapter progresses.

Table 4.1: Glossary of terms associated with the estimated choice models

Term	Abbreviated term	Definition
Final log-likelihood	Final LL	This is a measure of the goodness of fit of a choice model. The choice model is calibrated through a process of maximising the log-likelihood.
Adjusted rho-squared	Adj. ρ^2	This is a further goodness of fit measure derived from the log-likelihood, which 'adjusts' for the number of parameters in the model
Estimate	est.	This is a parameter of the choice model which is calibrated through the estimation process.
t-ratio	t-rat.	This is a measure of the statistical significance of an estimated parameter. In what follows, we typically test significant

Term	Abbreviated term	Definition
		difference from zero. If we apply tests against zero and one, then these are distinguished by the notation *t*-rat(0)) and *t*-rat(1).
Robust t-ratio	rob t-rat.	This is a 'robust' version of the t-ratio defined earlier. Robust statistics seek to provide methods that emulate popular statistical methods, but which are not unduly affected by outliers or other small departures from model assumptions.
Likelihood ratio probability value against previous specification	LR p-value against previous specification	This is a measure of the relative goodness of fit of one model specification over another.
Value of travel time	*VTT*	The monetary value of travel time.
Value function for a cost change	$v(\Delta c)$	A mathematical expression representing the monetary value of a change in travel cost from a given reference cost.
Value function for a time change	$v(\theta\Delta t)$	A mathematical expression representing the monetary value of a change in travel time from a given reference time.
'Underlying' VTT parameter	θ	This parameter represents the 'underlying' estimate of VTT.
Scale parameter	μ	This parameter represents the scale of the estimated VTT, which is related to the variance of the associated error term.
Elasticities	λ	These parameters represent the responsiveness (or 'elasticity') of VTT with respect to income, distance, cost and time.
Covariate parameters	ξ	These parameters represent the interaction of VTT with various features of the traveller and trip.
dBF parameters	β, γ, η	These parameters represent various effects of size and sign on VTT.

4.3 SP data cleaning and diagnostics

The SP data from the field surveys has been subjected to detailed examination in order to establish its quality and reliability for modelling. Data 'cleaning' has been conducted systematically and in a consistent way across the dataset; where a concern has been identified for a given individual, data for that individual has been removed for all games in which he/she participated.

It is important to clarify that this process covered all SP data, giving rise to the core SP models described in this chapter and the auxiliary SP-based models described in **Chapter 5**. The one exception to this is the RP rail operator choice model, which entailed different cleaning criteria, as explained in **Section 5.3.1**.

For audit purposes, a detailed record of each person eliminated from the analysis has been kept. Our overall approach to data cleaning was that records were removed only when absolutely necessary and with a view to avoiding bias. In what follows, we outline two SP cleaning scenarios which we tested. The first scenario involved 'basic' cleaning, and the second involved 'additional' (i.e. more intensive) cleaning.

4.3.1 Basic cleaning criteria

Basic cleaning involved the application of the following criteria, for all modes and purposes:

- **Eliminating those individuals who reported a zero travel time or zero travel cost for their current trip:** These levels were incompatible with the research methods adopted for the study and would have prevented the presentation of meaningful choice scenarios, especially for SP1.
- **Checking the distance measure and ratios (i.e. miles per hour, £ per mile):** Whilst the Google maps tool was largely effective in recording distance data for the reference trip, we encountered some complications where the trip involved multiple stages. In such cases, the door-to-door trip distance recorded by the Google maps tool could be an inaccurate indicator of the distance of trip leg which formed the basis of the SP games. Against this background, some individuals were excluded on the basis of extreme absolute distances.
- **Applying rule-based mode-specific thresholds for the key variables:** These thresholds (shown in **Other covariates:** it was decided that the few people who did not report age or gender should be excluded from the analysis. For employees' business, the few people who answered *"no"* to the question *"are you in paid employment or self-employed"* were also removed since, as a consequence of this response, they were not presented with the specific business travel questions. It should be noted that the need for this criterion was identified (and implemented) relatively late in the modelling work, and did not form part of the tests reported in Table 4.2. In total, this criterion excluded 10 people for car, 17 for rail and 9 for 'other PT'.
- **Table 4.2** below, helped to protect the quality of the dataset, at the expense of only a modest number of exclusions. We checked that the imposition of these thresholds did not unduly affect the substantive results for any mode, purpose or game.
- **Additional exclusions for 'extreme' absolute time and cost:** commuting trips of over 3 hours were removed, and due to their more 'local' nature, bus and 'other PT' trips for any purpose lasting 2 hours or more were removed. Rail commuting costs over £80 (one way) were also removed, where these were likely caused by incorrectly entered ticket types or confusion about season ticket costs.
- **Other covariates:** it was decided that the few people who did not report age or gender should be excluded from the analysis. For employees' business, the few people who answered *"no"* to the question *"are you in paid employment or self-employed"* were also removed since, as a consequence of this response, they were not presented with the specific business travel questions. It should be noted that the need for this criterion was identified (and implemented) relatively late in the modelling work, and did not form part of the tests reported in **Table 4.2.** In total, this criterion excluded 10 people for car, 17 for rail and 9 for 'other PT'.

Table 4.2: Basic cleaning criteria

	DISTANCE (miles)	TIME (hours)	COST (£)	£ per hour
Car				
Lower bound	<0.5	-	<0.10	<2
Upper bound	>700	-	-	>40
Rail				
Lower bound	<0.5	-	<1	-
Upper bound	>700	-	>200	>120
Bus				
Lower bound	-	-	<0.50	<0.50
Upper bound	>700	-	>30	>50
'Other' PT				
Lower bound	-	-	<0.50	<0.50
Upper bound	>700	-	>10	>50

Throughout the cleaning process, models were estimated with and without potential exclusions. This reassured us that the relevant exclusions improved the ability of models to explain the behavioural responses, or at the very least did not reduce it, and did not compromise the validity of the substantive results.

4.3.2 Additional cleaning criteria

As a means of comparison against the basic cleaning criteria scenario outlined above, we experimented with additional constraints on the relationship between distance and the base levels of cost and time in the data; that is to say, one might expect the ratios £/mile and miles/hour to be within 'reasonable' levels. More specifically, the additional cleaning criteria involved the cleaning thresholds given in **Table 4.3** *in addition* to those given in **Table 4.2**.

Table 4.3: Additional cleaning criteria

	£ per mile	Miles per hour
Car		
Lower bound	<0.02	<5
Upper bound	>1	>75
Rail		
Lower bound	-	-
Upper bound	-	>120
Bus		
Lower bound	-	<5
Upper bound	>1	-
'Other PT'		
Lower bound	-	<5
Upper bound	>1	-

4.3.3 Comparative testing of the cleaning criteria

Table 4.4 summarises, by mode, the sample sizes resulting from application of the two alternative cleaning criteria. In addition, it summarises model results obtained from the SP1 data, using the multiplicative specification outlined later in this section, where we relied on SP1 for these tests given the generic nature of these experiments across modes. For model outputs, we give an indicative model fit measure (in the form of the adjusted ρ^2) along with the estimated value of travel time (VTT) from a model without additional covariates. The table also includes a separate column for 'usable data', which is the data for trips with positive values for cost, time and distance.

Table 4.4: Comparison of the two sets of cleaning criteria

Mode		Raw data	Usable data	Basic cleaning	Additional cleaning
Car	Individuals	3017	2950	2816	2552
	Adj. ρ^2		0.153	0.157	0.176
	VTT (£/h)		7.95	7.92	8.12
Rail	Individuals	3044	2961	2788	2727
	Adj. ρ^2		0.197	0.209	0.213
	VTT (£/h)		11.74	11.69	11.76
Bus	Individuals	945	867	836	674
	Adj. ρ^2		0.167	0.180	0.201
	VTT (£/h)		1.71	1.78	1.96
'Other PT'	Individuals	1397	1313	1252	1117
	Adj. ρ^2		0.229	0.242	0.246
	VTT (£/h)		3.8	3.64	3.65

As can be seen from **Table 4.4**, the additional cleaning results in substantially more individuals being dropped from the analysis. As might be expected, it also leads – in some cases (e.g. car) at least – to an improvement in model fit. Reassuringly, however, the more intensive cleaning does not lead to statistically significant changes in the VTT.

Overall, we judged that the gains in fit and the variations in VTT estimates were not so significant as to justify the additional cleaning criteria, and we proceeded therefore with only basic cleaning. On this basis, the mode most affected by cleaning was bus, with slightly over 11.5% individuals being removed from the analysis. 'Other PT' followed with 10.3% removal, whilst for car and train only 8.4% and 6.6% individuals respectively were dropped from the analysis. In general, the most common cleaning factors were missing distance information (the reason for 206 out of the total 711 exclusions across all modes) and missing cost information (the reason for 88 exclusions). For bus, the main reason for exclusion was due to cost being below 50 pence (73 respondents, out of which the vast majority reported a zero cost). Reported cost was also the key reason for exclusion in 'other PT': 37 cases with cost below 50 pence, and 38 cases with cost greater than £10, which were considered unrealistic. For car and rail, the main factor was distance (zero or below 0.5 miles for over 100 respondents per mode), followed by unrealistic commutes of over 3 hours (68 cases in car and 38 in rail). Rail costs either below £1 or over £80 for one way were the other major causes in rail. Overall, most of the exclusions driven by unrealistic cost also implied an unrealistic cost/time ratio (in £/hour), which would have made the SP choice scenarios for that respondent unrealistic.

4.3.4 Diagnostics of respondent behaviour

Diagnostic tests were performed on the data relating to a number of behavioural traits that can adversely affect model estimation[33]. The results of this process are summarised in **Table 4.5**. As with the cleaning process described above, it should be clarified that these tests refer to all of the SP data, but not the RP.

With our main focus in the analysis being on the estimation of joint models across the three SP games, the four main measures are calculated using responses from all games combined (i.e. 15 SP tasks per respondent if the three SP games are used; 10 SP tasks if only SP1 and SP2 are used for estimation). Our observations for these four measures are as follows:

- **Rate of left or right non-traders:** the proportion of people who always chose the option presented on the left (or always choose the one on the right) across all SP choice tasks presented is negligible.
- **Rate of time non-traders:** for most cases, only 1-2% of the respondents consistently chose the fastest travel option. For bus, the rate is even lower (being zero for non-work travellers), while only for 'other PT' for employees' business do we see a value exceeding 3%. These rates are sufficiently low to confirm that the ranges presented in the survey were wide enough, and offers reassurance regarding the ability to estimate the tail of the VTT distribution in Mixed Logit models[34].

Table 4.5: Data diagnostics

		Commute	Employees' business	Other non-work
CAR	Number of respondents	922	917	977
	Number of included choices	13,830	13,755	14,655
	Rate of left non-traders (combined SPs)	0.54%	0.22%	0.10%
	Rate of right non-traders (combined SPs)	0.11%	0.44%	0.20%
	Rate of time non-traders (combined SPs)	1.52%	1.20%	0.82%
	Rate of cost non-traders (combined SPs)	5.21%	4.03%	5.73%
	Rate of inconsistent traders (SP1)	26.46%	18.54%	23.13%
TRAIN	Number of respondents	847	945	996
	Number of included choices	12,340	13,390	14,275
	Rate of left non-traders (combined SPs)	0.12%	0.63%	0.20%
	Rate of right non-traders (combined SPs)	0.00%	0.00%	0.00%
	Rate of time non-traders (combined SPs)	2.24%	2.22%	1.00%
	Rate of cost non-traders (combined SPs)	8.26%	5.71%	7.93%

[33] Hess, S., Rose, J.M. and Polak, J.W. (2010) 'Non-trading, lexicographic and inconsistent behaviour in stated choice data', Transportation Research Part D, 15(7), pp405-417.

[34] Börjesson, M., Fosgerau, M. and Algers, S. (2012) 'Catching the tail: Empirical identification of the distribution of the value of travel time', Transportation Research. Part A: Policy and Practice, Vol. 46, No. 2, pp378-391.

	Rate of inconsistent traders (SP1)	17.95%	19.47%	19.48%
BUS	Number of respondents	334	20	482
	Number of included choices	5,010	300	7,230
	Rate of left non-traders (combined SPs)	0.90%	0.00%	0.21%
	Rate of right non-traders (combined SPs)	0.00%	0.00%	0.00%
	Rate of time non-traders (combined SPs)	0.30%	0.00%	0.00%
	Rate of cost non-traders (combined SPs)	3.29%	5.00%	8.71%
	Rate of inconsistent traders (SP1)	26.65%	35.00%	22.82%
'OTHER PT'	Number of respondents	565	206	481
	Number of included choices	7,030	2,515	6,115
	Rate of left non-traders (combined SPs)	0.35%	0.49%	0.42%
	Rate of right non-traders (combined SPs)	0.00%	0.97%	0.42%
	Rate of time non-traders (combined SPs)	1.77%	3.40%	0.83%
	Rate of cost non-traders (combined SPs)	5.66%	4.85%	6.24%
	Rate of inconsistent traders (SP1)	20.35%	16.02%	17.67%

- **Rate of cost non-traders:** the proportion of people consistently choosing the cheapest option around all tasks is around 5% in most samples, being always between 3% (bus commuters or car business travellers) and 8% (train commute and non-work, and bus non-work). These rates are lower than in many other studies and support the ranges presented in the trade-offs.
- **Rate of inconsistent traders:** this diagnostic measure is only applicable to SP1 responses. 'Consistent' behaviour would imply that all boundary value of time measures rejected by a respondent (i.e. those trade-offs where the respondent chooses the cheaper option) are higher than those boundary value of time measures accepted by the respondent (i.e. those trade-offs where the respondents chooses the faster option). A respondent is inconsistent if at least one accepted boundary value is higher than the lowest rejected boundary value. It should be noted that any such inconsistencies may be related to the sign or size of the travel time change, or the sign or size of the travel cost change, which are plausible sources of intra-individual heterogeneity in valuation; that is, whether a time or cost difference reflects a loss or gain or is large or small may cause differences in the valuation that may give rise to apparently inconsistent behaviour. This ratio is around 20% on average, being generally lower for the train and 'other PT' segments and higher for the bus and car segments. This is not too dissimilar from what has been observed for other datasets of the same type.

4.4 Generic model structure assumptions

The field of choice modelling has evolved substantially since the 2003 national value of time study was conducted in the UK. The present study makes use of many of these developments, relating to error structure of the models, the treatment of reference dependence (size and sign effects) and the incorporation of flexible random heterogeneity in valuations. These developments are addressed in turn in the following sub-sections. This specific section is concerned with a discussion of the error structure of the models.

4.4.1 Multiplicative vs additive error structures

The utility in a choice model is decomposed into a deterministic and a random component, the error term. The models used in the 1994 and 2003 VTT studies made use of standard additive error structures, used in most VTT studies world-wide since the pioneering UK work[35], where $U = V + \varepsilon$, with V and ε giving the deterministic and random components of utility, respectively.

In the present study, we move away from this assumption by relying on models using a multiplicative formulation[36]. The multiplicative formulation represents the state-of-the-art in VTT estimation for experiments of the SP1 type, and the advantages have also been verified by tests made for this study, as reported below. A corresponding approach for SP2 and SP3 is also possible, and used here, in common with the recent Danish national VTT study, but further developed here for reference dependence.

In a multiplicative formulation, we replace the typical additive specification of the utility of an alternative $U = V + \varepsilon$ by $U = V.\varepsilon$, where V and ε are still defined as the deterministic and random components of utility, respectively. That is, the random (error) component of utility is taken to multiply the deterministic component, rather than be added to it.

The practical advantage given by the multiplicative approach is that it becomes much easier to make an assumption of constant variance for ε. In general, it is found that utility variance increases as utility increases and this is handled automatically in the multiplicative form of the model. This benefit is confirmed by the improved results given by multiplicative models in this context.

In the multiplicative model, it is practical[37] to work with $\log U = \log V + \varepsilon$, the log function having no impact on the ranking of utilities, since it is a monotonic transformation. Technically, the assumptions regarding the distributions of ε are different in these cases. In practice, it is assumed that ε follows a Gumbel distribution in the multiplicative model as in the additive model, so that the simple logit model can be used to calculate probabilities.

4.4.2 Multiplicative specification for SP1

For SP1, the analysis is quite simple because there are just two attributes: time and cost. For this reason, it is possible to formulate the econometric model in a multiplicative log willingness-to-pay form, as shown below. This formulation has proved successful in several previous studies and has been shown to outperform the other alternatives tested for the present data.

With this specification, the error in the models is then proportional to the boundary value of time, i.e. the trade-offs faced by respondents. This is consistent

[35] Daly, A. and Zachary, S. (1975) 'Commuters' Values of Time', Report to Department of the Environment, LGORU report T55, January 1975.

[36] See Harris, A.J. and Tanner, J.C. (1974) 'Transport demand models based on personal characteristics', Transport and Road Research Laboratory Supplementary Report SR65UC, Crowthorne, UK and Fosgerau, M. and Bierlaire, M. (2009) 'Discrete choice models with multiplicative error terms', Transportation Research Part B, 43, pp494-505.

[37] It is not difficult in practice to arrange that $V > 0$ and $\varepsilon > 0$.

with the notion that the main source of error in simple time vs. money trade-offs would be unexplained heterogeneity in VTT measures, which would thus lead to larger error in scenarios where the value of time required to choose the expensive option is larger.

For ease of implementation, the model is estimated using an additive structure, where a logarithm applied to the deterministic utility incorporates the heteroskedasticity. Specifically, we now use (avoiding for now additional subscripts for respondents and choice tasks):

$$V_s = \mu_{SP1}.\log\left(-\frac{cost_1 - cost_2}{time_1 - time_2}\right) \quad (4.1)$$

$$V_e = \mu_{SP1}.\log \omega \quad (4.2)$$

where the alternatives are reordered so that V_s gives the utility of the cheaper but slower option and V_e gives the utility of the faster but more expensive option, ω is the estimated VTT (for reductions in time, i.e. a positive value of time) and μ_{SP1} is an estimated scale parameter to accommodate the error in the model. Comparing these utilities, we see that if the VTT is greater than the 'bid', i.e. the implied boundary VTT presented to the respondent, then the expensive option will be chosen, otherwise the cheap option is chosen.

We now have that the probability of the observed sequence of T ($t = 1, \dots T$) choices for person n for SP1 is given by the product of logit probabilities:

$$P_{SP1,n} = \prod_{t=1}^{T}\left(\frac{e^{V_{snt}}}{e^{V_{snt}}+e^{V_{ent}}}\right)^{\delta_{snt,SP1}}\left(\frac{e^{V_{ent}}}{e^{V_{snt}}+e^{V_{ent}}}\right)^{\delta_{ent,SP1}} \quad (4.3)$$

where $\delta_{snt,SP1}$ is 1 if and only if the slow option is chosen by respondent n in task t, with a corresponding definition applying for $\delta_{ent,SP1}$. Only one of these indices is 1 in any given choice task, of course.

It is important to note that, with a Multinomial Logit (MNL) specification of the model, the use of ω alone as an estimate of the VTT is likely to underestimate the true mean VTT. Indeed, in the model above, the error term likely captures not just noise but also heterogeneity in the VTT (given that this model works in relative valuations). As such, for an MNL model, ω relates more to a median than a mean VTT, a point explained in more detail in **Section 4.4.4** below. Given the reliance on Mixed Logit models later in the analysis, where we explicitly incorporate random heterogeneity in VTT rather than rely on the error term to capture it, this issue disappears as the additional random components distributed across respondents then capture the random variation in VTT.

4.4.3 Multiplicative specification for SP2 and SP3

For SP2 and SP3, log willingness-to-pay is not a feasible approach as there are multiple attributes and the signs of differences from the reference value are not consistent. For these games, therefore, the model used is a multiplicative utility model. This formulation has previously been used in the most recent Danish study for games of this type. In SP2 and SP3, the error is then proportional to the overall utility, which essentially means the model implies greater error on longer trips.

In an additive model expressed in what is referred to as preference space, we would have that:

$$V_j = \tau_{cost} cost_j + \sum_k \tau_k x_{jk}, \quad (4.4)$$

where x_{jk} refers to *K* different non-cost attributes for alternative *j*, and where the τ parameters are estimated marginal utilities. The τ parameters would be negative for undesirable attributes, and positive for desirable attributes. This can be rewritten for a mathematically equivalent specification[38] in valuation space as:

$$V_j = -\mu_{SP2}\left(cost_j + \sum_k \omega_k x_{jk}\right), \quad (4.5)$$

where μ_{SP2} is now a positive scale parameter (using here SP2 as the example), and where ω_k is now a directly estimated monetary valuation for changes in x_{jk}[39]. The negative sign on the entire utility now means that the ω_k are positive for undesirable attributes, i.e. they relate to a willingness-to-pay for avoiding positive changes in an attribute. This is appropriate for time, but less intuitive for other attributes.

It then makes sense to replace $\sum_k \omega_k x_{jk}$ by $\sum_{k_T} \omega_{k_T} x_{jk_T} - \sum_{k_{NT}} \omega_{k_{NT}} x_{jk_{NT}}$, where x_{jk_T} are time attributes and $x_{jk_{NT}}$ are non-time attributes. For our analysis, we include attributes such as travel time variability, headway and delays in x_{jk_T}.

In the multiplicative model, we would then use:

$$V_j = -\mu_{SP2} \cdot \log\left(cost_j + \sum_{k_T} \omega_{k_T} x_{jk_T} - \sum_{k_{NT}} \omega_{k_{NT}} x_{jk_{NT}}\right) \quad (4.6)$$

where ω_{k_T} remains the directly estimated monetary value (i.e. willingness-to-pay) for reductions in time components, and $\omega_{k_{NT}}$ is the directly estimated monetary value for increases in non-time components $x_{k_{NT}}$. The use of the log transform enables the estimation of the multiplicative model using an additive software specification.

Again using a Gumbel distribution for ε, the probability for the observed sequence of choices for respondent *n* in say SP2 is now given by:

$$P_{SP2,n} = \prod_{t=1}^{T} \prod_{j=1}^{2} \left(\frac{e^{V_{jnt}}}{e^{V_{1nt}} + e^{V_{2nt}}}\right)^{\delta_{jnt,SP2}} \quad (4.7)$$

where $\delta_{jnt,SP2}$=1 if and only if alternative *j* is chosen by respondent *n* in task *t* in SP2. Analogous equations apply for SP3.

4.4.4 Empirical comparisons between multiplicative and additive specifications

To justify the use of the multiplicative modelling approaches, we now present a brief empirical comparison between additive and multiplicative models for the

[38] Train, K.E. and Weeks, M. (2006) ‘Discrete choice models in preference space and willingness-to-pay space’, in R. Scarpa and A. Alberini (eds), ‘Applications of simulation methods in environmental and resource economics’, Springer.

[39] Of course $\mu_{SP2} = -\tau_{cost}$ and $\omega_k = \tau_k / \tau_{cost}$.

data collected for the present analysis. These comparisons are conducted independently for each of the three main SP games, where, for SP3, we focus on the crowding games for public transport for this presentation. The models used here are free of any covariates, as no separate specification search was conducted for the additive models.

While the multiplicative models give direct estimates of WTP measures, the additive models were estimated in preference space, hence giving estimates for the component marginal utilities. These are then used to calculate WTP measures, by taking ratios against the cost coefficients, and standard errors for the WTP measures are computed using the Delta method. This is known to give the exact same result as working in preference space, both for the estimates and the standard errors, using MNL models[40].

SP1

For SP1, the models make use of two parameters, a time and cost coefficient for the additive model, and a scale parameter and estimated VTT for the multiplicative model. In our comparison, we focus on two key measures, namely model fit and the implied VTT measures. For the former, we present both the log-likelihood as well as the adjusted ρ^2 measure. For the VTT, it should once again be noted that, with an MNL structure, the SP1 estimates for the multiplicative models, i.e. ω, do not themselves take into account the error distribution and are thus effectively medians rather than means. On the basis of the assumption that the main source for error in the log-WTP space model is unexplained heterogeneity in the VTT, we can calculate:

$$VTT = e^{\log \omega + \frac{1}{\mu_{SP1}(\varepsilon_1 - \varepsilon_2)}} \quad (4.8)$$

where ε_1 and ε_2 follow a Gumbel distribution. This assigns the entire error term of the model to the VTT distribution and leads to a very long tail. In the presentation of the results for these tests, we have censored the distribution of VTT at a value of £50/hr. This is a rather arbitrary assumption made here for the sake of presentation and comparability across games, but gives a good coverage of the overall distribution of boundary trade-offs presented across the SP1 games.

Looking at the results (**Table 4.6**), it is clear that, for all four modes, the multiplicative model obtains substantially better fit with the same number of parameters. Across all four modes, the multiplicative model also obtains much smaller standard errors[41] for the VTT measures, which, even with the lower median values in the multiplicative models, still lead to higher t-ratios. It is worth highlighting here that, for bus, the additive model fails to produce a significant

[40] See Daly, A.J., Hess, S. and de Jong, G. (2012) 'Calculating errors for measures derived from choice modelling estimates', Transportation Research Part B 46(2), pp333-341, and Daly, A.J., Hess, S. and Train, K.E. (2012) 'Assuring finite moments for willingness-to-pay in random coefficients models', Transportation 39(1), pp19-31, for points on the equivalence of the standard errors and estimates, respectively.

[41] Model estimates are accompanied by 'robust' standard errors. These are obtained using the 'sandwich' estimator and account for general model specification. When comparing fit for two models estimated on the same data, and where one model is a more general version of the other one, a likelihood ratio test was used. For comparisons of non-nested models (i.e. where it is not the case that one model is more general than the other), the adjusted ρ^2 measure can be used.

estimate of the VTT. Using (4.8), and censoring the resulting distribution at £50/hr, a point which captures the majority of presented SP1 trade-offs, we obtain higher means (denoted $\overline{VTT}$ in the table) than medians in the multiplicative model, reflecting the skewed distribution.

Table 4.6: Comparison of additive and multiplicative models for SP1

	Car				**Rail**					
	Additive		Multiplicative		Additive		Multiplicative			
Respondents	2,816				2,788					
Observations	14,080				13,940					
Final LL	-8,989.97		-8,803.59		-8,526.53		-8,320.27			
Adj. ρ^2	0.0786		0.0977		0.1174		0.1387			
	est.	t-rat.	est. (ω)	t-rat.	$\overline{VTT}$	est.	t-rat.	est. (ω)	t-rat.	$\overline{VTT}$
VTT (£/hr)	7.57	23.30	4.84	26.30	8.07	11.64	25.68	7.94	30.16	10.31
	Bus				**'Other PT'**					
	Additive		Multiplicative		Additive		Multiplicative			
Respondents	836				1,252					
Observations	4,180				6,260					
Final LL	-2,494.13		-2,384.89		-3,474.36		-3,336.49			
Adj. ρ^2	0.1385		0.1762		0.1988		0.2306			
	est.	t-rat.	est. (ω)	t-rat.	$\overline{VTT}$	est.	t-rat.	est. (ω)	t-rat.	$\overline{VTT}$
VTT (£/hr)	0.34	0.61	1.46	13.42	4.43	4.66	22.53	3.10	25.03	5.98

SP2

For SP2, the additive models estimated travel time coefficients across all modes, along with coefficients for the standard deviation of travel time for car, and the mean lateness and standard deviation of lateness for public transport modes. In the multiplicative models, the corresponding monetary valuations were directly estimated, i.e. the VTT, the value of the standard deviation of travel time (VSDTT) for car, and the value of lateness (VLATE) and the standard deviation of lateness (VSDLATE) for public transport. For car, we can then calculate a reliability ratio (relative sensitivity to standard deviation of travel time and mean travel time), while, for public transport, we can calculate a lateness multiplier (value of mean lateness vs. value of time). Notwithstanding the existence of theoretical relationships between the scheduling and mean-variance approaches[42], we adopt here the specifications which directly produce the outputs used in current UK appraisal for different modes, and these differ between public and private transport.

With reference to **Table 4.7** we again see that, across modes, the multiplicative model obtains substantially better model fit with the same number of parameters.

[42] Fosgerau, M. and Karlström, A. (2009) 'The value of reliability'. Transportation Research Part B, 44 (1), pp38-49.

In terms of actual valuations coming out of the models, we note the lower VTT in the multiplicative models for car and rail, as well as relatively greater sensitivity to travel time variability. For bus and ‘other PT’, the VTT measures are comparable between additive and multiplicative, but the valuations for reliability are lower in the latter. The actual source of the differences between the two structures in terms of valuations are difficult to explain with certainty, but a likely contributing factor is that the unexplained heteroskedasticity in the additive model leads to bias in the parameter estimates. Across the different modes and valuations, the t-ratios tend to be higher in the multiplicative models, in line with the SP1 findings.

Table 4.7: Comparison of additive and multiplicative models for SP2

	Car				**Rail**			
	Additive		Multiplicative		Additive		Multiplicative	
Respondents	2,816				2,788			
Observations	14,080				13,940			
Final LL	-8,770.84		-8,539.86		-8,953.74		-8,481.61	
Adj. ρ^2	0.1010		0.1247		0.0729		0.1218	
	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
VTT (£/hr)	14.47	12.55	7.45	22.79	14.35	20.71	7.68	23.43
VSDTT (£/hr)	4.03	9.14	2.49	11.80	-	-	-	-
VLATE (£/hr)	-	-	-	-	31.69	11.85	18.21	15.30
VSDLATE (£/hr)	-	-	-	-	26.99	9.56	9.61	9.09
reliability ratio	0.28		0.34		-		-	
lateness multiplier	-		-		2.21		2.37	

	Bus				**‘Other PT’**			
	Additive		Multiplicative		Additive		Multiplicative	
Respondents	836				1,252			
Observations	4,180				6,260			
Final LL	-2,650.41		-2,569.73		-3,713.21		-3,658.10	
Adj. ρ^2	0.0839		0.1117		0.1433		0.1560	
	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
VTT (£/hr)	2.25	7.17	2.16	9.35	3.70	16.71	3.96	16.16
VSDTT (£/hr)	-	-	-	-	-	-	-	-
VLATE (£/hr)	7.86	7.88	6.60	9.78	9.28	13.57	8.96	14.29
VSDLATE (£/hr)	3.01	4.29	3.26	5.98	4.17	6.78	3.70	7.45
reliability ratio	-		-		-		-	
lateness multiplier	3.49		3.06		2.51		2.26	

SP3

For SP3, we obtain estimates of three types of time components (free-flow; ff, light traffic; lc, and heavy traffic; hc) for car, while for public transport, we obtain valuations of travel in up to 10 different crowding conditions (VCROW1 to VCROW10).

Table 4.8: Comparison of additive and multiplicative models for SP3

	Car				**Rail**			
	Additive		Multiplicative		Additive		Multiplicative	
Respondents	2,816				2,425			
Observations	14,080				12,125			
Final LL	-8,608.02		-8,238.82		-8,136.51		-7,636.80	
Adj. ρ^2	0.1176		0.1554		0.0306		0.0900	
	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
VFF (£/hr)	2.30	4.26	1.96	9.09	-	-	-	-
VLC (£/hr)	5.03	9.93	3.37	14.71	-	-	-	-
VHC (£/hr)	11.89	14.71	6.96	21.58	-	-	-	-
VCROW1 (£/hr)	-	-	-	-	8.02	4.87	2.16	5.73
VCROW2 (£/hr)	-	-	-	-	5.60	2.85	1.36	3.49
VCROW3 (£/hr)	-	-	-	-	9.21	4.71	3.24	7.64
VCROW4 (£/hr)	-	-	-	-	9.54	4.53	3.31	7.31
VCROW5 (£/hr)	-	-	-	-	9.54	3.89	3.62	6.90
VCROW6 (£/hr)	-	-	-	-	-0.48	-0.13	2.17	4.26
VCROW7 (£/hr)	-	-	-	-	-8.43	-1.65	1.78	3.28
VCROW8 (£/hr)	-	-	-	-	-13.83	-2.21	2.02	3.41
VCROW9 (£/hr)	-	-	-	-	-18.99	-2.45	2.04	3.27
VCROW10 (£/hr)	-	-	-	-	-17.06	-2.20	2.62	3.52
	Bus				**‘Other PT’**			
	Additive		Multiplicative		Additive		Multiplicative	
Respondents	419				628			
Observations	2,095				3,140			
Final LL	-1,373.37		-1,299.72		-2,010.17		-1,946.90	
Adj. ρ^2	0.0501		0.1008		0.0737		0.1027	
	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
VCROW1 (£/hr)	1.42	2.53	1.02	4.32	2.53	10.01	2.75	10.90
VCROW2 (£/hr)	1.32	2.28	1.03	4.26	2.54	9.77	2.72	11.12
VCROW3 (£/hr)	1.37	2.24	1.03	3.86	2.60	10.79	2.67	11.36
VCROW4 (£/hr)	1.84	2.94	1.44	4.96	3.07	12.29	2.98	12.14
VCROW5 (£/hr)	3.28	4.31	2.90	6.36	4.49	14.13	4.87	13.58

For each mode (**Table 4.8**), we again see substantial gains in fit for the multiplicative models, with the same number of parameters. For car, the range in valuations across the three levels of congestion becomes narrower and arguably more realistic, given the ratio of over 5 in the additive model. We also see increases in statistical robustness. For rail, we note that while there are some issues with a non-monotonic trend for the multiplicative model (something addressed later on in this chapter), the valuations are all of the correct sign, unlike for the additive model, showing the benefit of capturing the heteroskedasticity. For bus and 'other PT', the results are quite consistent between the two models, notwithstanding the big difference in fit. We again generally observe higher t-ratios for the multiplicative models.

Conclusions on error structure

The findings from this empirical comparison between the additive and multiplicative models confirm the clear advantages of the latter in terms of fit as well as generally more reasonable and robust estimates. The remainder of the analysis therefore relies on multiplicative specifications.

4.5 Additional considerations for SP2 and SP3

We have already discussed the specification of the utility functions for SP1 in detail. For SP3, the components inside the logarithmic term, i.e.

$$cost_j + \sum_{k_T} \omega_{k_T} x_{jk_T} - \sum_{k_{NT}} \omega_{k_{NT}} x_{jk_{NT}} \quad (4.9)$$

relate simply to the individual attributes.

For SP2 however, an additional complication arises as the respondents are presented, for each alternative, with five different possible outcomes in terms of travel time (with constant travel cost). For car, we simply work with the mean travel time and the standard deviation in (4.9). However, for public transport, we work with the *"usual travel time"* as well as valuations for early and for late arrivals. It should be noted that this is different from early and late schedule delay as the comparison is not with some preferred arrival time, but with a published arrival time, and thus relates to increases and decreases in travel time.

We would thus get the value for outcome 1 (out of 5) for alternative j, say $\chi_{j,1}$ is given by:

$$\chi_{j,1} = cost_j + \omega_{time} time_j + \omega_{early} earlyness_{j,1} + \omega_{late} lateness_{j,1} - \omega_{late\ penalty} late_{j,1} \quad (4.10)$$

where $time_j$ is the usual travel time, $earlyness_{j,1}$ and $lateness_{j,1}$ relate to the amount of early or late delay in outcome 1 for alternative j, and $late_{j,1}$ is equal to 1 if the first outcome has late arrival, and 0 otherwise. The estimates for ω_{time}, ω_{early} and ω_{late} are value of time measures, which we expect to be positive for ω_{time} and ω_{late}, as reductions would imply shorter travel time, while $\omega_{late\ penalty}$ is the willingness-to-pay for late arrival, which we expect to be negative. For ω_{early}, we also expect a negative estimate as reductions in earliness would mean longer travel time. There should be no strong *a priori* expectation that the estimates of ω_{early} and ω_{time} are simple opposites, given non-linearities in sensitivities, but also given the potential for different behavioural responses to scheduled travel time and unscheduled early arrival.

On the assumption of equal weight being given to the five outcomes, we can then use:

$$V_j = -\mu_{SP2} \cdot \log\left(\sum_s \frac{\chi_s}{5}\right) \quad (4.11)$$

where s is an index over the five possible outcomes. Building on the work of Liu and Polak[43], we allow for differential weights for the five outcomes using the constant absolute risk aversion (CARA) specification, and instead specify the model as:

$$V_j = -\mu_{SP2} \cdot \log\left(\sum_s \frac{1-e^{-\alpha\chi_1}}{\alpha}\frac{1}{5}\right) \quad (4.12)$$

where, with α approaching zero, we get a risk neutral model corresponding to the earlier specification. If α becomes positive, bad outcomes are valued more strongly and good outcomes less strongly (i.e. risk averseness), with the opposite applying for negative α.

A similar specification for car was not possible in the context of wanting to include a standard deviation of travel time (which obviously does not apply at the individual outcome level), as this is a required output of the analysis.

4.6 Treatment of size and sign effects

Many SP-based VTT studies, including the previous UK study, have found that the values obtained depend on the sign and size of time and cost changes *relative* to a 'reference' value. These findings can be related to Prospect Theory, e.g. that gains are attributed a lower absolute value than equivalent losses[44]. When travellers are interviewed relative to a specific trip, the reference value for a given attribute is often and reasonably taken to be the corresponding value on that trip. In the present study, reference-dependent effects have been tested for all of the SP models, but not for the RP models, since in the latter case it is unclear what constitutes the reference value. Similarly, reference-dependent effects could not be included for those attributes in the SP survey where no immediate reference values were available, such as reliability, or where the number of possible effects due to reference dependence would be too large to test efficiently and too difficult to implement in model application, such as with crowding.

We have adopted the principles of the approach to modelling reference dependence set out by de Borger and Fosgerau[45] (dBF), which appears to be the most sophisticated practical method to date, and have further developed it for this study. We will first discuss the general approach and explain the implications for VTT calculations before turning to the implementation of the approach for the different games.

---

[43] Liu, X. and Polak, J.W. (2007) 'Nonlinearity and the specification of attitudes towards risk in discrete choice models', Transportation Research Record 2014, pp27-31.

[44] Kahneman, D. and Tversky, A. (1979) 'Prospect theory: an analysis of decision under risk, Econometrica', 47 (2), pp263-291.

[45] de Borger, B. and Fosgerau, M. (2008) 'The trade-off between money and travel time: A test of the theory of reference-dependent preferences', Journal of Urban Economics, Vol. 64, pp101-115.

4.6.1 Notation and degrees of freedom

In a binary choice, for an attribute x, define the differences:

$$\Delta x_1 = x_1 - x_0, \quad (4.13)$$

$$\Delta x_2 = x_2 - x_0, \quad (4.14)$$

$$\Delta x = x_1 - x_2 \quad (4.15)$$

where the subscript 0 refers to the base (current) value and 1 and 2 refer to the alternatives presented to the respondent. Note that $\Delta x \equiv \Delta x_1 - \Delta x_2$.

Then we define increases and decreases:

$$\Delta x_1^+ = \max(0, \Delta x_1) \quad (4.16)$$

$$\Delta x_1^- = \min(0, \Delta x_1) \quad (4.17)$$

$$\Delta x_2^+ = \max(0, \Delta x_2) \quad (4.18)$$

$$\Delta x_2^- = \min(0, \Delta x_2) \quad (4.19)$$

Note that $\Delta x_1^+ + \Delta x_1^- = \Delta x_1$ and similarly for alternative 2, so that $\Delta x \equiv \Delta x_1^+ + \Delta x_1^- - \Delta x_2^+ - \Delta x_2^-$.

When we compare alternatives, we need to define the difference in increases and decreases between the alternatives:

$$\Delta x^+ = \Delta x_1^+ - \Delta x_2^+ \quad (4.20)$$

$$\Delta x^- = \Delta x_1^- - \Delta x_2^- \quad (4.21)$$

Note that $\Delta x \equiv \Delta x^+ + \Delta x^-$.

Considering the degrees of freedom, for a given choice task, there are just three measurements of x: the current value and the presented values. Any attempt to introduce more than three variables by calculating differences etc. is liable to fail by introducing over-specification. Note also that raising such variables to powers does not help, because the power 1 has to be part of the domain of estimation and the risk of over-specification is not avoided. The equivalences above point to some of the potential pitfalls.

In a model of SP, we wish to introduce the concept of gains and losses, as well as the basic difference between the alternatives. The question arises as to whether we want to measure differences from the base value or differences between the alternatives. The discussion above, and in particular the last equivalence, show that we cannot include both of these differences.

A natural approach would be to model the value (or marginal value) of an alternative by x_0, Δx^+ and Δx^-. In this study, the base value x_0 is treated as a background variable, while the changes Δx are treated by allowing for different valuations for positive and negative changes.

The key point here is that it is not possible to attach separate values to gains, losses and differences between the alternatives. In previous studies such as the 2003 UK study and the more recent Scandinavian studies, these effects would have been perfectly confounded as the reference values for both time and cost appeared in every choice tasks. This means that the difference between alternatives is the same as the difference from the reference alternative. Previous studies have interpreted the estimates of the size effects as relating to the differences between the alternatives, but the alternative interpretation would have been equally plausible. Given the vast evidence on the importance of reference dependence, our work focusses on that, rather than on differences between the alternatives. This is further supported by our findings of significant asymmetry in relation to gains and losses in the empirical work reported later on.

4.6.2 Defining the value functions

While the inclusion of reference dependence presents major complexities from the point of view of practical implementation, it is necessary to represent the complexities of response that are observed. If this is not done, then VTT risks being dependent on the experimental design.

The concept here is to introduce non-linear functions that express the possibility that size and sign effects exist. This is done by defining a function that gives the value of a change Δx relative to the reference value x_0 of a given attribute. dBF formulate the value functions as:

$$v(\Delta x) = S(\Delta x).\exp(\eta\, S(\Delta x)).|\Delta x|^{\alpha} \quad (4.22)$$

where $\Delta x = x - x_0$, $\alpha = 1 - \beta - \gamma S(\Delta x)$

$S(\Delta x)$ is the sign function, defined for $\Delta x \neq 0$ by $S(\Delta x) = \Delta x\, /\, (|\Delta x|)$, i.e. it takes the values ± 1 with the same sign as Δx; for convenience we also specify that $S(0) = 0$.

η gives the difference of gain value and loss value from an ‘underlying’ value. It is explicitly assumed by dBF that gains and losses exactly bracket this underlying value. The parameter η measures the sign effect. It is expected that $\eta > 0$, so that the value of losses (increases in Δx) is greater than the value of gains.

β allows the impact of gains and losses to be non-linear. If $\beta > 0$, the marginal value of changes decreases as the change increases, i.e. the value is ‘damped’. This is the main measure of the size effect. Generally we anticipate that β should be larger for cost than for time, so that VTT increases as the changes increase, while small time savings have lower monetary value.

γ allows the non-linearity of value to be different for gains and losses. Essentially, this gives an interaction between the sign and size effects. A negative value for γ would for example mean that any damping (i.e. decreasing marginal effects for larger changes) would be smaller for increases (losses) than for decreases (gains) from the reference value, or that any increasing sensitivity for larger changes (i.e. with $\alpha > 1$) would be stronger for increases than for decreases from the reference value.

The value functions are defined to have arguments denominated in cost units. Thus the value of a cost change Δc is given by $v(\Delta c)$, while the value of a time change t is given by $v(\theta\Delta t)$, where θ is the 'underlying' value of time. The value functions are of the same general form for time and cost (and potentially for other utility components) but the arguments η, β and γ are specific to each utility component. Differently from the dBF work, and from how it was used in the Danish work, we are able to estimate separate arguments for both time and cost also in SP1, as our design does not impose the presence of the reference value for one of the two alternatives in the choice task[46].

While these functions are rather sophisticated, we have chosen to use them for this study on the basis that they have been shown to work in the dBF work (which was based on VTT SP data), that they operate satisfactorily for our own data, and that they incorporate the full size and sign effects that we might wish to investigate.

4.6.3 Deriving the VTT

A simple way to see the derivation of VTT (and other WTP measures) is to think of the values of Δc and Δt that would maintain indifference with the base situation in which $\Delta t = \Delta c = 0$ and the total value is of course zero. Thus when we have a specific value $\Delta t'$, and we have estimated the parameters of the value functions v, we can find the value $\Delta c'$ such that $v(\Delta c') + v(\theta\Delta t') = 0$. The average willingness-to-pay per unit of time is then $\Delta c'/\Delta t'$. Again this follows the ideas of dBF, though they are not clear about this calculation.

It is reasonable to extend the method of dBF in taking the average of the gain value and the loss value to express an 'underlying' VTT. In fact, it is difficult to formulate an alternative: in the SP context we obtain gain values and loss values and these need to be averaged in some way, not least as we do not know whether these are real world or SP effects, or, as likely, a mixture of the two. That is, to obtain a reference-free value we need to calculate the average of the loss value of a given Δx and the gain value of the same Δx to obtain a reference-free value of Δx. As in dBF, we calculate the *geometric* mean[47] of $v(\Delta x)$ and $-v(-\Delta x)$:

$$\sqrt{v(\Delta x).-v(-\Delta x)} = \sqrt{\exp(\eta).|\Delta x|^{1-\beta-\gamma}.\exp(-\eta).|\Delta x|^{1-\beta+\gamma}}$$

$$= \sqrt{|\Delta x|^{2-2\beta}} = |\Delta x|^{1-\beta} \quad (4.23)$$

[46] The DATIV data they used was based on a design in which the current time and cost always appeared in one or other of the alternatives presented. This design has the effect that it is not possible to make separate identifications of β for both time and cost. However, with the new UK data the design is more varied and it is possible to make these identifications.

[47] dBF assumed (unnecessarily) that γ was zero to calculate the geometric mean. It would also be possible to work with the *arithmetic* mean, in which case value functions could be defined as $v(\Delta x) = S(x)|\Delta x|^{1-\beta}(1 + S(x)(\eta + \gamma\Delta x))$, which would give an analogous simplified result.

To obtain a 'reference free' value, this argument suggests it is reasonably simple to omit the asymmetry parameters η and γ. However, there is no analogous argument to eliminate β and the value remains a function of Δx.

Solving the equation $v(\Delta c') + v(\theta \Delta t') = 0$, for the simplified value functions $v(\Delta x) = S(\Delta x)|\Delta x|^{1-\beta}$ we obtain, for oppositely signed Δc and Δt:

$$|\Delta c|^{1-\beta_c} = (\theta|\Delta t|)^{1-\beta_t} \quad (4.24)$$

$$|\Delta c| = (\theta|\Delta t|)^{\frac{1-\beta_t}{1-\beta_c}} = (\theta|\Delta t|)^{\kappa} \quad (4.25)$$

where $\kappa = \frac{1-\beta_t}{1-\beta_c}$, so we can calculate the VTT (per unit of time) as:

$$VTT = {|\Delta c|}/{|\Delta t|} = \theta^{\kappa}|\Delta t|^{\kappa-1} \quad (4.26)$$

Here it is obvious that if $\beta = 0$ the VTT is simply θ. More generally, if $\beta_c = \beta_t$ VTT is independent of Δt, as the time and cost damping cancel out, i.e. we get that $\kappa = 1$. However, in general the β values will not be equal and VTT is not equal to θ. It is for this reason that we change the notation from ω in the non-reference-dependent models, which is always the VTT, to θ in these models, noting that the estimate of θ then needs to be used in (4.26) to calculate the VTT.

Note that in the formulation using value functions it is not appropriate to obtain VTT from strictly marginal valuations, as would be found by differentiation. The concept is to determine the value of a finite amount of time Δt, where the marginal value of both time and cost varies continuously. The use of differentials would, for instance, give the value of changing Δt from 10 to 11 minutes, whereas what is required is the value of the change from 0 to 10 minutes. Moreover, the differential of the time value depends on Δt, whereas the differential of the cost value depends on Δc, so that the ratio of differentials varies in two dimensions.

4.6.4 Implementation for SP1

In **Section 4.4.2** the model specification for SP1 was written as:

$$V_s = \mu_{SP1}.\log\left(-\frac{cost_1 - cost_2}{time_1 - time_2}\right) \quad (4.27)$$

$$V_e = \mu_{SP1}.\log\omega \quad (4.28)$$

Then using the Δ notation and subtracting $\mu_{SP1}\log\omega$, (4.27) and (4.28) can be reformulated without changing their meaning as:

$$V_1 = \mu_{SP1}.\log\left(-\frac{\Delta c_1 - \Delta c_2}{\omega\Delta t_1 - \omega\Delta t_2}\right) \quad (4.29)$$

$$V_2 = 0 \quad (4.30)$$

This purely technical reformulation allows us to extend the model to include reference dependence.

For SP1 the design ensures that the value differences have opposite signs, like the differences themselves. Comparing alternatives s and e, respectively the 'slow' and 'expensive' alternatives in SP1, respondents value the cost difference they are offered by $\left(v(\Delta c_1) - v(\Delta c_2)\right)$ and the time difference by $\left(v(\theta\Delta t_1) - v(\theta\Delta t_2)\right)$. It is then 'rational' to choose the slow alternative if $|v(\Delta c_1) - v(\Delta c_2)| > |v(\theta\Delta t_1) - v(\theta\Delta t_2)|$. This implies a model form:

$$V_s = \mu_{SP1}.\log\left(-\frac{v(\Delta c_1)-v(\Delta c_2)}{v(\theta\Delta t_1)-v(\theta\Delta t_2)}\right) \quad (4.31)$$

$$V_e = 0 \quad (4.32)$$

This is the model that is estimated, with separate dBF parameters for time and for cost, using the reported time and cost for the respondent's recent trip as the reference points.

4.6.5 Implementation for SP2 and SP3

The reference-free specification for SP2 and SP3 was derived in **Section 4.3.3** as (using SP2 as the example and ignoring for now the risk aversion flexibility introduced by the CARA approach):

$$V_j = -\mu_{SP2} \cdot \log\left(cost_j + \sum_{k_T} \omega_{k_T} x_{jk_T} - \sum_{k_{NT}} \omega_{k_{NT}} x_{jk_{NT}}\right) \quad (4.33)$$

To introduce reference dependence we can replace any of the terms in the utility functions by the corresponding v for changes in the associated attribute, noting that if η, β, γ are constrained to zero, we get $v(\theta\Delta x) = \theta\Delta x$. For example, instead of a utility function of the form in (4.33), in cost, time and delay (d), we could write:

$$V_j = -\mu_{SP2} \log\left(c_j + \theta_t t_j + \theta_d d_j\right) \quad (4.34)$$

and we could substitute value functions for these components:

$$V_j = -\mu_{SP2} \log\left(v(\Delta c_j) + v(\theta_t \Delta t_j) + v(\theta_d \Delta d_j) + \theta_t t_0 + \theta_d d_0 + c_0\right) \quad (4.35)$$

The inclusion of the base values in addition to the value functions inside V_j are required, as, in contrast with the model for SP1, we are not working in relative valuation space.

We can get the required generalisation by estimating or eliminating some or all of the parameters expressing reference dependence. In each case θ relates to the willingness-to-pay for changes in the specific attribute.

An important discussion relates to the choice of reference values for the individual non-cost attributes[48] in SP2 and SP3, where the situation is not as straightforward as for SP1. The following approach was used, listing the various components in turn:

[48] Cost was obviously treated the same way as in SP1.

- For car SP2, we allowed for reference dependence only for travel time, not for the standard deviation of travel time, in the absence of a value for the reference trip. The same reference point as in SP1 was used for travel time.
- For the three public transport SP2 games, we used the time for the reference trip as the reference value for the *usual* travel time, with no reference points being available for early or late arrival. Here, it is important to remember that the CARA specification was used for the public transport versions of SP2, meaning that for each alternative, we have five different possible outcomes. The use of the CARA specification does not make the implementation substantially more difficult. Indeed, we then simply work with reference dependence at the level of each of the five outcomes for a given alternative.

  Let V_{js} be the term associated with outcome *s* for alternative j. We then have that:

$$V_{js} = v(\Delta c_j) + v(\theta_t \Delta t_j) + \theta_t t_0 + c_0 + \theta_t t_0 + \theta_e t_{e,j,s} + \theta_l t_{l,j,s} + \delta_l late_{l,j,s} \tag{4.36}$$

  where:

  - Δc_j is defined as before,
  - Δt_j is calculated as the difference between the *usual* time for alternative *j* and the reference time for the person
  - $t_{e,j,s}$ is the amount of early arrival for outcome *s* for alternative *j*, compared to the *usual* travel time
  - $t_{l,j,s}$ is a corresponding value for late arrival
  - $late_{l,j,s}$ is a 0/1 dummy term indicating late arrival

  The CARA formulation then implies that we use:

$$V_j = -\mu_{SP2} \log\left(\sum_s \frac{1-e^{-\alpha V_{js}}}{\alpha} \cdot \frac{1}{5}\right) \tag{4.37}$$

  where α is the previously defined risk aversion parameter.

- For car SP3, we initially attempted the use of separate reference values for the three individual time components, but better fit and more reasonable results were obtained by applying the reference dependence to the total travel time, in relation to the base travel time, thus summing up the three components, while allowing for different θ values for the three valuations within the total travel time. In mathematical terms, this means that, for a given alternative *j* in task *t*, the value function for total time would use:

$$\theta_{TT,j} = \frac{\theta_{FFT} FFT_{jt} + \theta_{LCT} LCT_{jt} + \theta_{HCT} HCT_{jt}}{FFT_{jt} + LCT_{jt} + HCT_{jt}} \tag{4.38}$$

  i.e. the weighted average of the three θ values as a function of the shares of the total time for alternative *j* in task *t* in each of the three conditions. The approach in equation (4.34) can then be used, with $\theta_{TT,j}$ being specific to alternative *j*.

- For SP3 crowding games, we used the travel time for the reference trip as the reference value, with the specific θ being used depending on the crowding level presented for the alternative at hand. This means that the reference dependence effects are the same across all crowding levels. To be specific, we would have, for alternative *j*:

$$\theta_{TT,j} = \sum_{k=1}^{K} \theta_{T,crowding_k} \delta_{crowding_k,j} \quad (4.39)$$

  where $\delta_{crowding_k,j} = 1$ if and only if crowding level *k* applies for alternative *j*. The approach in equation (4.34) can then be used, with $\theta_{TT,j}$ being specific to alternative *j*.

- For the bus time components SP3 game, no reference dependence was used for headway, while, for the three travel time components (free-flow, slowed down and dwell time), after initial tests using attribute specific reference points, an approach corresponding to that for car SP3 above was used.

4.6.6 An illustration of size and sign effects with the dBF specification

To illustrate the performance and interpretation of the dBF specification, we now proceed with step-by step tests on the SP1 data for car commuters. In this process, we compare five models incrementally increasing in complexity:

- a base specification excluding covariates and with η, β and γ all constrained to zero.
- a specification adding in elasticities in relation to household income, and current trip cost, time and distance (given as λ_{income}, $\lambda_{distance}$, λ_{cost}, and λ_{time}, respectively, with additional multipliers for three groups of non-reporters for income, namely $\zeta_{income\ not\ stated}$, $\zeta_{income\ unknown}$, and $\zeta_{income\ refused}$). The specification used is the same one as that explained in detail later on, in equation (4.41).
- a model allowing for simple gain-loss (GL) asymmetry, i.e. sign effects, by estimating values for η_t and η_c.
- a model additionally allowing for damping in sensitivities depending on the size of changes, i.e. size effects by estimating values for β_t and β_c.
- a model additionally allowing for asymmetric damping, i.e. estimating values for γ_t and γ_c.

Table 4.9 shows the results obtained with these five specifications. It is clear from the likelihood ratio test p-values that each additional generalisation of the model leads to significant gains in log-likelihood. While the biggest gain is obtained by allowing for the initial covariate effects, a different conclusion could be reached if the order of model refinements was changed. The crucial finding is that all of these effects lead to improvements.

Before turning to a detailed analysis of the VTT findings, we now look at the findings in terms of respondent behaviour and how the understanding of this changes from one model to the next.

Table 4.9: Testing the dBF specification on SP1

	Base model		**Adding basic covariates**		**Adding sign effects**		**Adding size effects**		**Adding asymmetry in sign effects**	
Final LL	-2,905.31		-2,753.87		-2,744.34		-2,734.65		-2,728.56	
No of parameters	2		9		11		13		15	
Adj. ρ^2	0.0902		0.1354		0.1377		0.1401		0.1414	
LR p-value against previous specification	-		0.0000		0.0001		0.0001		0.0023	
	est.	**t-rat. (0)**	**est.**	**t-rat. (0)**	**est.**	**t-rat. (0)**	**est.**	**t-rat. (0)**	**est.**	**t-rat. (0)**
Θ	3.9687	15.89	4.8084	8.15	4.4907	7.54	5.6861	9.35	5.7202	9.62
λ_{income}			0.3777	4.19	0.4049	4.19	0.4761	3.92	0.4638	3.92
$\lambda_{\text{distance}}$			-0.0149	-0.23	-0.0121	-0.18	-0.0201	-0.26	-0.0151	-0.20
λ_{cost}			0.4432	2.52	0.4377	2.35	0.4768	2.18	0.4742	2.22
λ_{time}			0.1587	0.75	0.2157	0.95	-0.1501	-0.52	-0.1992	-0.71
M	0.8004	20.70	0.8968	20.99	0.8454	18.96	0.7911	14.79	0.7957	14.60
β_t							-0.4514	-3.03	-0.4396	-3.00
β_c							0.0876	1.14	0.1017	1.35
γ_t									-0.2296	-2.63
γ_c									-0.0714	-1.35
η_t					0.2363	4.29	0.1923	3.10	0.2335	3.44
η_c					0.0635	1.20	0.1257	2.03	0.1571	2.40
	est.	**t-rat. (1)**	**est.**	**t-rat. (1)**	**est.**	**t-rat. (1)**	**est.**	**t-rat. (1)**	**est.**	**t-rat. (1)**
$\zeta_{\text{income not stated}}$			1.7057	1.27	1.7097	1.19	1.9672	1.15	2.1896	1.46
$\zeta_{\text{income unknown}}$			0.5526	-1.66	0.5383	-1.66	0.4881	-1.72	0.5304	-1.52
$\zeta_{\text{income refused}}$			0.9848	-0.12	0.9820	-0.07	0.9788	-0.07	0.9756	-0.08

Starting with the model with basic covariates, we observe a positive income elasticity, along with a positive and significant elasticity of the VTT in relation to the cost of the reference trip. No significant impacts are found in relation to trip distance and reference trip time. The multipliers on the VTT for respondents who do not provide income information are not significantly different from one.

The next model shows positive estimates for ηt and ηc, although only the former is statistically significant at usual confidence levels. These findings indicate the presence of sign effects, where the positive estimates show that losses in time/cost are valued more negatively than gains are valued positively.

To illustrate the role of these additional parameters, we now plot graphs showing the value functions for $\theta\Delta t$ and Δc obtained with the different models. We look at changes in time by up to 60 minutes either side of the reference value, and changes in cost by up to £12 either side of the reference value. This is for illustration purposes and goes beyond the ranges faced by almost all commuters in the sample.

When comparing the model with and without sign effects (**Figure 4.2** for the $\theta\Delta t$ function and **Figure 4.3** for the Δc function), we observe the stronger asymmetry for time, with losses (i.e. positive Δ) being valued more negatively than gains are valued positively.

We next turn to the model adding in size effects, i.e. estimating β_t and β_c. A negative estimate means that larger changes are valued increasingly more highly than smaller changes, with the opposite applying for positive estimates. We observe a significant negative estimate for β_t, showing that larger changes in time from the reference value have a relatively larger impact than smaller changes (**Figure 4.4**). For cost, the estimate of β_c is positive, and while not significant, this points towards cost damping, with larger changes being valued less strongly per unit than smaller changes (**Figure 4.5**). The figures now compare the value functions to the model with sign effects only.

We now turn to the last model which allows for asymmetry in the damping to either side of the reference values. We observe a negative and significant estimate for γ_t, indicating stronger size effects for losses than for gains (**Figure 4.6**). The impact of γ_c is less strong and also not significant at usual levels of confidence (**Figure 4.7**).

As a final step, we consider the implications of these gradual model refinements on the VTT. As discussed in the theory sections, the value of θ itself is not a direct estimate of the VTT and needs to be transformed by incorporating β_t and β_c and the size of time changes, i.e. Δt. While the calculation of the VTT produces a value free of sign effects (and thus not requiring the inclusion of η_t, η_c, γ_t and γ_c), the incorporation of these effects in the models clearly has the potential to improve the estimates of θ and other parameters, and thus also impact on the VTT.

With reference to **Figure 4.8**, we calculate the VTT for a representative respondent with an annual household income of £40K, a £5 reference trip cost, 30 minute reference trip time and 20 miles reference trip distance. Higher VTT measures would be obtained with a higher income or a more expensive reference trip, with time and distance elasticities not being significant. To illustrate the role of Δt, we present VTT measures for Δt of up to 20 minutes, which covers 93% of car commuter choice tasks in our sample. As can be seen, adding basic covariates increases the VTT by 21% for the reference individual, who has a relatively high income and reference trip cost. Incorporating sign effects leads to a drop in the VTT by just under 7%. The VTT for the model with size effects corresponds to that of the model with sign effects only when using a value of Δ*t* of 7 minutes. It is lower than that with smaller Δ*t*, and higher with larger Δ*t*. The incorporation of asymmetry in the size effects has almost no impact on the VTT findings, but gives more confidence in the model estimation.

Figure 4.2: Value function for *θ*Δ*t* as a function of Δ*t*

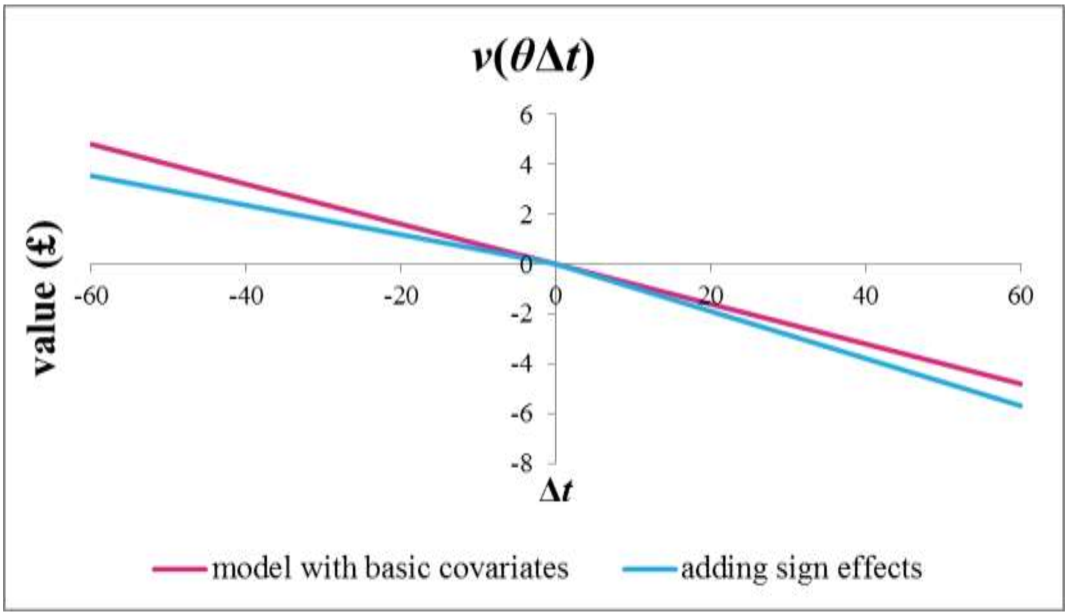


Figure 4.3: Value function for Δ*c* as a function of Δ*c*

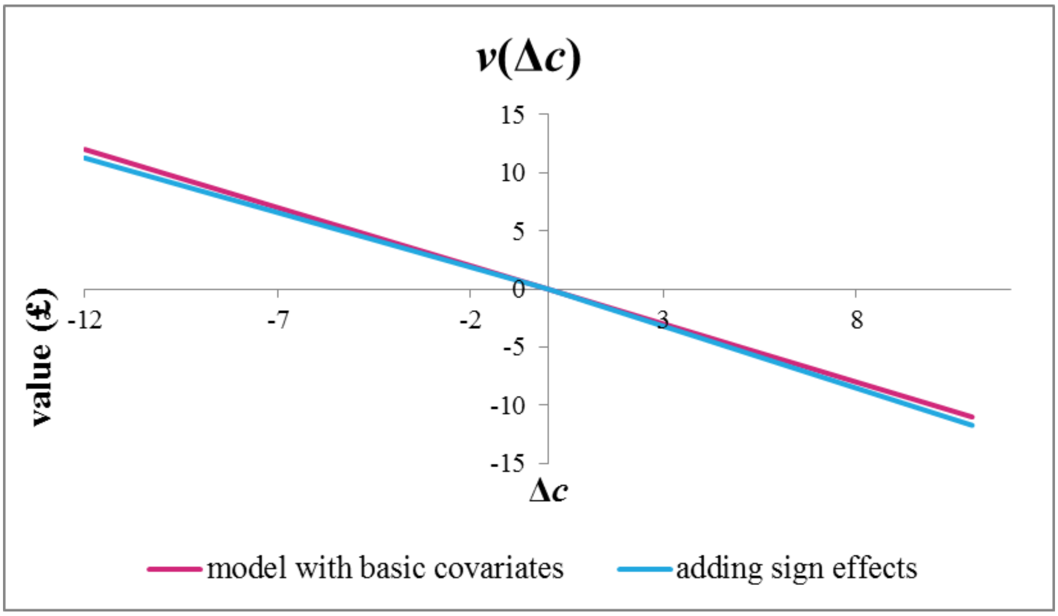


Figure 4.4: Value function for $\theta\Delta t$ as a function of Δt, adding size and sign effects

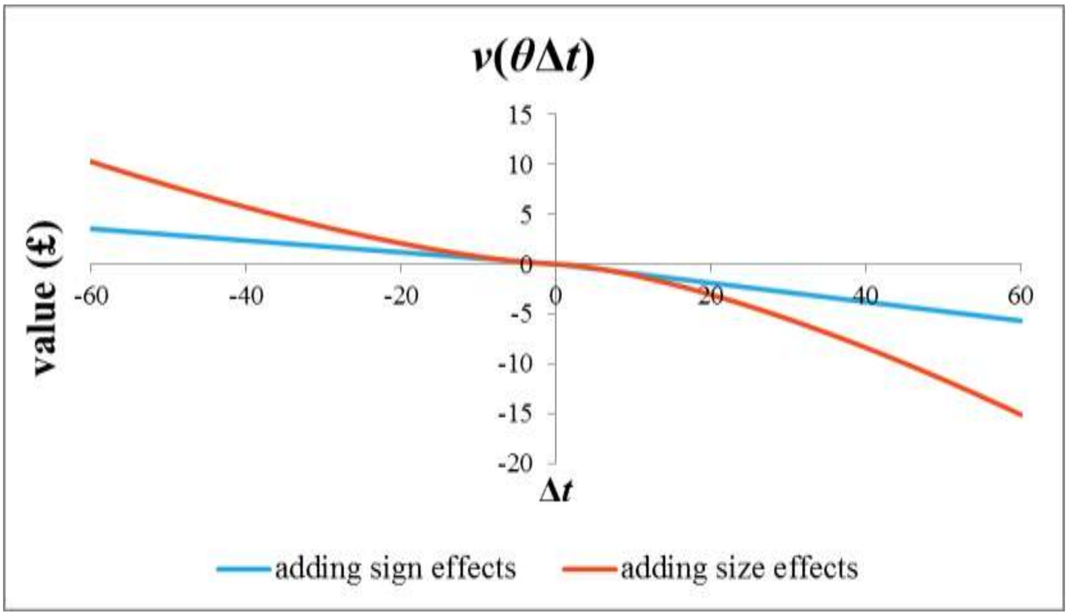


Figure 4.5: Value function for $\theta\Delta c$ as a function of Δc, adding size and sign effects

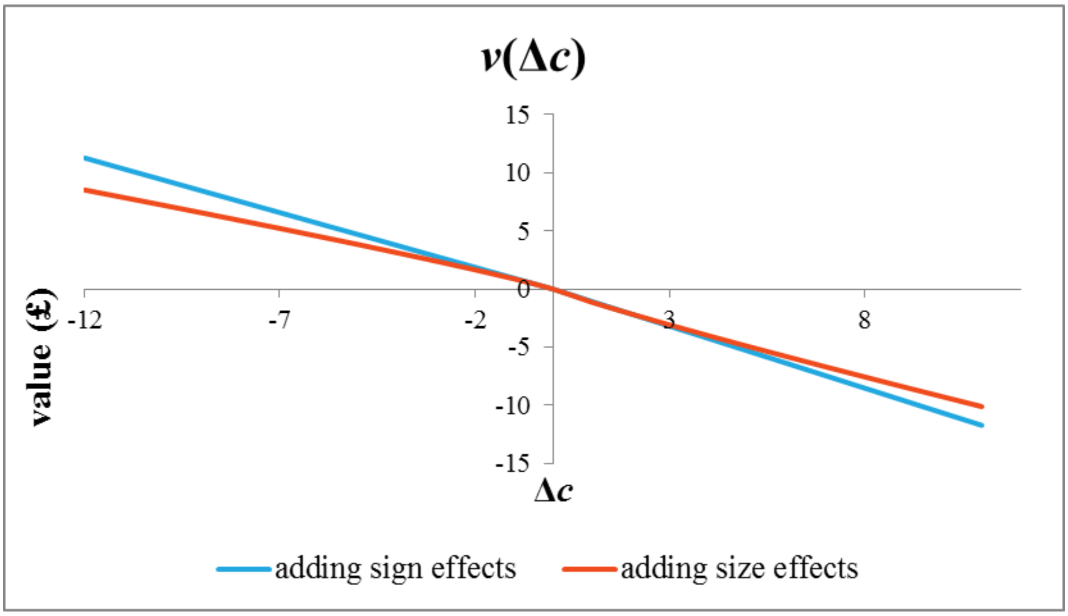


Figure 4.6: Value function for $\theta\Delta t$ as a function of Δt, adding size and asymmetric sign effects

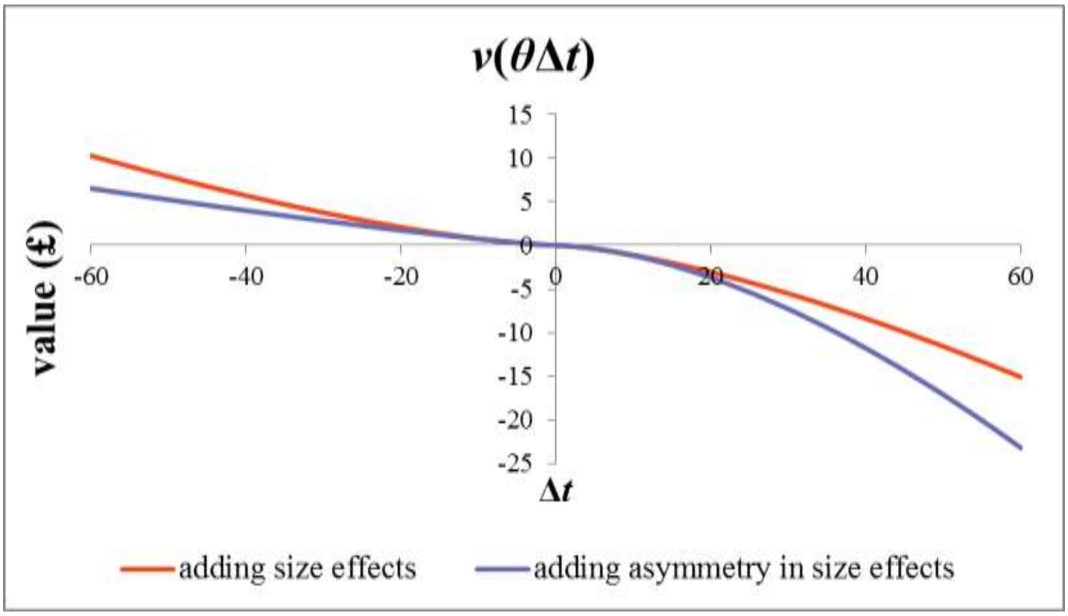


Figure 4.7: Value function for Δc as a function of Δc, adding size and asymmetric sign effects

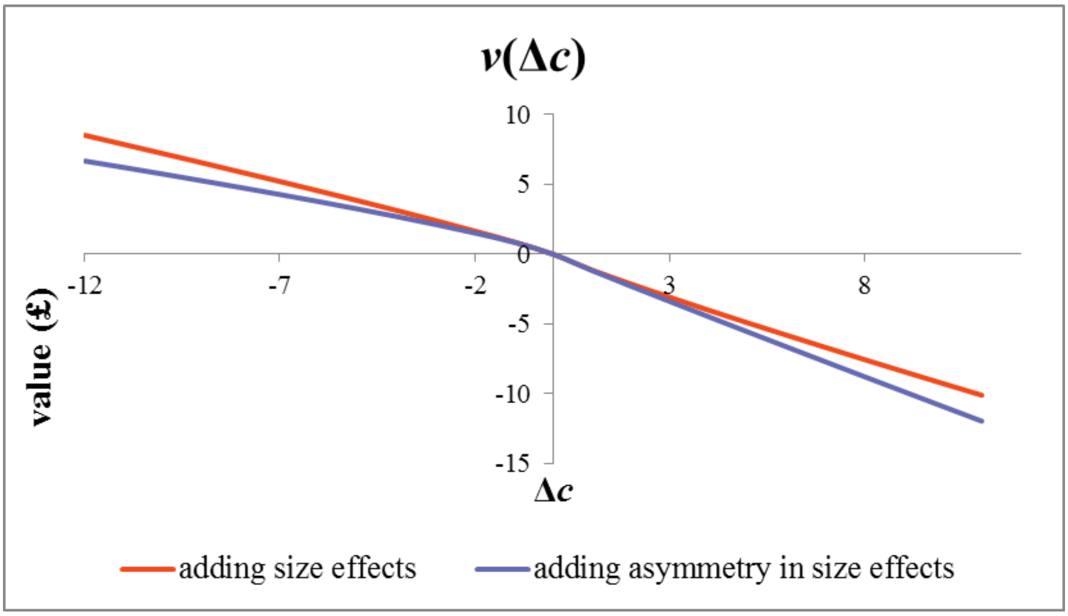
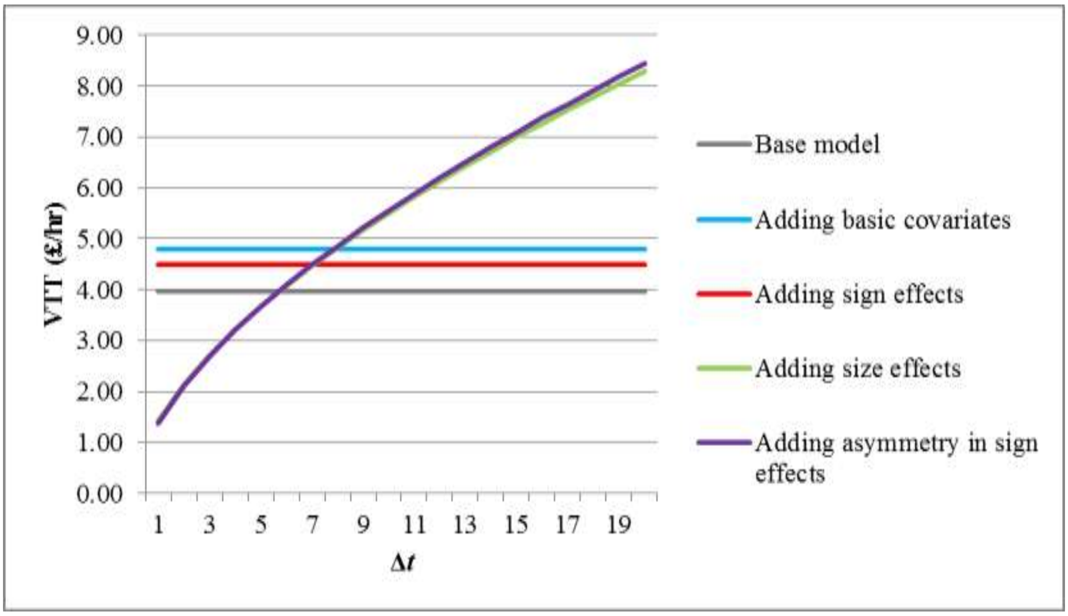


Figure 4.8: VTT (£/hr) for respondent with £40K income, £5 trip cost, 30 minute trip time and 20 miles trip distance



It is important to note that **Figure 4.8** is based on certain assumptions regarding the traveller and trip; we will revisit size and sign effects, and assess the implications for national average VTTs in **Section 7.6.1**.

4.7 Joint modelling approach

While initial tests were conducted separately on the individual SP games (as reported above when comparing the additive and multiplicative specifications), the majority of our work makes use of models jointly estimated on multiple games. This is in line with our decision at the outset to present each respondent with three games of five choice tasks each. The main benefit of the joint estimation is increased robustness for those parameters shared across games, which in our case is the set of covariates explaining deterministic heterogeneity in valuations as well as the random heterogeneity parameters. The details of this process are explained later on in this report.

In the joint estimation, we allow for differences in valuations across games by using separate multipliers for each valuation in our models (across the three games), relating it to the base θ, say θ_0. This allows us to capture differences in valuations that clearly relate to different components (e.g. reliability as opposed to travel time) but also to test for differences in interpretation for attributes that are common across games, such as generic travel time in SP1 and SP2. Using the example of car, where we then obtain six separate θ measures, as follows:

- $\theta_{SP1,VTT} = \zeta_{SP1,VTT}\theta_0$, for the valuation of travel time in SP1
- $\theta_{SP2,VTT} = \zeta_{SP2,VTT}\theta_0$, for the valuation of average travel time in SP2
- $\theta_{SP2,VSDTT} = \zeta_{SP2,VSDTT}\theta_0$, for the valuation of the standard deviation of travel time in SP2
- $\theta_{SP3,VFFT} = \zeta_{SP3,VFFT}\theta_0$, for the valuation of travel time in free-flow conditions in SP3

- $\theta_{SP3,VLCT} = \zeta_{SP3,VLCT}\theta_0$, for the valuation of travel time in light traffic in SP3
- $\theta_{SP3,VHCT} = \zeta_{SP3,VHCT}\theta_0$, for the valuation of travel time in heavy traffic in SP3

A normalisation is required here, and we thus set $\zeta_{SP1,VTT} = 1$, meaning that the base valuations relate most directly to SP1, albeit that we still allow for additional game-specific effects in what follows. A corresponding approach is also used for the other three modes, where the specific valuations differ across modes, but with no impact on the general method.

To alleviate concerns about the impact of individual games on the overall valuations, we performed a simple test where we ran the specification of the final car commute model on the SP1 data only, and this led to a change in the VTT by only 3.4%, suggesting that the use of game specific multipliers is sufficient to avoid any potential biasing impact of individual games, e.g. SP2, on the overall values. This justifies the use of a joint modelling approach, given the benefits in terms of covariates.

We also make use of game-specific scale parameters, as already outlined earlier on in the utility specifications for the separate games.

In most general terms, let us return to the definition of $P_{SP1,n}$, $P_{SP2,n}$ and $P_{SP3,n}$ as being the likelihood of the observed set of choices for respondent *n* in the three sets of stated choice scenarios. We then have that the joint probability of the choices observed for respondent *n* is given by:

$$P_n = P_{SP1,n}P_{SP2,n}P_{SP3,n} \tag{4.40}$$

and the logarithm of this product is the contribution of respondent *n* to the log-likelihood function used in model estimation.

For most respondents, three games, and thus 15 choice tasks, were used in the joint models, with two exceptions. For rail respondents who were given the operator choice game instead of the crowding game, and for ‘other PT’ respondents who were presented with a mode choice game instead of the crowding game, only SP1 and SP2 were used in the joint model. This was motivated by the fact that the valuations in the labelled experiments (operator or mode choice) are substantially different from the other games, and could adversely impact on our ability to estimate consistent covariate effects across games.

4.7.1 Covariates tested for deterministic heterogeneity

An extensive specification search was undertaken to test the impact on valuations by a substantial range of person and trip covariates, as well as to account for potential design effects. These included the following:

Person characteristics:

- Age
- Gender
- Employment status

- Household composition
- Income (household income, personal income, a measure of household income corrected by household composition and an approximation to hourly wage)[49]
- Job category (business trips only)

Trip characteristics:

- Congestion on reference trip (car)
- Mix of different traffic conditions (bus)
- Trip duration (nights away)
- Travelling party composition
- Trip frequency
- Road type mix for reference trip (car)
- Outbound vs return vs one-way
- Driving vs. passenger vs partly driving (car)
- Weekend vs. weekday
- Who pays for the trip
- Time use during travel (slightly different categories for car)
- Peak vs. off-peak
- Cost of reference trip
- Travel time of reference trip
- Distance of reference trip
- Crowding on reference trip (bus, rail, 'other PT')
- Reserved seat (rail)
- Ticket type, single vs. return (bus, rail, 'other PT')
- Season tickets (bus, rail, 'other PT')
- Access mode vs. main mode trip (bus, 'other PT')
- Geography of the trip (London, Urban, Rural)
- Company travel policy (business trips)

Design effects:

- Size and sign effects, as discussed above, where it is not possible to determine whether these are SP artefacts or also apply in real life, or whether they relate to short term or long term effects.
- Position of time attribute relative to cost attribute in presentation of travel options (for SP1 only).
- Presentation of cheap option on left or right (i.e. as first or second alternative).

[49] The categorical income variables were turned into continuous variables, using DfT agreed 'midpoints' of £5K for the lowest category (£7.5K in the case of household income), and then £15K, £25K, £35K, £45K, £62.5K, £87.5K and finally £130K for the upper bracket.

4.7.2 Specification used for deterministic heterogeneity

All the covariates listed above in the person and trip characteristic groups were interacted in the same way for the individual θ measures used in different games. This equates to an assumption that their impact is consistent across the different types of components valued in our work, and that they relate primarily to an underlying willingness-to-pay measure, independently of the good being valued. This is of course a simplification, but one that was necessary in the context of this project. As we will see in **Section 4.7.3**, applying the multipliers to each individual θ is not the same as applying the multipliers to θ_0.

For the majority of the components above, multipliers on the VTT were estimated, with one category for the attribute being used as the base, for which the multiplier was then set to a value of 1. This means that the base estimates of θ relate to an individual and a trip at the base values for these covariates. We will return to this point in the discussion of the actual model results.

A different specification was used for four continuous effects, namely income and the cost, time and distance of the reference alternative. The multipliers used here were based on the elasticity approach from the 2003 UK study. In particular, taking income as the example, the multiplier on θ would be given by:

$$\left(\frac{inc}{40}\right)^{\lambda_{inc}} \delta_{income\ reported} + \zeta_{not\ stated}\delta_{income\ not\ stated} + \zeta_{unknown}\delta_{income\ unknown} + \zeta_{refused}\delta_{refused} \qquad (4.41)$$

With this specification, *inc* is a continuous income variable (expressed in £1,000s per annum), λ_{inc} is an estimated income elasticity, $\zeta_{not\ stated}$, $\zeta_{unknown}$ and $\zeta_{refused}$ are multipliers on the VTT for respondents with unreported income, and the four δ terms are dummy variables categorising the respondents according to whether income was reported or not. The value of 40 chosen as a denominator simply means that the base θ relates to a respondent with an annual income of £40,000. Tests were conducted to determine which of the different income variables was most appropriate for given purposes, where, across modes, we ended up with a specification using household income for commuting and for other non-work, with personal income used for business[50] .

A corresponding specification was used to estimate elasticities of θ with respect to the cost of the reference trip (with a base of £5), the trip time of the reference trip (with a base of 30 minutes) and the distance of the reference trip (with a base of 20 miles). Again, this means that the base estimates for θ relate to a trip with these base values for time, cost and distance. No missing information multipliers were needed for these three covariates.

A slightly different approach was used for design effects. We have already outlined in detail the approach to reference dependence that was used in our work, and how these effects were entered at the level of individual games, with different effects for different attributes. In addition, we tested for an impact of the relative position of the time and cost attributes in SP1, and the presentation of the cheap option on the left or the right (across games). Given that these are purely SP effects which we do not want to influence the estimated VTT, a multiplicative

[50] Note that the normalisation does not need to be adjusted when the income measure is changed, as it merely gives a rough scaling so that the multiplier is close to 1 for most people.

effects coding approach was used. Further, these effects were entered at the level of individual choices in individual games, unlike the other covariates. In particular, taking the example of whether the cheap option was presented on the left or the right, the additional multiplier on the value of the game and task specific θ in choice task t would be given by:

$$\zeta_{cheap\ left}\delta_{cheap\ left_t} + \frac{1}{\zeta_{cheap\ left}}\left(1-\delta_{cheap\ left_t}\right) \qquad (4.42)$$

where $\delta_{cheap\ left_t}$ is set to 1 if and only if the cheap option is presented on the left in task t. This specification ensures that the base estimates of θ relates to the average situation (geometric mean) in the data according to how often the cheap option is presented on the left or on the right.

Some of the covariates that are needed are thus continuous, while others are discrete. The discrete covariates require slightly different treatment within a joint model using a dBF specification, as is explained in the following section.

Other than the interaction of the covariates with θ, we also allowed for the order of the games to be interacted with the scale parameters. With SP1 always being presented first, this allowed us to test whether the scales for SP2 and SP3 were affected by which of these games was presented second, and which was presented third.

Finally, we also tested the inclusion of constants for the alternative presented on the left directly in the utility functions in SP2 and SP3, along with a constant for any alternatives with no travel time variability in the SP2 games. These terms, by being entered directly into the utility functions, do not affect the VTT measures.

4.7.3 Approach to covariates in joint model

The main aim of modelling all of the games together is to ensure a robust estimation of the person and trip type influences on the various valuations, giving the benefit of having 15 tasks per person for most respondents (except those getting for example the operator choice game). This is achieved by ensuring that the impact of the covariates is same for all θ parameters across all of the games. In the absence of a reference dependent specification, this would be straightforward, and we would simply apply them as multipliers on the base θ which is then taken forward into calculation of game specific θ values using the different ζ values discussed at the start of **Section 0**. However, in our work, the role of the β parameters needs to be accounted for (remembering that η and γ do not enter into the WTP calculations).

As discussed above, continuous covariates (e.g. income and trip length) are modelled to act as elasticities affecting the underlying value θ, while discrete covariates (e.g. age category) are modelled as multipliers of θ. Using the VTT in car SP1 as an example, we would then have:

$$\theta_{SP1,VTT} = \theta_0\left(\zeta_{SP1,VTT}\prod_m z_m^{\lambda_m}\prod_n \zeta_n^{z_n}\right)^{1/\kappa_{SP1,VTT}} \qquad (4.43)$$

Where:

θ_0 would relate to a respondent with the base values for all covariates in a single game model

$\zeta_{SP1,VTT}$ is the multiplier for SP1, which as discussed above is normalised to 1 (for SP1 only)

$\kappa_{SP1,VTT} = \frac{1-\beta_{t,SP1}}{1-\beta_{c,SP1}}$ as defined previously, with $\beta_{t,SP1}$ and $\beta_{c,SP1}$ being specific to SP1
λ is the elasticity for a continuous covariate z_m
ζ is the multiplier applied for a discrete covariate z_n when its value is 1
m and n run over the continuous and discrete covariates respectively.

The inclusion of the exponent $1/\kappa_{SP1,VTT}$ ensures that the impacts λ and ζ apply directly to the VTT for SP1, as follows. Indeed, given this parameterisation, we calculate VTT as before by:

$$VTT_{SP1} = \theta_{SP1,VTT}{}^{\kappa_{SP1,VTT}} |\Delta t|^{\kappa_{SP1,VTT}-1}$$

$$= \theta_0^{\kappa_{SP1,VTT}} \zeta_{SP1,VTT} \prod_m z_m^{\lambda_m} \prod_n \zeta_n^{z_n} \,.\, |\Delta t|^{\kappa_{SP1,VTT}-1}, \qquad (4.44)$$

This calculation does not incorporate any of the multipliers relating to design effects, e.g. whether time is shown above cost or whether the cheap option is shown on the left. The multiplicative effects coding used for these covariates already ensures that the base θ_0 relates to the geometric mean in the data for these covariates. This formulation simplifies the reporting and testing of the models. In particular, it ensures that the estimates of the impacts of the covariates as well as the game specific multipliers relate directly to all the individual valuations, as do the standard errors. This re-parameterisation is required given that the value of κ can differ across games and across valuations within a given game. A corresponding approach to that above is used for all the separate valuations in SP2 and SP3.

One important reservation applies to this discussion. Our modelling framework allows for differences in size effects across different time valuations, a decision justified by our estimation results. However, as a result, the game-specific multipliers, e.g. $\zeta_{SP1,VTT}$, cannot be directly understood to explain the differences in the valuations across games. Indeed, the differences in say the VTT between SP1 and SP2 would be given by:

$$VTT_{SP1}/VTT_{SP2} = \zeta_{SP1,VTT}/\zeta_{SP2,VTT}\ \theta_0^{\kappa_{SP1,VTT}-\kappa_{SP2,VTT}} |\Delta t|^{\kappa_{SP1,VTT}-\kappa_{SP2,VTT}} \qquad (4.45)$$

which is thus not as simple as $\zeta_{SP1,VTT}/\zeta_{SP2,VTT}$. While $\theta_0^{\kappa_{SP1,VTT}-\kappa_{SP2,VTT}}$ is simply a constant which can be calculated, $|\Delta t|^{\kappa_{SP1,VTT}-\kappa_{SP2,VTT}}$ is a function of Δt, and as a result, the ratios of different valuations depend on the assumptions made in relation to Δt. This point is addressed in **Chapter 7** by making specific assumptions about Δt.

4.7.4 Incorporating random heterogeneity

The incorporation of random heterogeneity is rather straightforward, and is accommodated by allowing for random heterogeneity in θ_0, i.e. the base value before the incorporation of covariates and reference dependence. This means that the contribution to the likelihood function by person *n* is now given by:

$$P_n = \int_{\theta_0} P_{SP1,n}(\theta_0) P_{SP2,n}(\theta_0) P_{SP3,n}(\theta_0)\, f(\theta_0) d\theta_0 \qquad (4.46)$$

where $f(\theta_0)$ is the density function for θ_0.

For the present study, after testing popular alternatives such as the log-normal distribution, we settled on the use of a log-uniform distribution, which has a somewhat shorter tail than the log-normal distribution and for which any differences in fit were very small, with the log-uniform avoiding problems with extreme values.

In the same way that a variate *x* has a log-normal distribution if *y* = *log*(*x*) is normally distributed, we define *x* as log-uniformly distributed if *y* = *log*(*x*) is uniformly distributed.

Denote a as the lower bound and b as the spread of a uniform distribution, such that, with $U(a, b)$ denoting a uniform distribution between a and a+b, and with $\vartheta = \log \theta_0 \sim U(a, b)$, we have that the mean of the resulting log-uniform distribution is given by:

$$E(\theta_0) = \int_a^{a+b} \exp(\vartheta)\, f(\vartheta|a, b) d\vartheta = \frac{1}{b}\int_a^{a+b} \exp(\vartheta)\, d\vartheta = \frac{\exp(a+b)-\exp(a)}{b} \qquad (4.47)$$

where $f(\vartheta|a, b)$ is a uniform density function between a and a+b, which are estimated parameters.

We also have that:

$$E(\theta_0^2) = \int_a^{a+b} \exp(2\vartheta)\, f(\vartheta|a, b)\, d\theta_0 = \frac{1}{b}\int_a^{a+b} \exp(2\vartheta) d\vartheta = \frac{\exp\big(2(a+b)\big) - \exp(2a)}{2b} \qquad (4.48)$$

$$E(\theta_0^3) = \int_a^{a+b} \exp(3\vartheta)\, f(\vartheta|a, b) d\vartheta = \frac{1}{b}\int_a^{a+b} \exp(3\vartheta) d\vartheta = \frac{\exp(3a+b) - \exp(3a)}{3b} \qquad (4.49)$$

From this, we can then calculate the variance and skewness as:

$$Var(\theta_0) = E(\theta_0^2) - \big(E(\theta_0)\big)^2 = \exp(2a)\left[\frac{\exp(2b)-1}{2b} - \frac{(\exp(b)-1)^2}{b^2}\right] \qquad (4.50)$$

$$Skew(\theta_0) = \frac{E(\theta_0^3) - 3E(\theta_0)Var(\theta_0) - E(\theta_0)^3}{\big(Var(\theta_0)\big)^{\frac{3}{2}}} \qquad (4.51)$$

Finally, the probability density function (pdf) and cumulative distribution function (cdf) are given by:

pdf: $g(\theta_0) = \frac{1}{\theta_0 b}$ for $\exp(a) \leq \theta_0 \leq \exp(a+b)$; 0 otherwise (4.52)

cdf: $G(\theta_0) = \frac{\log(\theta_0) - a}{b}$ for $\exp(a) \leq \theta_0 \leq \exp(a+b)$ (4.53)

With the above specification, the heterogeneity is entered at the respondent level rather than the game level, meaning that the integration over the distribution of θ_0 is carried out over all choices for the respondent in the joint game. We estimated the model using simulated log-likelihood, with 500 Halton draws used per respondent to ensure very small simulation error.

To investigate the appropriateness of using the log-uniform distribution, we report here a simple test showing comparisons between the log-uniform and log-normal distributions, as well as testing the suitability of the log-uniform distribution by using the Fosgerau and Mabit approach[51]. These tests were run on a model for car commuters and SP1 only, but using the full specification reported later in **Section 4.8** for covariates.

Table 4.10: Testing log-uniform against log-normal distributions

				VTT (£/hr) sample enumeration results on estimation sample, using numerical simulation				
Model	**Final LL**	**No. of param.**	**Adj. ρ^2**	**min**	**mean**	**median**	**std. dev**	**max**
Log-normal	-2,308.03	22	0.2708	2.83	18.82	13.20	17.15	132.96
Log-uniform	-2,318.25	22	0.2676	2.24	13.53	9.40	12.17	92.15
Log-uniform with 1 additional polynomial term (K=2)	-2,317.21	23	0.2676	2.39	14.83	10.30	13.31	96.91
Log-uniform with 2 additional polynomial terms (K=3)	-2,302.85	24	0.2718	3.26	21.43	15.34	18.73	160.53
Log-uniform with 3 additional polynomial terms (K=4)	-2,302.24	25	0.2717	2.77	18.57	13.26	16.11	138.37

These tests are reported in **Table 4.10**, where we show the results for a model using a log-normal distribution, a model using a log-uniform distribution and three models with additional polynomial terms. Draws for the standard log-uniform distribution are obtained as $\exp(a + b\xi_U)$, where ξ_U is a uniform random variable between a and a+b. The addition of polynomial terms using the Fosgerau and Mabit approach changes this to $\exp\left(a + \sum_{k=1}^{K} b_k {\xi_U}^k\right)$. Alongside model fit statistics, we also report characteristics for the VTT distribution. These are obtained through numerical simulation on the estimation sample (using 500 Halton draws per respondent), where values for the VTT are calculated separately for each individual, taking into account their socio-economic and trip

[51] Fosgerau, M. and Mabit, S. (2013) 'Easy and flexible mixture distributions', MPRA Paper 46078, University Library of Munich, Germany.

characteristics. The use of numerical simulation means that the range of the log-normal distribution is of course not infinite, as it would be analytically.

Comparing first the two base models, we note that the log-normal distribution obtains a log-likelihood that is 10.22 units better than that for the log-uniform distribution, with the same number of parameters. However, the differences in implied VTT measures are substantially more important than the relatively small difference in fit, with for example the mean being almost 40% higher. This is a direct result of the long tail of the log-normal distribution, and the reason for the censoring discussions in the Scandinavian work. What should be noted here is that applying censoring to the log-normal distribution would inevitably mean that the log-likelihood would be lower than reported here, and reduce (or reverse) the difference between the two distributions. It is also worth noting that for the remaining 20 estimates (i.e. those not linked to the parameters of the random distributions), only minor differences were noted, with none of the differences being larger than half a standard error for the estimates of either model.

Turning next to the inclusion of additional polynomial terms, we see that significant improvements in fit are only obtained for the model with two additional polynomial terms (with no significant improvements with just one additional term, and no further gains for a third term). However, this model shows similar (or worse) problems to that of the log-normal in terms of an excessively long tail and hence higher mean values.

We are of the opinion that the gains in fit for the log-normal model and for the model using two additional polynomial terms are caused by a greater ability to accommodate the time non-traders in the model, at the expense of the overall shape of the distribution and inflated central moments. It should also be noted that the incidence of such non-traders is of course very much lower in the joint models, which is what the remainder of our estimates relate to.

The Danish and Swedish studies discussed the issue of tails of random distributions in great detail and advocated the use of censoring. In the present study, we proceeded without censoring of results for a number of key reasons.

Firstly, censoring is an inherently unsatisfactory process. It implies the estimation of a model and then changing the outputs from that estimation process with a view to rejecting a few ‘inconvenient’ values. The issue with this process is that if censoring is applied post-estimation, it is then possible that the results for other model parameters do not relate to the censored results; i.e. the estimated model no longer gives an optimal fit to the data. In other words, if say censoring was applied during estimation for the value of time distribution, then it is also likely that different values would be obtained for key covariates such as income elasticities. Censoring during estimation is however substantially more difficult than censoring afterwards. Attempts were made with the addition of semi-nonparametric terms to change the distribution, but no meaningful differences in results were obtained, as discussed above.

Secondly, our models are estimated jointly across all games, while the other studies used random heterogeneity only at the level of an individual game, in particular SP1. This substantially increases the scope for non-trading on time which in turn could lead to long tails of the estimated distribution. In our study, this is not the case thanks to the estimation across all three games, with the rates of time non-trading especially being very small. The joint estimation on all games also prevents us from easily conducting non-parametric analysis of the data as performed in the Scandinavian work, i.e. inferring the distribution of the VTT from 'looking' at the data.

Thirdly, the use of the log-uniform distribution itself reduces the issues with extreme tails, with the upper limit not being infinity. This in effect means that the parameters are estimated conditional on the censoring, avoiding the issues of estimating incorrect parameters when censoring is applied after estimation.

Fourthly, the calculation of an appropriate censoring point would be very arbitrary. Even in the Scandinavian work, using the boundary VTT is a somewhat arbitrary input to censoring as the data revealed that some people did indeed have higher values than that boundary, although this was mitigated in the Swedish work by having presented much wider trade-offs allowing a greater investigation of the true truncation point. This also meant that the distribution was rather insensitive to the actual choice of a truncation point. But in the case of models estimated jointly across three games, the calculation of such boundaries is not easily possible, and the use of an SP1 boundary would not be appropriate as SP2 and SP3 might well allow for the estimation of a wider range of VTT measures.

Finally, the wide ranges used in our design work give extensive coverage to the domain of possible VTT values to be revealed in the models.

For bus and 'other PT', we were additionally able to estimate random scale heterogeneity across respondents, thus allowing for variations in the degree of error in the models, net of the damping effects already introduced by the multiplicative model. A possible reason for our inability to estimate such scale heterogeneity also for car and rail is that, with the longer trips for these modes, the main source of scale heterogeneity is distance based, and this is already captured by the multiplicative models.

In particular, to introduce random scale heterogeneity into the bus and 'other PT' games, we used:

$$\mu_{SPj} = e^{\mu_{log(\mu_{SPj})} + \sigma_{log(\mu)} \xi_N} \quad (4.54)$$

where ξ_N is a standard normal random variate, and where the resulting scale parameters now follow a log-normal distribution. The means of the normal distribution for the logs of the scale parameters are game specific, while a common standard deviation term was used, i.e. $\sigma_{log(\mu)}$, to capture the fact that scale heterogeneity should be common across games, but proportional to the scale.

4.8 Estimation results for joint models

This section focusses only on the estimates in the final models for each purpose. The process used in the specification search was iterative, gradually building up model complexity and using results from intermediary models as starting values for more complex models, leading to substantial reductions in estimation time, which was essential in the context of a tight analysis schedule. We started with models excluding covariates and first added elasticities in relation to travel time, travel cost, distance and income. This was then followed by adding in size and sign effects, before gradually adding in the remaining covariates. These were added in batches, each time removing insignificant parameters before adding additional ones, and every time using the final estimates from the previous model as starting values.

In general, parameters with a low level of statistical significance were removed from the model with the exception of a number of key multipliers where we did not wish to impose equality with base categories, for example. Parameters that did not have a significant estimate across any of the three purposes in estimation are not shown in the tables which follow; this applies for example to a large number of the dBF terms, in addition to many of the covariates tested, which could be traveller, trip or design related.

The presentation of the results is divided into a number of key categories which are now looked at in turn. For each parameter, we present the estimate along with either the t-ratio against 0 or the t-ratio against 1, whichever is more appropriate for the parameter in question (e.g. 1 for any multipliers). The standard errors used in these calculations are robust standard errors obtained using the sandwich estimator, taking into account the repeated choice nature of the data in the calculation of the log-likelihood function.

4.8.1 Results for car

The results for the car models are presented in **Table 4.11**, combining data across SP1, SP2 and SP3.

The findings can be summarised as follows:

parameters of base θ_0 distribution

- The estimates here relate to the uniformly distributed logarithm of the θ_0 parameter, meaning that θ_0 itself follows a log-uniform distribution. The models retrieve significant random heterogeneity for all three purpose segments as can be seen by the statistically significant and large estimates for $b_{\log(\theta_0)}$, which gives the range of the underlying uniform distributions, while $a_{\log(\theta_0)}$ gives the minimum value of the underlying distribution, so that its exponential gives the minimum value of θ_0.

game specific θ_0 multipliers

- As discussed above (cf. equation (4.46)), the game specific multipliers cannot directly be interpreted as the differences in the valuations across games as these differences are a function of Δt and the impact of this differs across the valuations in the different games. As such, analysis of the differences across games will happen at the implementation stage in **Chapter 7**, on the basis of specific assumptions about Δt.

key elasticities

- Significant positive income elasticities on the various VTT measures are obtained across all three purposes, which are highest for other non-work, and lowest for employees' business, remembering that the latter uses personal as opposed to household income. It is instructive to draw comparison to the income elasticities reported in the 2003 study[52], although it should be borne in mind that the latter study was restricted to the car mode, to the non-work trip purposes, and employed a different model specification to the present study. The 2003 study reported income elasticities of 0.36 and 0.16 for commute and other non-work respectively, somewhat lower than the corresponding elasticities (0.58 and 0.68) reported in the present study. Whilst theory and evidence point to a temporal income elasticity of VTT with respect to income of one, cross-sectional elasticities estimated on SP data are usually somewhat less than one. In particular, Wardman and Abrantes (2011)[53] reported an income elasticity from meta-analysis of 0.5, and the elasticities given in **Table 4.11** would by comparison seem credible.

- There are significant positive elasticities on the valuations in relation to reference cost and negative elasticities in relation to reference time, with the cost elasticities generally larger in absolute value than the time elasticities, implying that the VTT is higher for longer distance trips (which tend to be more expensive and take longer). The elasticities are strongest (in terms of magnitude and significance) for other non-work, and weakest for employees' business. Cost elasticities were also reported in 2003, although the elasticities reported in **Table 4.11**, 0.68 and 1.05 for commute and other non-work respectively, are somewhat higher than the corresponding estimates (0.42 and 0.31) from 2003, possibly as the 2003 study was not able/did not allow jointly for the impact of trip cost and trip time on valuations.
- These time and cost elasticities can be related to the damping effects on longer trips, where heterogeneity in valuations implies that sensitivity to both time and cost diminishes on longer trips, so cost increases would *increase* VTT, while time increases would *reduce* VTT. Note that if the elasticities are highly negatively correlated, as would be expected, then errors in their estimation will 'cancel out' in the VTT error calculation.
- A significant distance elasticity is observed only for employees' business, which is positive, leading to higher valuations on longer trips even with time and cost being held constant. The 2003 study specified the cost elasticity as a proxy for distance, as no distance information had been collected in the survey, and hence could not attempt to distinguish between the two effects; the high correlation between cost and distance will largely explain the insignificance of the distance elasticity for non-work in **Table 4.11**, but it has also been argued before that impacts of cost and time make more behavioural sense than a pure distance impact[54].

traveller covariates

- Respondents who refuse to provide an income measure tend to have lower VTT measures, while a less clear picture emerges for those with unstated

[52] Mackie, P.J., Wardman, M.R., Fowkes, A.S., Whelan, G.A., Nellthorp, J., and Bates, J.J. (2003) 'Value of Travel Time Savings in the UK'. Report to Department for Transport. Available at: http://eprints.whiterose.ac.uk/2079/2/Value_of_travel_time_savings_in_the_UK_protected.pdf

[53] Abrantes, P.A.L. and Wardman, M.R. (2011) 'Meta-analysis of UK values of travel time: an update'. Transportation Research Part A, 45, pp1-17.

[54] See Daly, A. (2010) 'Cost Damping in Travel Demand Models'. Report of a RAND Europe study for the Department for Transport: http://webarchive.nationalarchives.gov.uk/20110202223908/http://www.dft.gov.uk/pgr/economics/rdg/costdamping/.

income or unknown income, where the very high value (and associated high standard error) for unknown income for employees' business suggests the presence of a small number of respondents with outlying behaviour, justifying the inclusion of this multiplier. Even when the multiplier is technically insignificant, it has been retained to avoid these people biasing the model; the multiplier is *not* used in calculating the output VTT.

- Female commuters show higher valuations, while younger travellers, all else being equal, also have higher valuations for commute and for other non-work, where, for the latter, this captures both of the lower two age categories.
- There are lower valuations for households with two or more adults in the other non-work segment.
- Households owning at least one car have higher valuations for other non-work trips, while households with two or more motorcycles have much lower valuations in the same segment.
- Self-employed commuters have higher valuations, as do commuters where travel costs are paid by the company.
- For employees' business trips, the valuations are higher if the company buys savings come what may, and lower if it does not buy time savings.
- For employees' business trips, valuations are lower if self-employed, and then lower still if self-employed costs are not covered; these features of business travel are discussed in more detail in our synthesis of the business evidence in **Chapter 6** of this report.

trip covariates

- Valuations are higher on other non-work trips with at least one night away from home.
- Commuters travelling with others have lower valuations.
- Commuters have lower valuations when driving on rural roads.
- The impact of congestion during the reference trip is only evident for commuters and other non-work, and the statistical significance of the effects is low, albeit that we see higher valuations for respondents with more congestion on their reference trips.
- For employees' business travel with a London origin and destination, we observe higher valuations, although the statistical significance of the effect is low.

design covariates

- We observe lower valuations when the cheap option is presented on the left for both SP1 and SP3 for commute and other non-work, where the effects are less strong and not as significant for SP3 as for SP1, or for other non-work as opposed to commute.
- Valuations are lower in SP1 for employees' business if time is shown above cost.
- The scale (i.e. the inverse of the variance of the random error term) for SP2 is lower for commuters if SP2 is shown before SP3, and higher for employees' business.

scale parameters

- The values for the scale parameters should not be compared between SP1 and SP2/SP3 given the different modelling approach that was used for SP1.
- No specific insights could be obtained from comparing the relative values of μ_{SP2} and μ_{SP3} across purposes.

dBF parameters

- We observe size effects for time in SP1 and SP2 across all three purposes, and for cost in SP1 for employees' business and other non-work. This means that, in the contexts noted, valuations vary by the size of the time/cost change away from the reference values; we will account for this phenomenon when calculating final values in **Chapter 7**.
- There is asymmetric damping for time in SP1 across all three purposes, and in SP2 for employees' business and other non-work, while, for cost, asymmetric damping is observed only for employees' business and then only in SP3.
- There are sign effects (gain-loss asymmetry) for time in SP1 and SP2 for commuters and SP1 for other non-work, while we note sign effects for cost for commuters in SP1 and SP3, and in SP2 for employees' business and other non-work. This means that, in the contexts noted, valuations vary depending on whether there is a time/money gain or loss; with the specification used in our work, these differences cancel out and we obtain a reference free value, as well as such a value can be determined.

Table 4.11: Estimation results for joint car models

	Commute		Employees' business		Other non-work	
Respondents	922		917		977	
Observations	13,830		13,755		14,655	
Final LL	-7,332.67		-6,933.43		-7,585.74	
Adj. ρ^2	0.23		0.27		0.25	
parameters of base θ_0 distribution	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$a_{\log(\theta_0)}$	-0.3559	-1.74	0.5150	3.61	-0.8840	-2.65
$b_{\log(\theta_0)}$	3.7060	15.62	3.3727	18.31	3.7141	19.16
game specific θ_0 multipliers	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
SP1 travel time	1	-	1	-	1	-
SP2 travel time	1.5988	4.05	1.1396	0.82	2.1875	5.52
SP2 std dev of travel time	0.5803	-4.75	0.8765	-1.04	0.8118	-1.48
SP3 free-flow	0.6968	-2.26	0.5718	-4.54	0.5008	-4.43
SP3 light traffic	0.9770	-0.14	0.9206	-0.74	0.8801	-0.90
SP3 heavy traffic	1.8557	2.98	1.7076	4.23	1.9955	4.05
key elasticities	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
income elasticity (λ_{income})	0.5797	6.10	0.3003	3.64	0.6819	7.76
distance elasticity ($\lambda_{\text{distance}}$)	0	-	0.2390	3.41	0	-
cost elasticity (λ_{cost})	0.6790	3.70	0.4511	2.63	1.0492	6.56
time elasticity (λ_{time})	-0.6241	-2.62	-0.4538	-2.29	-0.9273	-4.72
traveller covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
unstated income ($\zeta_{income\ not\ stated}$)	2.4775	0.65	0.5034	-2.41	1.0117	0.03

	Commute		Employees' business		Other non-work	
unknown income ($\zeta_{income\ unknown}$)	1.4264	1.16	9.3098	2.90	0.2998	-5.89
refused income ($\zeta_{income\ refused}$)	0.7697	-1.30	0.5812	-1.43	0.8644	-0.77
female (base=male)	1.3674	2.26				
aged 17-29 (base=30+)	1.3645	1.76				
aged 17-39 (base=40+)					1.4530	2.52
household with 2+ adults (base=1 or no adults)					0.6980	-3.47
1+ car owned (base=no cars)					2.6826	1.91
2+ motorcycles owned (base=1 or 0 motorcycles)					0.4668	-1.65
Self-employed (base=any other)	1.6669	1.97				
Travel costs paid by company (base=respondent or other paid)	2.2194	3.09				
Company would buy savings come what may (base=buys if benefits>costs, or unknown)			1.3044	1.34		
Company would not buy time savings (base=buys if benefits>costs, or unknown)			0.4435	-9.51		
Self-employed costs not covered (base=costs covered)			0.5629	-3.04		
Self-employed (base=paid employment)			0.6767	-2.56		
trip covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
1+ nights away (base=day return)					1.5522	2.14
travelling with others (base=travelling alone)	0.6690	-3.37				
driving on rural roads (base=urban or motorway)	0.8119	-1.31				
light traffic (base=free flow)	1.4025	1.57			1.3554	1.51
heavy traffic (base=free-flow)	1.5604	1.78			1.4621	1.57
trip with London base origin & destination (base=any other)			1.7530	1.42		
design covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
SP1 cheap option on left (multiplicative effects coding)	0.8842	-3.26			0.9259	-2.00
SP3 cheap option on left (multiplicative effects coding)	0.9282	-1.37			0.9537	-0.85
SP1 time shown above cost (multiplicative effects coding)			0.8878	-2.52		

	Commute		Employees' business		Other non-work	
SP2 scale (μ_{SP2}) multiplier if SP2 before SP3 (multipl. effects coding)	0.8938	-2.63	1.1531	2.52		
scale parameters	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
μ_{SP1}	1.1975	14.71	1.7354	16.90	1.3014	16.53
μ_{SP2}	7.7383	18.05	6.3695	10.95	7.5389	16.92
μ_{SP3}	5.6636	14.65	7.2603	16.16	5.9410	17.07
dBF parameters	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$\beta_{t,SP1}$	-0.4000	-3.64	-0.1141	-1.61	-0.1366	-1.91
$\beta_{t,SP2}$	-0.1564	-2.84	-0.4487	-5.10	-0.2435	-4.96
$\beta_{c,SP1}$			0.1013	1.83	0.1032	1.78
$\gamma_{t,SP1}$	-0.2127	-3.52	-0.1293	-3.58	-0.1075	-2.75
$\gamma_{t,SP2}$			-0.0627	-1.94	-0.0606	-1.79
$\gamma_{c,SP3}$			-0.1581	-2.11		
$\eta_{t,SP1}$	0.2573	4.34			0.2237	4.20
$\eta_{t,SP2}$	0.0874	1.43				
$\eta_{c,SP1}$	0.1267	2.18				
$\eta_{c,SP2}$			0.1959	2.15	0.2244	2.88
$\eta_{c,SP3}$	0.2771	1.51				

4.8.2 Results for rail

The results for the rail models are presented in **Table 4.12**, combining data across SP1, SP2 and the SP3 crowding games, meaning that respondents with the operator choice games only have 10 observations in these joint models as they did not obtain the crowding game, and as operator choice was treated in a separate model. The findings can be summarised as follows:

parameters of base θ_0 distribution

- The models retrieve significant random heterogeneity for all three purpose segments as can be seen by the statistically significant and large estimates for $b_{\log(\theta_0)}$, which gives the range of the underlying uniform distributions.

game specific θ_0 multipliers

- The multipliers (relative to the time measured in SP1) are all of the correct sign, with the negative multipliers for early arrival relating not to schedule delay but to a reduction in trip time, which is desirable.
- As discussed above (cf. equation (4.46)), the game specific multipliers cannot directly be interpreted as the differences in the valuations across games as these differences are a function of Δt, where the impact of this differs across the valuations in the different games. As was noted above when discussing car, we will defer interpretation of the multipliers until the implementation stage in **Chapter 7**.
- We can however draw some conclusions in relation to the crowding multipliers in SP3 as the same dBF parameters apply to all these multipliers. We note a monotonic increase in sensitivity to different levels of crowding for both seated and standing passengers. Across purposes however, the sensitivity to highest crowding level for seated passengers is higher than the sensitivity to the lowest crowding level for standing passengers, and in fact for the two lowest levels of crowding for standing passengers for commute and other non-work. For other non-work, the sensitivity to the two lowest levels of crowding for seated passengers is constant.
- Our presentation of crowding within the SP entailed a simplified version of that which underpins current PDFH values. Overall, the crowding multipliers estimated in the course of the present study are largely supportive of PDFH guidance, but we will return to this comparison in **Chapter 7** when deriving national average VTTs.

key elasticities

- Significant positive income elasticities on the various VTT measures are obtained across all three purposes, which are highest for employees' business, remembering that this uses personal as opposed to household income.
- A distance elasticity is observed only for employees' business, which is positive but not highly significant, leading to higher valuations on longer trips even with time and cost being held constant.
- There are significant positive elasticities on the valuations in relation to cost, and negative elasticities in relation to time, implying that the VTT is higher for longer distance trips (which tend to be more expensive and take longer).

traveller covariates

- A diverse picture emerges for the various multipliers for respondents without income information, where the only highly significant effect is a much lower set of valuations for commuters with unstated income.

- Female respondents on other non-work trips have lower valuations.
- There are higher valuations for households with three or more children in the other non-work segment, and lower valuations for households with three or more adults in the commute segment.
- For commute trips, valuations are higher if costs are paid by the company or any other party, while, for other non-work trips, they are higher if costs are paid by the company.
- For employees' business trips, the valuations are higher if the company buys savings come what may, and lower if it does not buy time savings or if the policy is unknown to the respondent.
- For employees' business trips, valuations are lower if self-employed, especially for blue collar.

trip covariates

- Valuations are lower for commuters on trips with overnight stays, higher for other non-work for one night return trips, and lower for employees' business on trips with multiple nights away from home.
- Lower frequency leads to lower valuations for other non-work (if less than daily) and commute (if less than monthly).
- Valuations are lower for commute and employees' business for one-way trips.
- Valuations are lower for weekend travel for other non-work.
- Valuations are lower for travellers without a reserved seat for other non-work.
- For employees' business travel with a London origin and destination, we observe higher valuations.

design covariates

- We observe lower valuations when the cheap option is presented on the left for SP1 for other non-work.

SP2 specific effects

- The negative values for α show risk seeking behaviour across all purposes.
- There is an overall preference for alternatives with constant travel times, i.e. no variability.

scale parameters

- The values for the scale parameters should not be compared between SP1 and SP2/SP3 given the different modelling approach that was used for SP1.

dBF parameters

- We observe size effects for time in SP1 and SP3 across all three purposes, and for cost in SP1 for employees' business and in SP2 for all purposes.
- There is asymmetric damping for time in SP1 for commute and other non-work, and for time in SP3 and cost in SP2 for commute and employees' business.
- There are sign effects (gain-loss asymmetry) for time in SP1 for commuters and employees' business, for time in SP2 for commute and other non-work, for cost in SP2 for commute and other non-work, and for cost for commute in SP3.

Table 4.12: Estimation results for joint rail models

	Commute		Employees' business		Other non-work	
Respondents	847		945		996	
Observations	12,340		13,390		14.275	
Final LL	-6,016.62		-6,903.61		-7,371.27	
Adj. ρ^2	0.29		0.25		0.25	
parameters of base θ_0 distribution	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$a_{\log(\theta_0)}$	0.4305	4.15	0.6025	5.21	0.8655	4.55
$b_{\log(\theta_0)}$	2.7356	18.92	2.6219	20.66	2.8442	22.52
game specific θ_0 multipliers	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
SP1 travel time	1	-	1	-	1	-
SP2 travel time	1.2356	2.58	1.8071	4.45	1.0877	0.70
SP2 early delay	-2.1925	-2.83 (vs -1)	-2.7967	-3.28 (vs -1)	-2.5482	-3.49 (vs -1)
SP2 late delay	3.5360	5.07	4.9920	5.36	3.4967	6.10
SP3 seated with 50% Load Factor	0.7033	-3.07	0.8509	-1.20	0.7336	-2.42
SP3 seated with 75% Load Factor	0.7621	-2.45	0.8618	-1.08		
SP3 seated with 100% Load Factor	0.9695	-0.29	1.1280	0.89	1.0242	0.19
SP3 seated with 1 pass standing per m^2	1.0543	0.50	1.2790	1.83	1.1642	1.17
SP3 seated with 3 pass standing per m^2	1.2704	2.24	1.5289	3.13	1.4280	2.69
SP3 standing with 0.5 pass per m^2	1.1216	0.97	1.4506	2.40	1.2417	1.51
SP3 standing with 1 pass per m^2	1.1574	1.19	1.5612	2.79	1.2962	1.69
SP3 standing with 2 pass per m^2	1.2750	1.88	1.7642	3.20	1.6055	3.04
SP3 standing with 3 pass per m^2	1.5246	3.25	1.8148	3.00	1.8298	3.66
SP3 standing with 4 pass per m^2	1.8026	4.49	2.2878	4.05	2.2197	4.74
key elasticities	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**

	Commute		Employees' business		Other non-work	
income elasticity (λ_{income})	0.2979	4.44	0.3566	5.69	0.2936	6.10
distance elasticity ($\lambda_{\text{distance}}$)	0	-	0.0585	1.15	0	-
cost elasticity (λ_{cost})	0.6640	7.82	0.7428	12.67	0.5983	10.35
time elasticity (λ_{time})	-0.2753	-2.49	-0.3479	-3.87	-0.5406	-7.20
traveller covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
unstated income ($\zeta_{income\ not\ stated}$)	0.3871	-5.58	1.0728	0.24	0.6251	-1.50
unknown income ($\zeta_{income\ unknown}$)	1.5947	0.82	1	-	1.1516	0.52
refused income ($\zeta_{income\ refused}$)	0.6543	-1.62	2.4822	0.93	1.1719	0.59
female (base=male)					0.8429	-2.37
household with 3+ children (base=2 or fewer children)					1.6202	2.10
household with 3+ adults (base=2 or fewer adults)	0.8539	-1.64				
Travel costs paid by company or other (base=respondent paid)	2.0689	4.16				
Travel costs paid by company (base=respondent or other paid)					1.6473	2.39
Company would buy savings come what may (base=buys if benefits>costs)			1.4692	2.09		
Company policy on savings unknown (base=buys if benefits>costs)			0.5616	-2.62		
Company would not buy time savings (base=buys if benefits>costs)			0.3504	-19.00		
Self-employed briefcase (base=paid employment)			0.5489	-5.82		
Self-employed blue collar (base=paid employment)			0.3612	-6.64		
trip covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
1+ nights away (base=day return)	0.5676	-5.22				
1 night away (base=day return or 2+ nights away)					1.4008	2.40
2+ nights away (base=day return or 1 night away)			0.7714	-2.29		
frequency less than once per day (base=daily)					0.6477	-2.75
frequency less than once per month (base=1 or more times per month)	0.7227	-2.60				
one way trip (base=return trip)	0.6023	-2.80	0.6996	-2.04		
weekend travel (base=weekday travel)					0.6439	-3.30

	Commute		Employees' business		Other non-work	
no reserved seat (base=reserved seat)					0.8514	-1.90
trip with London base origin & destination (base=any other)			1.9416	2.86		
design covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
SP1 cheap option on left (multiplicative effects coding)					0.9490	-2.13
SP2 specific effects	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
risk averseness parameter (α)	-0.0395	-3.11	-0.0070	-2.48	-0.0279	-3.81
constant for alternatives with zero variability (expressed in £)	0.4465	4.60	0.3814	2.12	0.5682	5.44
scale parameters	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
μ_{SP1}	1.8210	18.90	2.1018	20.06	1.8857	21.80
μ_{SP2}	7.2521	9.91	10.8840	10.40	6.9026	12.09
μ_{SP3}	6.5578	13.48	7.1558	11.31	6.6187	12.93
dBF parameters	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$\beta_{t,SP1}$	-0.2137	-3.31	-0.1462	-3.62	-0.1327	-2.79
$\beta_{t,SP3}$	-0.2418	-2.51	-0.1421	-3.14	-0.1861	-3.72
$\beta_{c,SP1}$			0.0683	1.75		
$\beta_{c,SP2}$	0.2478	3.78	0.1351	1.95	0.1587	2.84
$\gamma_{t,SP1}$	-0.1093	-2.39			-0.1364	-5.19
$\gamma_{t,SP3}$	-0.1216	-2.00	-0.1016	-3.48		
$\gamma_{c,SP2}$	0.3456	3.40	-0.2641	-3.61		
$\eta_{t,SP1}$	0.0951	1.63	0.1109	3.87		
$\eta_{t,SP2}$	0.1408	2.18			0.2858	4.40
$\eta_{c,SP2}$	0.2058	2.01			0.2319	1.54
$\eta_{c,SP3}$	0.1379	1.90				

4.8.3 Results for 'other PT'

The results for the 'other' PT models are presented in **Table 4.13**, combining data across SP1, SP2 and the SP3 crowding games, meaning that respondents with the mode choice games only have 10 observations in these models.

The findings can be summarised as follows:

parameters of base θ_0 distribution

- The models retrieve significant random heterogeneity for all three purpose segments as can be seen by the statistically significant and large estimates for $b_{\log(\theta_0)}$, which gives the range of the underlying uniform distributions.

game specific θ_0 multipliers

- The multipliers are all of the correct sign, with the negative multipliers for early arrival relating not to schedule delay but to a reduction in trip time, which is desirable.
- As discussed above (cf. equation (4.46)), the game specific multipliers cannot directly be interpreted as the differences in the valuations across games as these differences are a function of Δt, where the impact of this differs across the valuations in the different games.
- We can however draw some conclusions in relation to the crowding multipliers in SP3 as the same dBF parameters apply to all these multipliers. We note a monotonic increase in sensitivity to different levels of crowding for commute trips, while, for employees' business and other non-work, the lowest three levels (seated) had to be combined.
- ITF (2014)[55] helpfully summarises official UK guidance on public transport crowding multipliers, giving the ranges 1.0-1.05 for seated travel at load factors of 70-100%, 1.06-2.12 for seated travel at passenger densities of 1-3 pass/m^2, and 1.45-2.80 for standing travel at passenger densities of 1-3 pass/m^2. Whilst the SP presentation employed in the present study does not map exactly to the ITF segmentations, the clear finding is that the multipliers reported in **Table 4.13** are somewhat more conservative than those given in UK guidance.

key elasticities

- Significant positive income elasticities on the various VTT measures are obtained for employees' business and other non-work, with a low and not very significant income elasticity for commute.
- No significant distance elasticity is observed for any segment.
- There are significant positive elasticities on the valuations in relation to cost for commute and other non-work, and negative elasticities in relation to time for all purposes. This implies that the VTT is higher for longer distance trips (which tend to be more expensive and take longer) for commute and other non-work, but decreases for employees' business in the absence of a cost elasticity.

traveller covariates

- A diverse picture emerges for the various multipliers for respondents without income information, with non-reporters in the employees' business segment having higher valuations, but with low significance levels for the multipliers.

[55] International Transport Forum (ITF) (2015) 'Valuing Convenience in Public Transport. Roundtable Report 156'. OECD. Paris, France.

- The valuations are lower for younger respondents for commute and other non-work.
- The valuations for commute trips with shared costs are lower.
- The valuations for other non-work for people without full time employment are lower.
- The valuations are lower for employees' business trips where the company would not buy time savings or for self-employed where costs are not covered.

trip covariates

- Valuations are lower for commuters on trips with overnight stays, and for employees' business travel with more than two nights away from home.
- Valuations are lower for employees' business travel when accompanied by two or more other travellers.
- Valuations are higher for employees' business travel when the 'other PT' leg is an access trip.
- Valuations are lower for other non-work trips with a frequency of less than twice per week, but higher for peak time trips.
- Commute valuations are higher on trips that are non-London based and have a rural origin or destination.

design covariates

- We observe lower valuations when the cheap option is presented on the left for SP1 for commute and other non-work, and for SP2 for commute and employees' business.
- The scale for SP2 is higher for other non-work when SP2 is presented before SP3.

SP2 specific effects

- The negative value for α show risk seeking behaviour for commute and employees' business.
- There is an overall preference for alternatives with constant travel times, i.e. no variability.
- There is a penalty for late arrival for commute trips.
- All else being equal, there is a preference for the left option for other non-work.

scale parameters

- With the models incorporating random scale heterogeneity, the values for the scale parameters now relate to the parameters for the normally distributed logarithms of the scale parameter.
- We observe significant random heterogeneity in scale across respondents for all three purposes.

dBF parameters

- We observe size effects for time in SP1 for employees' business and other non-work, and for cost in SP1 for commute and employees' business.
- There is asymmetric damping for time in SP1 for commute.
- There are sign effects (gain-loss asymmetry) for time in SP1 for commuters and other non-work, and for cost in SP1 only for commute.

Table 4.13: Estimation results for joint 'other PT' models

	Commute		**Employees' business**		**Other non-work**	
Respondents	565		206		481	
Observations	7030		2,515		6,115	
Final LL	-3,512.9		-1,205.37		-3,064.08	
Adj. ρ^2	0.27		0.29		0.27	
parameters of base θ_0 distribution	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$a_{\log(\theta_0)}$	0.2065	1.57	0.1933	0.95	0.3498	1.81
$b_{\log(\theta_0)}$	3.1141	15.61	2.6011	10.69	2.8908	14.42
game specific θ_0 multipliers	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
SP1 travel time	1	-	1	-	1	-
SP2 travel time	1.1200	1.27	1.0622	0.46	1.0659	0.61
SP2 early delay	-2.6879	-5.35 (vs -1)	-1.7667	-1.68 (vs -1)	-3.1717	-5.34 (vs -1)
SP2 late delay	1.9652	4.61	2.0722	2.93	2.3928	4.61
SP3 plenty of seats free and did not have to sit next to anyone	0.7812	-2.46				
SP3 a few seats free but had to sit next to someone/could not sit with people travelling with	0.7926	-2.31	0.8401	-1.01	0.8699	-1.35
SP3 a few seats free but had to sit next to someone/could not sit with people travelling with, some standing	0.8210	-1.96				
SP3 no seats free – a few others standing	0.9274	-0.73	0.9868	-0.07	0.9594	-0.37
SP3 no seats free – densely packed	1.3983	2.74	1.4960	1.90	1.6284	3.26
key elasticities	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
income elasticity (λ_{income})	0.1170	1.26	0.3948	3.36	0.2707	3.96
distance elasticity ($\lambda_{distance}$)	0	-	0	-	0	-
cost elasticity (λ_{cost})	0.4088	3.19	0	-	0.2102	2.02

	Commute		Employees' business		Other non-work	
time elasticity (λ_{time})	-0.2674	-2.82	-0.4877	-3.61	-0.2174	-2.77
traveller covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
unstated income ($\zeta_{income\ not\ stated}$)	0.3786	-2.97	6.1003	0.68	0.5735	-2.36
unknown income ($\zeta_{income\ unknown}$)	0.7388	-0.67			0.6612	-0.65
refused income ($\zeta_{income\ refused}$)	0.6576	-1.56	4.4233	1.71	0.5594	-2.38
aged 17-20 (base=30+)	0.6849	-2.11				
aged 17-20 (base=21+)					0.6465	-4.31
aged 21-29 (base=30+)	0.7042	-2.65				
travel costs shared (base=any other)	0.3638	-3.00				
not full time employee (base=any other)					0.6733	-3.95
Company would not buy time savings (base=any other)			0.5512	-3.78		
Self-employed costs not covered (base=costs covered)			0.4940	-4.28		
trip covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
1+ nights away (base=day return)	0.5902	-1.81				
2+ nights away (base=day return or 1 night away)			0.1980	-6.09		
2+ additional travellers (base=one or no additional traveller)			0.3539	-3.28		
access trip (base=main mode)			2.1088	2.71		
frequency less than twice per week (base=more than twice per week)					0.7455	-2.57
peak travel (base=off-peak travel)					1.2248	1.61
trip with rural component and non-London (base=any other)	1.7790	1.49				
design covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
SP1 cheap option on left (multiplicative effects coding)	0.9407	-1.86			0.9483	-1.60
SP2 cheap option on left (multiplicative effects coding)	0.8872	-2.01	0.7698	-3.59		
SP2 scale (μ_{SP2}) multiplier if SP2 before SP3 (multipl. effects coding)					1.1272	1.59

	Commute		Employees’ business		Other non-work	
SP2 specific effects	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
risk averseness parameter (α)	-0.0634	-1.56	-0.0601	-1.65		
constant for alternatives with zero variability (expressed in £)	0.3606	5.29	0.6036	3.41	0.2740	6.48
late penalty (expressed in £)	-0.1814	-2.00				
constant for left option (expressed in £)					0.0658	2.57
scale parameters	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$\mu_{log(\mu_{SP1})}$	0.7775	9.87	1.1329	8.34	0.7727	7.90
$\mu_{log(\mu_{SP2})}$	1.9019	14.21	1.7755	9.46	1.8539	25.12
$\mu_{log(\mu_{SP3})}$	1.8976	17.72	1.9207	11.91	1.8893	17.93
$\sigma_{log(\mu)}$	0.6577	9.82	0.7007	5.47	0.6692	8.63
dBF parameters	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$\beta_{t,SP1}$			-0.1536	-1.86	-0.1171	-1.61
$\beta_{c,SP1}$	0.1409	2.55	0.1742	2.68		
$\gamma_{t,SP1}$	-0.1133	-1.67				
$\eta_{t,SP1}$	0.1422	2.53			0.1057	1.77
$\eta_{c,SP1}$	0.0715	1.38				

4.8.4 Results for bus

The results for the bus models are presented in **Table 4.14**, combining data across SP1, SP2 and the SP3, where SP3c refers to crowding games and SP3nc to the time components games.

The findings can be summarised as follows:

parameters of base θ_0 distribution

- The models retrieve significant random heterogeneity for both purpose segments as can be seen by the statistically significant and large estimates for $b_{\log(\theta_0)}$, which gives the range of the underlying uniform distributions.

game specific θ_0 multipliers

- The multipliers are all of the correct sign, with the negative multipliers for early arrival relating not to schedule delay but to a reduction in trip time, which is desirable.
- As discussed above (cf. equation (4.46)), the game specific multipliers cannot directly be interpreted as the differences in the valuations across games as these differences are a function of Δt, where the impact of this differs across the valuations in the different games.
- We can however draw some conclusions in relation to the crowding multipliers in SP3c as the same dBF parameters apply to all these multipliers. We note a monotonic increase in sensitivity to different levels of crowding for both segments.
- Again drawing reference to the ITF (2014) report cited in the discussion of ‘other PT’ above, the crowding multipliers for bus reported in **Table 4.14** are generally more conservative than official guidance. However, it is noticeable that the bus multipliers are consistently higher than those for ‘other PT’, especially so for standing travel at high passenger densities.

key elasticities

- A significant positive income elasticity on the various VTT measures was only obtained for other non-work trips.
- A positive distance elasticity is observed for both segments, but with low significance levels.
- There are significant positive elasticities on the valuations in relation to cost for both segments, and negative elasticities in relation to time. This implies that the VTT is higher for longer distance trips (which tend to be more expensive and take longer).

traveller covariates

- Two of the multipliers for respondents without income information are significant and several values are large enough to justify their inclusion with a view to avoiding bias.
- The valuations are lower for commuters in ‘other’ employment.

- The valuations are higher for commute trips when the company or another party pays for travel.

trip covariates

- Valuations are lower for commuters on one-way trips.
- Valuations are lower for other non-work trips with a frequency of less than three times per month, for weekend travel and for respondents who currently experience more dwell time on their trips.

design covariates

- We observe lower valuations when the cheap option is presented on the left for SP2 for other non-work.

SP2 specific effects

- There is an overall preference for alternatives with constant travel times, i.e. no variability.

scale parameters

- With the models incorporating random scale heterogeneity, the values for the scale parameters now relate to the parameters for the normally distributed logarithms of the scale parameter.
- We observe significant random heterogeneity in scale across respondents for all three purposes.

dBF parameters

- No significant estimates were obtained for any of the dBF parameters, potentially as a result of shorter overall trips than for other modes and a more homogeneous group of reference trips.

Table 4.14: Estimation results for joint bus models

	Commute		Other non-work	
Respondents	334		482	
Observations	5,010		7,230	
Final LL	-2,719.95		-3,831.33	
Adj. ρ^2	0.21		0.23	
parameters of base θ_0 distribution	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$a_{\log(\theta_0)}$	-0.1319	-0.52	-0.1227	-0.48
$b_{\log(\theta_0)}$	3.3893	12.22	4.0826	14.69
game specific θ_0 multipliers	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
SP1 travel time	1	-	1	-
SP2 travel time	1.4678	2.15	1.0521	0.37
SP2 early delay	-3.9541	-4.13 (vs -1)	-3.3623	-3.45 (vs -1)
SP2 late delay	4.2279	4.81	2.6525	4.37
SP3c plenty of seats free and did not have to sit next to anyone	0.9411	-0.34	0.7745	-1.30
SP3c a few seats free but had to sit next to someone/could not sit with people travelling with	0.9840	-0.09	0.7832	-1.07
SP3c a few seats free but had to sit next to someone/could not sit with people travelling with, some standing	1.1015	0.52	0.9294	-0.39
SP3c no seats free – a few others standing	1.3604	1.64	1.2114	0.92
SP3c no seats free – densely packed	2.3565	3.75	2.1522	2.81
SP3nc free-flow time	1.0912	0.36	1.1342	0.44
SP3nc slowed down time	1.5358	1.50	1.2685	0.88
SP3nc dwell time	0.7466	-0.39	1.4558	0.93
SP3nc headway	1.8543	2.64	1.4840	1.87
key elasticities	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
income elasticity (λ_{income})	0	-	0.4572	4.22
distance elasticity ($\lambda_{\text{distance}}$)	0.1525	1.88	0.0716	1.36

	Commute		Other non-work	
cost elasticity (λ_{cost})	0.5231	3.14	0.5653	3.55
time elasticity (λ_{time})	-0.5762	-4.39	-0.3470	-2.94
traveller covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
unstated income ($\zeta_{income\ not\ stated}$)	1.3895	0.54	0.3709	-3.71
unknown income ($\zeta_{income\ unknown}$)	0.3644	-4.42	0.7337	-1.04
refused income ($\zeta_{income\ refused}$)	1.1882	0.43	0.7332	-1.38
other employment (base=any other)	0.4640	-4.09		
Company or other party pays (base=respondent pays)	1.9992	1.30		
trip covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
one way trip (base=return trip)	0.5241	-2.66		
frequency less than three times per month (base=three or more per month)			0.6196	-3.61
weekend travel (base=weekday travel)			0.7241	-1.98
share of dwell time for reference trip			0.3824	-2.67
design covariates (multipliers on θ unless stated)	**est.**	**rob t-rat. (1)**	**est.**	**rob t-rat. (1)**
SP2 cheap option on left (multiplicative effects coding)			0.8298	-2.08
SP2 specific effects	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
constant for alternatives with zero variability (expressed in £)	0.2441	5.49	0.2019	5.40
scale parameters	**est.**	**rob t-rat. (0)**	**est.**	**rob t-rat. (0)**
$\mu_{log(\mu_{SP1})}$	0.4919	5.39	0.5387	5.81
$\mu_{log(\mu_{SP2})}$	1.7508	19.75	1.6278	20.82
$\mu_{log(\mu_{SP3c})}$	1.6195	12.25	1.5227	12.83
$\mu_{log(\mu_{SP3nc})}$	1.5671	11.39	1.7334	14.62
$\sigma_{log(\mu)}$	0.5721	8.00	0.7056	10.41

4.8.5 Covariates not found to be significant

It should be noted that throughout the modelling analysis, little or no effect was found in most models in relation to:

- time use
- geography (i.e. area)
- current journey conditions and current road types

We interpret the lack of effect of geography as an indication that most of the differentiation the VTT across regions is in fact due to differences in key covariates such as income and trip length effects, which we capture directly in our models.

In relation to time use, current journey conditions and current road types, we believe that our results should not necessarily be interpreted as saying that these factors do not affect the VTT in real life. Instead, we believe that, when faced with abstract choices in a SP context, it is difficult for a respondent to make a "leap of faith" and relate the presented choices to the specific context of the real life journey that the choices are based around. A key example of this is journey conditions. The SP3 results clearly show major differences in VTT depending on congestion, but congestion on the reference trip plays only a very small role in the heterogeneity in VTT measures in the behavioural models, and even then only for some segments.

5 Development of 'Auxiliary' Model Specifications

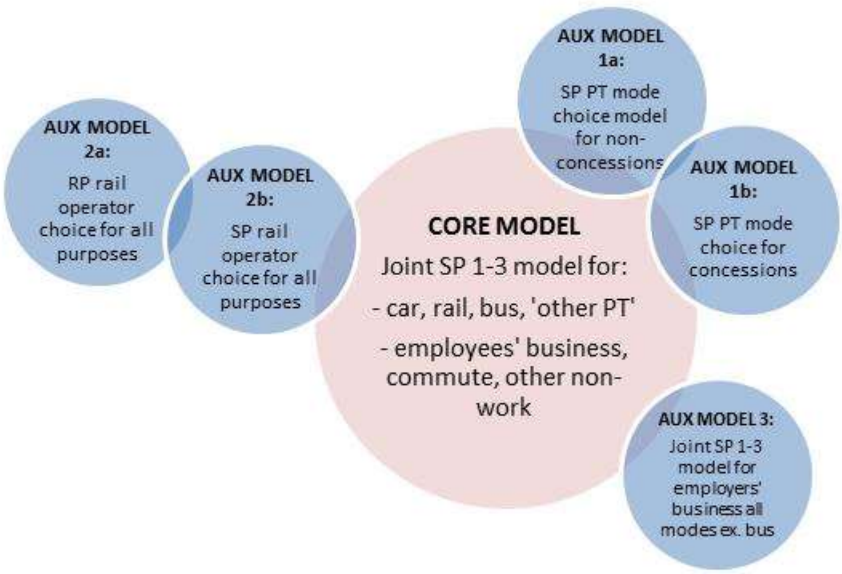
5.1 Context and overview of model development process

Whereas **Chapter 4** has set out the process for developing the 'core' behavioural model specifications – which cover most of the study scope – the present chapter will describe a number of 'auxiliary' model specifications which cover specific additional strands, namely:

- SP-based public transport mode choice, for both concessionary and non-concessionary travellers.
- Rail operator choice, both RP and SP.
- SP-based abstract choice for employers' business.

With reference to **Figure 5.1**, it can be seen that these different strands relate differently to the core specifications, and to each other.

Figure 5.1: Schematic showing organisation of model specifications



In particular:

- The two PT mode choice models are based around a common choice context, but distinguish between concessionary travellers who face a zero fare for bus and non-concessionary travellers who pay fares as normal. These models were subject to the same cleaning criteria as the core models. These models are described more fully in **Section 5.2** to follow.
- The rail operator choice models also based around a common choice context (**Section 5.3**). Whereas the SP-based model was cleaned in the same manner

as the core model, the RP-based model was subject to different cleaning criteria which are detailed below.

- The employers' business SP model was designed in broadly the same manner as the employee model, and was subject to the same cleaning criteria, but was administered to company representatives rather than to the travellers themselves (**Section 5.4**).

5.2 Specification and estimation of PT mode choice models

This section summarises the results from two strands of SP modelling outside of the core modelling reported in **Chapter 4**, namely:

- non-concessionary PT mode choice
- concessionary PT mode choice

As noted above, the data supporting these models was cleaned and subjected to diagnostic testing as part of the process covering all SP data (**Section 4.2)**.

Both of these strands entailed departure from the SP approach underpinning the core modelling, by considering a binary mode choice rather than a binary abstract choice. Because of constraints on the time available to undertake modelling, and certain technical reasons (which will be expanded upon below), the modelling of these alternative choice contexts is rather less developed than the core modelling. The results that follow are based upon a multiplicative random utility specification (i.e. analogous to the specification applied to SP2 and SP3, which is outlined in **Section 4.4.3**), but it should be qualified that:

- The dBF analysis of size and sign effects has not been implemented.
- The models reported are here of the MNL form, and Mixed Logit has not been implemented.

Furthermore, the results for these models represent 'raw' behavioural estimates, and have not been subject to post-processing by the Implementation Tool (see **Chapter 7**). Given these distinctions, it is difficult to objectively compare VTT and associated valuations reported in this section with those reported in the core modelling section (**Chapter 4**).

5.2.1 Non-concessionary PT mode choice

As outlined in **Chapter 2**, there were two variants of this SP game ('other PT' vs. bus and 'other PT' vs. rail), and the game was administered as SP3 for around 25% of the 'other PT' sample – the remaining 75% were instead given the crowding game for SP3 which was used in the joint models reported in **Section 4.8**. **Table 5.1a** shows the results for all models, where, for employees' business, the paucity of data available forced us to develop a single model combining the 'other PT' vs. bus and 'other PT' vs. rail choices.

The results show the expected pattern that the VTT for employees' business is higher than for non-work. For the reasons outlined above, caution should be exercised in reading too much into the absolute valuations, but it might be noted that the estimates of VTT in **Table 5.1a** are greater than (less than) the appraisal values that will subsequently be reported for bus (‘other PT’) in **Chapter 7**. Of course, the latter appraisal values are subject to post-processing by the Implementation Tool, whilst the modelled values in **Table 5.1a** are not.

One of the motivations for analysing mode choice was to estimate headway values. Mindful that such values could vary by the level of headway, we experimented with various non-linear functional forms, and found that an additive specification including linear and squared terms for headway fitted the data better than exponential or logarithmic transformations, and little worse than a specification based on dummy variables (which would be less amenable to deriving valuations).

Table 5.1b gives headway values arising from the model given in **Table 5.1a**, by 5 minute increments from zero to 90 minutes in the case of ‘other PT’ vs. bus and from zero to 120 minutes in the case of ‘other PT’ vs. rail[56]. The general pattern that we observe is the value of headway increasing with headway to a maximum (70 minutes for ‘other PT’ vs. bus, and 115 minutes for ‘other PT’ vs. rail) before then decreasing. It is worth remarking that this pattern is a direct consequence of the adopted specification, and the overriding result is that values of headway stabilise at longer headway times.

With reference to ‘other PT’ vs. bus, we get average (across all increments) headway multipliers of 0.50 for commute and 0.67 for other non-work, while for ‘other PT’ vs. rail the corresponding multipliers are 0.59 and 0.95, and the combined multiplier for employees’ business is 0.54. These multipliers fall within the ballpark of existing evidence; see for example the comprehensive review of international evidence in the ITF (2014) report[57], as well as the UK-focussed (albeit now somewhat dated) TRL 595 (2004) report[58]. However, it is notable that:

- In conflict with existing evidence, the headway multipliers for other non-work are markedly higher than those for commute and employees’ business.
- Consistent with existing evidence, the multipliers for the mode choice involving rail are higher than those for the mode choice involving bus.
- Mindful that existing evidence suggests that headway multipliers reduce markedly with distance, the multipliers reported here fall within the upper extremes of existing evidence – for shorter trips of less than 10 miles.

Whilst we attempted to estimate income, distance, time and cost elasticities for all models, this was not always supported by the data. Income elasticities were estimated for the ‘other PT’ vs. bus (both commute and other non-work) and for the combined modes employees’ business models, but found to be significant only for the latter case. For ‘other PT’ vs. bus, the cost elasticity is significantly different from zero for other non-work for not (quite) significant for commuting; this could reflect the fact that other non-work is more discretionary. By contrast, the cost elasticity is significant for ‘other PT’ vs. rail, both for commuting and other non-work. Where estimated, the distance and time elasticities are not significant.

---

56 This is because of the different ranges of boundary values used for different modes.

57 International Transport Forum (ITF) (2015) ‘Valuing Convenience in Public Transport. Roundtable Report 156’. OECD. Paris, France.

58 TRL (2004) ‘The demand for public transport: a practical guide’. TRL report 563.

Table 5.1a: Estimation results for non-concession mode choice models

	Commute				Employees' business		Other non-work			
	'Other PT' vs. bus		'Other PT' vs. rail		Combined games		'Other PT' vs. bus		'Other PT' vs. rail	
Respondents	180		109		120		128		92	
Observations	900		545		600		640		460	
Final LL	-407.02		-238.34		-293.99		-308.83		-208.57	
Adj. ρ^2	0.33		0.35		0.27		0.28		0.32	
	est.	**rob.t-rat. (0)**	**est.**	**rob.t-rat. (0)**	**est.**	**rob.t-rat. (0)**	**est.**	**rob.t-rat. (0)**	**est.**	**rob.t-rat. (0)**
value of free-flow time (£/hr)	4.7084	5.4945	5.7413	2.2396	6.773	4.1503	4.7826	4.0052	4.5516	1.8342
value of headway - linear term	5.8181	3.699	5.9872	2.3692	4.7508	4.1535	8.1656	3.0277	7.4811	2.3622
value of headway - square term	-2.7415	-2.1576	-1.9962	-1.9918	-0.90687	-1.5076	-4.0309	-2.0685	-2.4409	-1.9353
μ	9.1748	10.488	8.4406	6.7001	8.1092	7.7524	7.2047	7.4232	5.7949	4.2841
WTP for bus (£)	-0.5459	-3.0915	-	-	-	-	-0.9768	-2.798	-	-
WTP for bus/rail (£)	-	-	-	-	-0.4824	-2.5266	-	-	-	-
WTP for rail (£)	-	-	-0.5242	-1.4973	-	-	-	-	0.2664	-0.7990
income elasticity (λ_{income})	0.1971	1.5909	-	-	0.73859	3.2424	0.0735	0.4496	-	-
multiplier for unstated income	0.3777	1.9771	-	-	-	-	0.2876	1.3834	-	-
multiplier for refused income	1.7186	1.2868	-	-	-	-	0.35241	3.1658	-	-
distance elasticity ($\lambda_{distance}$)	-0.0967	-1.0535	-0.0664	-0.3819	-	-	-0.1259	-1.4829	-0.0792	-0.05
cost elasticity (λ_{cost})	0.6283	1.9314	0.7919	2.2369	0.21814	0.7275	1.1796	4.3319	0.9731	2.33
time elasticity (λ_{time})	0.0002	0.00123	0.5266	1.5267	0.00254	0.0126	0.1558	0.7388	0.4654	0.82

Table 5.1b: Implied values of headway for different levels of headway

	Commute				Employees' business		Other non-work			
	'Other PT' vs. bus		'Other PT' vs. rail		Combined games		'Other PT' vs. bus		'Other PT' vs. rail	
Minutes	Headway (£/hr)	s.e	Headway (£/hr)	s.e	Headway (£/hr)	s.e	Headway (£/hr)	s.e	Headway (£/hr)	s.e
5	0.47	0.123	0.49	0.204	0.39	0.094	0.65	0.212	0.61	0.256
10	0.89	0.229	0.94	0.394	0.77	0.185	1.25	0.399	1.18	0.496
15	1.28	0.320	1.37	0.571	1.13	0.273	1.79	0.562	1.72	0.720
20	1.63	0.395	1.77	0.735	1.48	0.360	2.27	0.700	2.22	0.929
25	1.95	0.455	2.15	0.885	1.82	0.446	2.70	0.815	2.69	1.123
30	2.22	0.501	2.49	1.023	2.15	0.531	3.08	0.907	3.13	1.301
35	2.46	0.535	2.81	1.147	2.46	0.616	3.39	0.978	3.53	1.465
40	2.66	0.556	3.10	1.258	2.76	0.701	3.65	1.030	3.90	1.614
45	2.82	0.568	3.37	1.356	3.05	0.788	3.86	1.064	4.24	1.749
50	2.94	0.575	3.60	1.442	3.33	0.877	4.01	1.084	4.54	1.871
55	3.03	0.579	3.81	1.515	3.59	0.968	4.10	1.094	4.81	1.980
60	3.08	0.587	3.99	1.575	3.84	1.062	4.13	1.100	5.04	2.076
65	3.09	0.606	4.14	1.624	4.08	1.159	4.12	1.109	5.24	2.160
70	3.06	0.643	4.27	1.661	4.31	1.260	4.04	1.130	5.41	2.234
75	2.99	0.704	4.36	1.686	4.52	1.366	3.91	1.173	5.54	2.297
80	2.88	0.792	4.43	1.700	4.72	1.477	3.72	1.248	5.64	2.351
85	2.74	0.908	4.48	1.704	4.91	1.593	3.48	1.362	5.70	2.398
90	2.56	1.051	4.49	1.699	5.09	1.715	3.18	1.519	5.73	2.438
95			4.48	1.684	5.25	1.843			5.73	2.474
100			4.43	1.662	5.40	1.978			5.69	2.507
105			4.36	1.633	5.54	2.119			5.62	2.540
110			4.27	1.600	5.66	2.267			5.51	2.575
115			4.14	1.565	5.77	2.422			5.37	2.614
120			3.99	1.530	5.87	2.584			5.20	2.661

5.2.2 Concessionary PT mode choice

As outlined in **Chapter 2**, this was a game designed specifically for concessionary passholders who incurred a zero fare for bus travel. In order to introduce time/money trade-offs to this context, the game entailed a choice between a relatively slow trip by bus at zero fare and a relatively fast trip by train at non-zero fare. The survey was targeted at micro-locations in West Yorkshire, South Yorkshire, Bristol and Brighton where such a choice exists in practice.

Given this focus on a mode choice involving bus, the scope of the survey was restricted to commuting and other non-work – since business was deemed to be out of scope for bus (see earlier discussion of the scope of the study in **Chapter 1**). In practice, only 7 commuters were identified. The rest, 113 individuals, travelled for non-work purposes. Moreover, the expectation was that this survey would capture mainly retired travellers. Given the modest sample size for this aspect of the survey (600 observations in total), the models reported in **Table 5.2** are pooled across the commute and other non-work purposes.

Table 5.2: Results for concessionary PT mode choice (commute and other non-work only)

	Includes obs where rail option is free		Excludes obs where rail options is free	
Respondents	120		74	
Observations	600		370	
Final LL	-248.45		-112.73	
Adj. ρ^2	0.39		0.54	
	est.	**rob. t-rat.**	**est.**	**rob. t-rat.**
VTT (£/hr)	1.2759	0.81	2.5415	0.94
Value of headway (£/hr)	4.5946	1.39	2.2475	0.97
Constant for bus	1.2598	3.72	0.5531	0.44
μ	0.8616	4.16	1.7495	4.94

An unfortunate finding was that, despite our best efforts to select micro-locations which offered a real choice between a zero fare bus trip and a non-zero fare rail trip, 230 of the 600 observations (some 38% of the sample) entailed a zero fare for the rail reference trip, as well as for the bus reference trip. Given the use of a 'pivot' approach for the SP design, the practical implication of this was that, for these 230 observations, zero fares were presented for both bus and rail throughout the SP games. In other words, there was no time vs. cost trade-off in the SP, and the only relevant trade-off was between in-vehicle time and headway.

Against this background, **Table 5.2** reports two models, one based on the full sample of 600 observations, and a second model on a restricted sample of 370 observations which entailed a non-zero fare for rail. Regrettably, little or nothing can be inferred from either model, since values of travel time saving and headway are insignificant in both models.

This is a result of very high rates of non-trading in the data, with respondents consistently choosing the free bus option, independently of the size of time savings they would obtain with the paid rail option. This high rate of bus inertia also led us to include a constant for bus outside the logarithmic transform in the multiplicative model, with a view to capturing the market shares. The presentation of cheaper rail options would potentially have avoided this issue to some extent, but our findings do highlight the strong modal inertia for concessionary travellers and the appeal of free travel.

5.3 Specification and estimation of operator choice models (RP and SP)

Mindful that SP represented the main analytical approach deployed by the present study, RP analysis was undertaken with the objective of seeking to validate the SP analysis. As discussed in **Chapter 2**, the Phase 1 research concluded that it would be most efficient and cost-effective to focus RP analysis on the rail mode, targeting specific locations where:

- there are real-world situations where spending money saves time;
- travellers are familiar with the opportunities available; and
- time vs. cost trade-offs offer real choices.

Corresponding to the rail SP3a experiment, we designed the RP survey around rail operator choices in carefully selected locations, collecting time, cost and headway data for the current trip (i.e. the chosen operator), as well as for the alternative trip (i.e. the alternative operator). That is to say, both the RP and SP considered the same basic choice context, namely a choice between operators for a trip to London, where one operator was cheaper but slower than the other(s), and where there were also potentially variations in headway and comfort.

As was explained in **Chapter 2**, we selected the following O-D trips as the basis for the RP experiment:

- Birmingham to London (served by three operators, Virgin, London Midland and Chiltern)
- Stoke to London (served by Virgin and London Midland)
- Peterborough to London (served by East Coast and Great Northern)

5.3.1 Revealed Preference data, cleaning and model specification

An overview of the RP dataset including descriptive statistics was reported in **Section 3.5**. Key observations which inform the modelling include the following:

- Of the 2,646 travellers surveyed, 50% were travelling on employees’ business, 17% were commuters, and 33% were travelling for other non-work purposes. Therefore, compared to the SP sample generally, this survey was more targeted on business travel.
- As might have been anticipated, the employees’ business and commuter samples had much higher household incomes than the other non-work sample: 64% of employees’ business and 58% of commuters reported household incomes of £50,000 per year or more, as compared with 31% for the other non-work sample.

- When intercepted, about two-thirds of travellers were on the outward leg of their trip; we will return to this point in what follows.
- Half the other non-work sample, four-tenths of the employees' business sample, and a quarter of the commuter sample, held advance purchase tickets – meaning that they were constrained to specific operators and departures. Relatively small proportions of travellers held full fare tickets (25% employees' business, 20% commuters and 8% other non-work), and 29% of rail commuters held season tickets.
- RP respondents were asked whether they compared the times and fares of operators prior to travelling. For the Peterborough flows, 43% of those on employees' business and commuting trips, and 61% of those on other non-work, said that they had made such a comparison. For the West Coast flows, 52% of those on employees' business, 43% on commuting trips and 61% on other non-work trips, said that they had made a comparison. Again, this will be an important point in what follows.

Consistent with the remainder of this study, the models employed in the analysis of the RP data used a multiplicative random utility specification, where the specification corresponded to that used for SP2 and SP3 (**Section 4.4.3**). With the amount and quality of the data available, only fixed coefficients models were estimated, thus not allowing for any random heterogeneity in valuations across respondents.

The RP modelling proved to be challenging, and a significant amount of cleaning was required in order to produce anything approaching a viable model. With reference to **Table 5.3**, cleaning was conducted in two stages, involving basic cleaning followed by additional cleaning.

Basic cleaning

The raw dataset received from Accent consisted of 3631 observations. This was subject to initial checks to confirm the validity of the data, prompting three basic actions.

1. Where possible, missing data (e.g. headway and travel time missing on one leg of journey) was in-filled using other data (e.g. headway and travel time from the other leg of journey).
2. Where possible, missing data was accommodated by adjusting the formulation of relevant observations (e.g. setting the third operator as 'unavailable' if data was only complete for two operators).
3. Where missing data could not be dealt with through actions 1 and 2 above, or extreme times or costs were reported by the respondent, data was removed altogether.

Basic cleaning resulted in the removal of 811 observations.

Table 5.3: Summary of RP data cleaning

Stage	Criteria	Description	Incremental number of exclusions	Remaining observations
Raw data				3631
Initial cleaning		Assemble, clean and check choice data. Make corrections to data to allow for obvious errors. Use sensible headways and times on one leg where missing for other leg. Where 2 operators, remove where missing information. Where 3 operators, make unavailable where missing information. Remove very large or small times and costs.	811	2820
Additional cleaning	Inconsistent purpose	State purpose is commute but frequency of trip is less than once a week or the respondent did not hold a season ticket.	2	2818
	Availability	Chosen alternative of the respondent has missing level of service attribute.	18	2800
	Out of scope	Respondent is not in scope of the time-cost trade-off. Examples: has been given a ticket by Virgin, possesses staff pass, no alternatives at that time.	63	2737
	Inconsistent differences in Out of Vehicle Times	Respondent has reported an unusual difference in Out of Vehicle Time (OVT) between the available alternatives high (>60 minutes).	118	2619
	Inconsistent differences in Out of Vehicle Costs	Respondent has reported an unusual difference in Out of Vehicle Cost (OVC) between the available alternatives (>£20).	43	2576
	Dominance check	Respondent has not chosen the alternative which is better in terms of time, cost and headway (quicker, cheaper and more frequent).	47	2529
	No choice	Respondent has only one alternative (unchosen alternatives have missing attributes).	51	2478
	Unsure of time or cost	Respondent is ‘very unsure’/ ‘unsure’/ ‘neither sure nor unsure’ about the reported time or cost of alternatives.	1753	725

Additional cleaning

Having undertaken basic cleaning, and assessed the plausibility of a preliminary model estimated on the dataset of 2820 observations, we then applied additional cleaning criteria incrementally, each time re-assessing the plausibility of the resulting model. In what follows, we will discuss each incremental criterion, but report only the model estimates for the final recommended model (i.e. once all cleaning steps have been applied).

1. Inconsistent purpose

In the RP questionnaire, respondents were asked for their trip purpose (and probed if there had been a mismatch with the purpose recorded in the recruitment questionnaire). Unfortunately, there proved to be significant disagreement in the data with regards to the two responses on trip purpose, i.e. at the recruitment and completion stages. This prompted us to conduct further investigation, checking the ‘stated’ trip purpose (Q8b: *“What was the main purpose of your rail trip?”*) against the stated frequency of the type of trip (Q9: *“How often do you make this trip?”*).

This revealed 175 (out of 370) cases where the respondent initially stated in the RP questionnaire that they were ‘Travelling to/from work’ or ‘Travelling to/from place of education’, but later stated in the same questionnaire that they make the trip less frequently than once a week. Given this inconsistency, we reverted to the trip purpose given in the recruitment questionnaire instead of using Q8b[59].

In addition, an exclusion criterion on frequency was used (where the recruitment purpose was commute, but the frequency of the trip was less than once a week and the respondent did not possess a season ticket). This resulted in the exclusion of 2 respondents.

2. *Availability*

In some cases, respondents did not supply times, costs and/or headway for the chosen alternatives. This prompted us to remove 18 observations.

3. *Out of scope*

In a number of cases, respondents were found to be outside the scope of the survey, since they did not have an alternative operator available to them. This accounted for 63 exclusions.

4. *Inconsistent differences in OVT and OVC*

Unusually high out-of-vehicle time (OVT) and out-of-vehicle cost (OVC) were observed in some cases. This may have been due to misunderstandings regarding the scope of the survey (which was intended to be restricted to the O-Ds cited earlier). Respondents with other O-Ds are likely to have been subject to other operator constraints, and this poses the risk of introducing noise to the estimation data. For this reason, the following observations were excluded:

- Observations with differences in OVT between alternatives > 60 minutes (resulting in 118 exclusions).
- Observations with differences in OVC between alternatives > £20 (resulting in 43 exclusions).

5. *Dominance*

The data was checked carefully for inconsistency and observations were excluded where respondents stated that Virgin was slower than London Midland or Chiltern, or that East Coast was slower than Great Northern. Also, observations were removed where respondents chose one operator when another operator dominated in terms of cost, IVT, headway, OVT and OVC. These criteria accounted for 47 exclusions.

---

[59] Accent conducted back-checking of respondents throughout the SP and RP data collections looking at cases where respondents had (apparently) reported different purposes at the recruitment and completion stages. However, this back-checking revealed no systematic problems, and in the vast majority of cases the purpose reported on completion was judged to be accurate.

6. No choice

Some respondents did not state the travel times and costs of non-chosen alternatives or declared zero times and costs for non-chosen alternatives. After data cleaning, this resulted in cases where some respondents were only left with a single available alternative. 51 observations were excluded on this basis, as they were not usable for modelling.

7. Uncertainty regarding travel times and costs of non-chosen alternatives

Of fundamental importance, given the particular RP approach adopted here, was that travellers were able to report – with confidence – not only the times and costs of their chosen operator, but also the times and costs of the alternative (i.e. non-chosen) operator(s). Without this information, it would be impossible to undertake discrete choice modelling and thereby infer the VTT. Against this background, we asked respondents about their level of confidence in the reported times and costs (**Table 5.4**).

Table 5.4: Uncertainty regarding attributes of non-chosen alternatives

	Level of confidence on costs of other alternatives	Level of confidence on travel time of alternative 2	Level of confidence on travel time of alternative 3 (if applicable)
1 = ‘Very sure’	237	401	74
2 = ‘Quite sure’	722	1283	257
3 = ‘Neither’	350	282	122
4 = ‘Unsure’	962	587	288
5 = ‘Very unsure’	545	263	173

The proportions of ‘unsure’ and ‘very unsure’ responses were highest in the case of employees’ business, with more than 68% of respondents being ‘neither sure nor unsure’, ‘unsure’ or ‘very unsure’ about the times and costs of non-chosen alternatives.

Several model structures were tested to account for this. The hypotheses that were tested in this regard on the employers’ business data, and the corresponding results, are presented in **Table 5.5.**

Table 5.5: Hypotheses tested concerning certainty of RP responses

	Hypothesis	Result
1	The scale parameter is different for respondents depending on levels of uncertainty (5 scale parameters).	Scale parameter was found to be significantly different but the resulting VTT was unusually high (>200£/hr).
2	The scale parameter is different for respondents depending on levels of uncertainty (5 scale parameters) and there is systematic taste heterogeneity across the 5 groups.	Encountered estimation problems due to high correlation between the scale and VTT parameters.
3	The scale parameter is the same for all observations but there is systematic taste heterogeneity across the 5 groups.	Very high and statistically insignificant VTT for the three 'unsure' groups ('neither', 'unsure', 'very unsure'). Therefore the three 'unsure' groups were dropped from further analysis. Note that this markedly reduced the number of observations.
4	The scale parameter is the same but there is systematic taste heterogeneity between the 2 groups 'very sure' and 'quite sure'. Respondents stating other levels of certainty are removed.	Encountered estimation problems. Correlation between the headway and travel times was suspected as the main source of the problems.
5	The scale parameter is the same but the VTT is different between the 'very sure' and 'quite sure' groups. Respondents stating other levels of certainty are removed.	The VTT was found to be significantly different between the groups only for the employees' business purpose.
6	Using only the 'very sure' data.	Estimation problems were again encountered, likely due to correlation between the headway and travel time attributes.

Other modelling considerations

Aside from the extensive cleaning process undertaken, it is appropriate to make a number of other comments regarding the development of the RP model.

1. Background information on attributes

The respondents were asked: *"Prior to making your trip, did you compare the fares and travel times?"* In 540 out of the 1,909 employees' business observations, the respondents replied that they did not check fares or travel times prior to making their trip (Q6B and Q6C). These proportions were much higher than for commute and other non-work, where 35 (out of 199) and 123 (out of 708) respondents respectively stated that they did not check fares or travel times prior to making their trip. Among the respondents who stated that they did not check fares or travel times prior to travelling (698 in total for all purposes), 133 had season tickets and 497 responded that they did not pay for the trips themselves (i.e. paid in full by other travellers / shared with other travellers, colleague / partner / family member paid or employer paid).

In a similar vein to the analysis of uncertainty reported above, model specifications were tested both in terms of differences in scale parameters and differences in VTT, and the former was found to yield the better results for employees' business. However, for commute and other non-work, the share of respondents who 'did not bother to check fares and travel times' was much lower than for employees' business. Subsequently, scale differences could not be estimated in the case of commute, and were statistically insignificant in case of other non-work. For the sake of consistency among purposes, the scale differences were not used in the final models.

2. *Single vs. return trip*

In the RP data, some people (92 out of 376 respondents in the cleaned data) contributed more than one observation – this arose because of reporting both legs of a return trip. Models were estimated both including and excluding the return leg, and no significant differences were observed. The return leg data were therefore retained in the final models, but account was taken of multiple observations for those respondents in the calculation of standard errors.

3. *Class of travel*

In the RP data, the class of travel (First and Standard for London Midland and Virgin, and Business and Standard for Chiltern) was available. Systematic taste heterogeneity based on class of travel was checked, but not found to be significant. Segmentation of the alternative-specific constants (ASCs) based on class of travel was also explored, and found to be significant for London Midland. Mindful however that it could present problems for model applicability, the ASC was not retained in the final specification.

4. *Who makes the decisions (employees' business)*

For employees' business, there are differences concerning who makes the decisions (Q67B1). Several specifications were tested to incorporate these in the model structure. These are presented in **Table 5.6**.

Table 5.6: Hypotheses tested on decision-making for employees' business

	Hypothesis	Result
1	There is systematic taste heterogeneity between EB respondents who make their own decisions, and EB respondents who are subject to the decision-making of their employers. The level of uncertainty on cost ('sure' vs. 'quite sure') is accounted for by scale differences.	The VTT was much lower for the case where employers make the decisions; £9.27/hr as opposed to £111.60/hr for the rest. However, the value of travel time and headway were insignificant in the former case. For the sake of consistency with SP, this segmentation was not retained in the final specification.
2	There is systematic taste heterogeneity between groups who make their own decisions independently (i.e. do not need to justify their decisions), and the rest (i.e. where others make the decisions on their behalf or need to approve their decisions). The level of uncertainty on cost ('sure' vs. 'quite sure') is accounted for by scale differences.	The difference in VTT between the two groups was not statistically significant.
3	There is systematic taste heterogeneity between respondents who: i) make their own decisions, ii) make their own decisions but are liable to the line manager, iii) are subject to company decision-making. The level of uncertainty on cost ('sure' vs. 'quite sure') is accounted for by scale differences.	Encountered estimation problems, probably due to insufficient observations.

5. Model specification

Given the data issues outlined above, the model specification was shaped by what we were able to estimate rather than what we wanted to estimate. In the final model, a single scale has been retained and the VTT has been segmented based on trip purpose. In addition, for employees' business, since the proportion of 'quite sure' responses was very high (75.5%), the VTT was further segmented based on associated uncertainty levels. Given its policy relevance, the headway time coefficient has been retained in the model despite being statistically insignificant. Sensitivities to OVT and OVC, reliability, crowding and quality of service were tested (both as continuous and categorical variables), but not found to be significant.

For employees' business, segmentation by who makes the decision, class of travel, who pays for the trip, and the form of employment (self-employed or not) were tested, but not found to be significant.

The effect of income was tested using household income for commute and other trips and found to be significant. For employees' business, the effects of both personal income and wage (derived by dividing personal income by hours per year and using average incomes for missing ones) were tested and personal income was found to better capture income effects. It might be noted that the average personal income of employees' business respondents – after cleaning – was £51,560, which is much higher than the national average, and may have been a cause of the relatively high VTT that we estimated.

Models accounting for unobserved heterogeneity were tested, both in the segmented and the un-segmented modes. In both cases, however, estimation problems were encountered due to the high standard deviation of VTT for the 'quite sure' group.

5.3.2 Stated Preference data and model specification

An overview of the SP dataset generally, including descriptive statistics, was presented in **Section 3.3**. Focussing specifically on the operator choice SP, the characteristics of the sample in terms of income and ticket type are summarised in **Table 5.7** and **Table 5.8** respectively.

We observe that income is lowest for other non-work respondents, and highest for employees' business respondents. Rates of season ticket usage are relatively low, as is perhaps to be expected given the length of the trips considered here.

Table 5.7: Income characteristics of the SP operator choice sample (no. respondents)

		Commute	Employee's business	Other non-work
	No. respondents	73	157	133
HH income	Not stated	0	5	1
	Unknown	0	0	0
	<10K	4	2	19
	10-20K	5	3	21
	20-30K	2	11	25
	30-40K	9	13	18
	40-50K	11	23	15
	50-75K	15	43	9
	75-100K	15	29	13
	<100K	12	28	11
	Refused	0	0	1
Pers income	Not stated	0	5	1
	Unknown	0	0	0
	<10K	9	3	48
	10-20K	6	7	23
	20-30K	12	24	24
	30-40K	13	31	13
	40-50K	12	28	8
	50-75K	12	30	6
	75-100K	5	17	7
	<100K	4	12	2
	Refused	0	0	1

Table 5.8: Ticket type characteristics of SP operator choice sample (no. respondents)

		Commute	Employee's business	Other non-work
First class	Weekly season	0	0	0
	Monthly season	0	0	0
	Annual season	1	0	0
	Advance	5	22	5
	Off-peak day ticket	3	4	5
	Anytime	1	1	0
	Other	0	1	0
Standard class	Weekly season	0	0	2
	Monthly season	6	1	0
	Annual season	8	0	0
	Weekly travel card	0	2	0
	Monthly travel card	0	0	0
	Annual travel card	0	0	0
	Travel card day ticket	0	0	1
	Advance	14	62	52
	Off-peak day ticket	19	40	50
	Anytime	8	23	13
	Other	8	1	5
Single/ return	Single	14	48	33
	Return	44	106	98

SP data cleaning and modelling

Unlike the RP models outlined above, the data supporting the corresponding SP models was cleaned and subjected to diagnostic testing as part of the process covering all SP data (**Section 4.2)**. Also consistent with the 'core' SP modelling, the operator choice model reported here used a multiplicative random utility specification, as already discussed in detail (**Section 4.4.3**). However, given the amount and quality of the data at our disposal, only fixed coefficients models were estimated, thereby precluding any random heterogeneity in valuations across respondents.

The data from the sampling locations were merged into a single dataset to improve estimation of purpose specific models. The estimated choices models are therefore a combination of binary (sampling at Peterborough, Rugby and Stoke-on-Trent) and multinomial choices (sampling at Birmingham New Street).

The value function comprised an additive set of operator constants, and constants denoting the current operator of the respondent. The latter in particular provided a large improvement in model fit and more reasonable valuations. The SP attributes, namely travel time, headway time and travel cost, were all included in an additive fashion.

The experimental design induced a price differential between peak and off peak travel. To account for this impact on choice behaviour, all operator constants and the main travel time component were initially interacted with a dummy variable representing peak travel. This was later dropped for comparability with the RP models, where no such information was available.

All non-cost elements in the value function, including the constants were interacted with reference time, cost and distance effects. Household income interaction effects were introduced for the commuting and other non-work purposes, whereas personal income effects were applied for employees' business, consistent with the approach used for the SP models in the core analysis. Note that size and sign effects were not implemented in the operator choice model.

Modelling the SP3a operator choice models turned out to be a challenging task. First of all, the attribute levels were found to be correlated with the operators. For example, Virgin is typically the fast and expensive option in Birmingham, whereas London Midland is the slow but cheap alternative. This caused confounding between the operator constants and the parameters for travel time savings. Confounding is a common phenomenon in RP studies, but in this case translates to the SP environment since the design was intended to mimic real world conditions. Estimating models with and without operator constants therefore had a large impact on the estimated value of time parameters.

Besides confounding, samples sizes are modest in size but rates of non-trading relatively high (14-29% in each sample). Non-trading also occurred on the fast and expensive operators, particularly in the commute and employees' business segments, and these effects serve to amplify the VTT estimates reported here. Furthermore, commute and employees' business travellers tend to use Virgin and East Coast during peak hours, which further induces confounding, and contributes to high VTT estimates. Splitting the samples into smaller cohorts to separate these effects turned out to be infeasible, due to small sample sizes in combination with non-trading.

5.3.3 Final specifications and results

For the purpose of testing the consistency between results for the RP and SP datasets, a relatively common specification was adopted for the final models. This included:

- A single scale parameters across all respondents in the given sample.
- Purpose-specific VTT measures only for travel time, where this was segmented into ‘very sure’ and ‘relatively sure’ respondents for employees’ business in the RP data.
- An income elasticity on all non-cost components, using household income for commute and other non-work, and personal income for employees’ business, with a single multiplier for all respondents with missing income.
- Constants for different operators, including inertia parameters in the SP data; any insignificant constants were dropped from the models.

OVT was dropped from the final RP models given poor performance in estimation.

The results for the final models are reported in **Table 5.9**, and give rise to the following findings:

- The VTT measures are higher overall for the operator choice games, as compared with the abstract choice games in **Chapter 4,** and this is likely be caused in part by confounding with operator effects.
- The VTT for commuters is remarkably similar in the two models, though it has a very large standard error in the RP model.
- The VTT for employees’ business is higher in the RP than the SP, though the value for those respondents who are ‘very sure’ about fares is within just one standard error of the value for SP. The value for respondents who are only ‘quite sure’ about fares is much higher, and probably should be disregarded.
- The VTT for other non-work is substantially higher in the SP data, and also attains a lower level of statistical significance in the RP data. However, the value from the RP data is much closer to that from the core SP games for rail in **Chapter 4**.
- No specific comparative insights can be gained from the operator constants.
- The income elasticity is much higher in the RP than the SP, and not significantly different from a unit income elasticity.

Table 5.9: Estimation results for RP and SP operator choice models

	SP			RP		
Respondents	363			578		
Observations	1,815			725		
Final LL	-1,029.21			-387.53		
Adj. ρ^2	0.17			0.21		
	est.	**t-rat. (0)**	**t-rat. (1)**	**est.**	**t-rat. (0)**	**t-rat. (1)**
VTT commute (£/hr)	40.59	7.09	-	39.95	0.64	-
VTT employees' business (£/hr)	41.10	7.81	-	-	-	-
VTT employees' business very sure (£/hr)	-	-	-	54.36	3.49	-
VTT employees' business quite sure (£/hr)	-	-	-	145.47	4.16	-
VTT other non-work (£/hr)	36.87	6.41	-	8.76	1.52	-
Value of headway (£/hr)	6.67	4.53	-	-0.74	-0.21	-
μ	6.89	13.95	-	4.66	9.52	-
WTP London Midland (off-peak constant only) (£)	-5.44	-2.07	-	-	-	-
London Midland inertia term Birmingham (£)	16.84	4.06	-	-	-	-
East Coast inertia term (£)	-15.35	-3.28	-	-	-	-
London Midland inertia term non-Birmingham (£)	13.72	3.23	-	-	-	-
WTP Chiltern (£)	-	-	-	8.29	1.41	-
WTP London Midland (£)	-	-	-	-3.38	-1.15	-
WTP Great Northern (£)	-	-	-	-17.03	-2.75	-
income elasticity (λ_{income})	0.34	3.37	-6.51	1.10	5.65	0.50
missing income multiplier	1.00	-	-	0.97	-	-0.04
cost elasticity (λ_{cost})	0.82	6.97	-	-	-	-
time elasticity (λ_{time})	-0.71	-4.46	-	-	-	-

5.4 Specification and estimation of employers' business model

The background to this model, which was motivated by a specific interest in business travel, is described more fully in **Chapter 6** which follows. The present discussion will restrict attention to the development of the model specification, and the 'raw' results arising from model estimation. Additional interpretation of the results will also be conducted in the following chapter. Recall from the discussion of method in **Chapter 2** that the employer survey was directed at company representatives with knowledge of the company travel policy and the travel behaviour of their employees.

For present purposes, it will suffice to note that the employer representative was presented with three SP exercises, depending upon whether they reported a car or rail trip by a relevant employee:

- Choices between time and cost only (SP1).
- Choice involving usual time, cost and a distribution of travel times (SP2).
- Choices involving travel time in various degrees of congestion and cost for car and travel time in various degrees of crowding and cost for rail (SP3).

Thus, employers were faced with essentially the same SP exercises as employees on business travel. The difference here was that the respondent was asked to indicate what the identified employee **should** have chosen, thereby revealing company willingness-to-pay.

5.4.1 Data and model specification

Separate models were estimated for car and rail trips (remembering from earlier discussion in **Chapter 3** that 'other PT' was dropped for business in the course of the field survey). In total, 387 interviews were achieved, of which 244 were car and 143 were rail.

After 'basic cleaning', as described in **Section 4.2**, the number of car users was reduced to 210 and the number of rail users to 138. These respondents yielded respectively 3150 and 2070 observations across the three SP exercises offered.

5.4.2 Final specification and results

Using the employer datasets for car and rail, a joint SP1-3 model was estimated in the same manner as for the 'core' modelling approach (**Section 0**), giving the results presented in **Table 5.10**.

We note that the goodness of fit (ρ^2) measures are modest. Contributory factors here are the absence of a large number of trip and socio-economic covariates, as well as any allowance for random taste heterogeneity. The reported models only contain effects over and above the type of time for trip distance, trip duration and trip cost.

Table 5.10: Estimation results for SP employers' business models

	Car			Rail		
Respondents	210			138		
Observations	3150			2070		
Adj. ρ^2	0.130			0.103		
Final LL	-1879.57			-1261.74		
	est.	**t-rat. (0)**	**t-rat. (1)**	**est.**	**t-rat. (0)**	**t-rat. (1)**
VTT (£/hr)	5.898	4.24	3.52	4.544	4.50	3.51
mult_TSP2	1.266	5.80	1.22	-	-	-
mult_FF	0.077	0.50	5.91	-	-	-
mult_LC	0.281	2.08	5.33	-	-	-
mult_HC	1.065	3.35	0.20	-	-	-
mult_SDTT	0.209	1.81	6.87	-	-	-
mult_SP2_time	-	-	-	1.613	5.61	2.14
mult_SP2_early	-	-	-	-3.627	2.18	2.78
mult_SP2_late	-	-	-	2.903	2.53	1.66
mult_SP3c_seating1	-	-	-	1.153	4.58	0.61
mult_SP3c_seating2	-	-	-	1.139	4.44	0.54
mult_SP3c_seating3	-	-	-	1.399	4.63	1.32
mult_SP3c_seating4	-	-	-	1.499	4.60	1.54
mult_SP3c_seating5	-	-	-	1.719	4.73	1.98
mult_SP3c_standing1	-	-	-	1.782	4.04	1.77
mult_SP3c_standing2	-	-	-	1.505	3.73	1.25
mult_SP3c_standing3	-	-	-	2.129	4.39	2.32
mult_SP3c_standing4	-	-	-	2.299	4.14	2.34
mult_SP3c_standing5	-	-	-	3.117	4.79	3.26
mult_inc_unknown	1.317	3.28	0.79	-	-	-
mult_inc_refused	1.406	2.41	0.70	-	-	-
distance elasticity ($\lambda_{distance}$)	-0.427	2.58	8.61	-	-	-
cost elasticity (λ_{cost})	0.241	0.71	2.23	1.029	8.43	0.23
time elasticity (λ_{time})	0.798	1.56	0.40	-0.562	2.83	7.86
μ_SP1	0.577	6.71	4.91	1.110	11.22	1.11
μ_SP2	5.581	6.45	5.29	8.122	5.32	4.67
μ_SP3	2.149	4.56	2.44	6.118	4.51	3.77
β_t_SP1	-0.561	1.90	5.30	-	-	-
β_t_SP2	-0.150	1.71	13.12	-	-	-
γ_c_SP1				0.116	2.43	18.42
γ_c_SP3	-0.525	3.41	9.91	-	-	-
mult_scale_SP2_first_effects_SP3	0.769	6.76	2.03	-	-	-
mult_cheap_left_effects_SP1	-	-	-	0.911	12.07	1.18
mult_cheap_left_effects_SP3c	-	-	-	0.439	3.98	5.09
mult_scale_SP2_first_effects_SP3c	-	-	-	0.790	4.81	1.28
asc_left_SP2	-0.167	1.89	13.23	-	-	-
no_variability_SP2	-0.392	1.64	5.81	0.936	2.65	0.18
late_penalty_SP2	-	-	-	-1.874	1.82	2.78

Our view is that the inability to detect significant effects for systematic and random taste variation was due in large part to the limited sample sizes. We were able to discern some promising effects for variables listed above but they were not precisely estimated, although we recognised from the outset that the sample size that could be afforded was unlikely to support detailed segmentation analysis. This is in the context where we would expect our sample of employers to have widely divergent willingness-to-pay, not least because of different employee income levels and productivity, and hence this inability to capture such heterogeneity will explain the relatively poor goodness of fit.

We should also point out that, given that the conventional Cost Saving Approach (CSA) is linked directly to wage rates, we were unable to here discern any link to the wage rate or income levels, although the VTT is found to be 30 to 40% larger when the employee's income level is unknown or not reported. If the employer SP values are valid at least in terms of relative values according to staff seniority, and we must recall that the employers were asked to undertake SP exercises related to a specific category of staff, then the implied valuations in the employer survey are not consistent with the CSA.

There are variations in VTT with distance (for car), time and cost, and the net effect will be for the VTT to be higher for longer distance trips. Again, such a relationship is not consistent with the CSA except insofar as longer distance trips are made by more senior staff.

We will discuss the above inferences in more detail in **Chapter 6** to follow.

6 Reconciliation of Business Values

Much of the background thinking to the recommended approach to valuing business travel time savings was undertaken in the scoping study which preceded the present study[60], and what is reported here is essentially the implementation of that thinking. However, relative to the analysis of non-work trip purposes where people are making their own travel decisions involving their own time and money, business trip-making is inherently more complex and models require a greater degree of interpretation and judgement which we aim to provide here. At the same time, it should be acknowledged that, from the standpoint of microeconomics, the methods employed in the present study – for both business and non-work – are faithful to conventional theory laid out in **Appendix A**.

Whilst interested readers may wish to refer to the detailed discussion in the report of the scoping study, it is perhaps useful to briefly summarise the conclusions of that study, before reporting the new analysis which has been undertaken in the course of the present study.

In the scoping study, we expressed reservations regarding the Cost Saving Approach (CSA) currently employed by the Department, essentially reiterating long-standing and well-rehearsed concerns that not all travel time is unproductive and not all time savings would be converted into productive use to the benefit of the company. In particular, the digital revolution has increased the potential for using travel time productively, and indeed can be expected to have increased the productivity of any such time spent working while travelling. Other arguments against the CSA surround difficulties in estimating the value of the marginal productivity of labour (which underpins the CSA), the benefits of spending more time at the destination (say with a client or at a sales pitch), and the benefits of avoiding overnight accommodation and travel in unsocial hours. By contrast, these effects should in principle work themselves through into a WTP-based valuation, thereby eliciting a reliable representation of what the company would pay.

We felt that an intuitively appealing approach would be to survey employers about how much they would be prepared to pay to reduce their employees' travel time. After all, if the CSA is a valid representation of the value of business travel time savings then the employer should simply express a WTP in line with the CSA. Nonetheless, the difficulties of, and uncertainties surrounding, a valuation approach based on surveying employers were recognised. For example, the data collection costs are high, there are challenges involved in identifying the appropriate employer agent, and even then the agent may not be entirely familiar with specific kinds of business trips.

A potentially complementary approach is to undertake employee surveys, using either RP or SP approaches, which are couched within an awareness of company travel policy. These have attractions provided that they are an accurate account of company decision-making, and the survey method should aim to maximise this reliability.

In the scoping study, we expressed a preference for WTP-based approaches, using different methods for corroborative and interpretive reasons. This reflected a

[60] https://www.gov.uk/government/publications/values-of-travel-time-savings-for-business-travellers

proposition that well designed and conducted quantitative research can provide a coherent ‘story’ as to how business travel time savings are valued, or better still, can elicit direct estimates of WTP that lend themselves to comparison against the CSA.

6.1 Overview of reconciliation

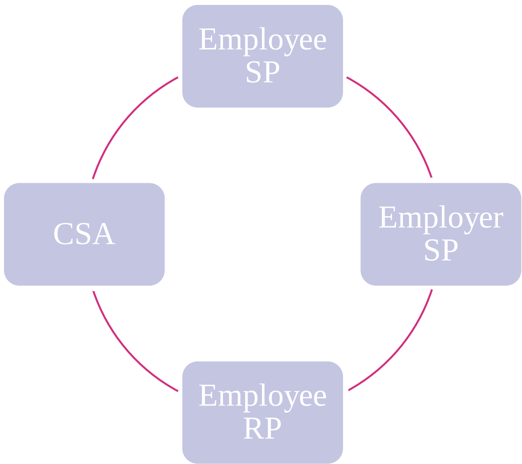
The business travel element of this study has three components based around the estimation of choice models. These are:

- Employer SP
- Employee SP
- Employee RP

The information collected on income and working hours also enables comparison with the Cost Savings Approach (CSA) based upon the wage rage.

Additionally we asked questions relating to the productive use of travel time and the use of saved time which supports the elicitation of the so-called ‘Hensher’ parameters. This evidence is discussed in **Annex H**.

Figure 6.1: Sources of evidence on business VTT



By drawing upon these three components, we are able to identify areas of corroboration between the different approaches, whilst also mitigating the risks of relying on any single approach (**Figure 6.1**).

From a conceptual point of view, it might be argued that employers should be the focus, since it is they who will actually be purchasing the time savings. However, data collection is in this context expensive, and there are significant challenges involved in achieving a representative sample of travel-using employers and identifying relevant decision makers within the firm.

By contrast, obtaining large samples of employees travelling on company[61] business is relatively straightforward, and indeed the business scoping study demonstrated that SP studies along these lines tend to be the norm. The concern here is whether employees are able to make choices in response to hypothetical

[61] Denoting all forms of public and private enterprise for which employee travel is an input.

scenarios that accurately represent the company's willingness-to-pay, or worse still simply represent their own willingness-to-pay. If the employee is to be an acceptable proxy for the employer, then we need employees to respond in accordance with the company's interests as opposed to their own private interests. An interesting special case here is self-employed business travellers, where it might be presumed that company and private interests are one-and-the-same, and that SP responses would therefore reflect what the company would pay.

One might suppose that real-world business travel decisions are sanctioned by companies, and hence if we can observe business travellers paying more money for faster options, then this would provide important insights into company willingness-to-pay. However, identifying contexts that support the estimation of robust RP models is not straightforward, whilst far less data – and generally of lesser information content – is collected than with SP methods. Furthermore, even where preferences are revealed, business travellers may be imperfectly informed as to the availability and nature of time vs. cost trade-offs.

6.2 Summary of approach

Given the concerns raised about the risks involved in obtaining choice data that reflect what the company would do in the light of purchasable time savings, the present study devised a series of questions which would allow interpretation of the robustness and appropriateness of the choice models developed. This was in addition to the more customary investigations of impacts on recovered values of time due to:

- SP design related issues;
- Individuals' socio-economic and trip-related characteristics, here focussing on employers, employees and factors specific to business travel, and;
- The perceived realism and difficulty of the exercises.

The employer and employee questionnaires have been supplied to the Department alongside this report (**Appendix E**).

6.3 Modelling and implementation approach

The present chapter exploits modelling outputs from Chapters 4 and 5. From **Chapter 4** we draw upon the employee SP model. From **Chapter 5** we draw upon both the employer SP model and the employee RP model.

As was qualified in the course of the modelling, it is not – strictly speaking – appropriate to draw interpretation from the 'raw' estimates of VTT from the respective models. Rather, the estimated models should be subjected to sample enumeration, so as to elicit an average valuation for the sample at hand. **Chapter 7** of this report outlines the development and application of an 'Implementation Tool', which was used to generate appraisal values for the **Chapter 4** models; this involved sample enumeration across the NTS sample of travellers (whilst introducing other aspects of functionality such as confidence intervals to the resulting estimates).

Since the discussion of business models in the present chapter considers additional models from **Chapter 5** (i.e. employer SP and employee RP), which did not fall within the scope of the Implementation Tool, the analysis proceeds as follows:

- In order to allow comparison – on a like-for-like basis – between the VTT estimates from the employee SP and employer SP models, we subjected both models to a simplified sample enumeration exercise based on their respective behavioural (i.e. SP) datasets. For reasons of representativeness, it would have been preferable to employ the NTS – as opposed to the behavioural – dataset for this exercise. However, the employer SP model was focussed around variables (e.g. on company travel policy and the employee's use of business travel time savings), which are not covered by NTS, and rescaling against NTS was not therefore feasible.
- As **Chapter 5** noted, the RP analysis conducted as part of this study was not as successful as might have been hoped, and the reliability of the resulting VTT estimates is therefore questionable. For this reason, the discussion of business values that follows makes only tentative reference to the 'raw' estimates of VTT from the employee RP model, and the RP model has not been subjected to any form of sample enumeration.

Before comparing the VTTs estimated using the different approaches, we will first draw some interesting comparisons with regard to company policy and other relevant issues such as time spent working while travelling and the use of saved time. As earlier discussion in this chapter has noted, these features give rise to important distinctions between CSA and WTP, and between employer and employee surveys.

6.3.1 Characteristics of employed travellers

This section highlights some key features of the results reported in **Table 6.1** which distinguish between the employer SP, employee SP and employee RP datasets, with further distinctions between car and rail users in the employer data, and between car blue collar, car briefcase, rail and 'other PT' in the employee data.

In terms of decision-making regarding the purchase of time savings, there is a broad degree of consistency across datasets. A common policy was that the employer allowed the employee to make decisions in the company's best interests, with similar proportions across categories, and few respondents reported that the company would make all the decisions leaving them no choice. Noticeably though, employers were somewhat more likely to say that the decision would have to be justified by a line or travel manager. Most were 'very sure' of the decision making process, with very few 'unsure'.

For the employee datasets – both SP and RP – there is broad similarity in terms of the company's position on buying time savings, with most saying the company would pay if the benefits exceeded the costs. Surprisingly though, significant numbers stated that their company would not be interested in buying a time saving for them. The employer responses were different, and this may be because they have responded in terms what would occur had the respondent bought a time saving. However, to the extent that, within reason, the company would pay come what may, the different views expressed here could be a source of different employer and employee valuations.

With the exception of the employee ‘other PT’ sample, where the surveys were focussed around suburban and shorter distance trips (and especially London), the datasets are dominated by longer distance trips and the mean trip times are therefore relatively high.

As expected, the proportion working while travelling by car is low (which gives credence to the CSA), but this proportion is much higher (and similar in magnitude) for employer SP and employee RP rail travellers. These figures translate into similar mean amounts of time spent working for rail in the latter two datasets, and a very small amount of working time for car. Across all datasets, a significant number of business trips are made – at least in part – outside of normal hours.

Remembering the conceptual underpinnings of the CSA discussed earlier, it is instructive to consider the characteristics of the datasets with regards to wages and incomes. As expected, blue collar employees have the lowest wage rates, whilst the high incomes of ‘other PT’ reflect the large number of London-based employees captured within this data. There is broad similarity of incomes between the employer and employee datasets.

6.3.2 Characteristics of self-employed travellers

Whilst **Table 6.1** is focussed on employees, **Table 6.2** reports analogous summary statistics for self-employed respondents from the employee datasets. To reiterate our earlier comment, this segment constitutes an interesting special case since private and company interests are in principle one-and-the-same. As expected, briefcase travellers have somewhat higher incomes than blue collar business travellers. Compared to **Table 6.1**, the largest differences are that self-employed briefcase rail users and blue collar ‘other PT’ users have somewhat higher incomes than their employed counterparts, whilst the reverse is the case for blue collar rail users.

With the exception of ‘other PT’, the self-employed have shorter trip times on average than the employed, with (as might be expected) blue collar shorter than briefcase.

The most prevalent category for covering travel costs is that they are ‘not covered’, with briefcase rail travellers having the greatest ability to charge to clients, and car users being most likely to have covered the costs in the bill.

The vast majority of self-employed business travellers state that they could not charge for travel time. Of those that could, travellers report a mix of undertaking the work on a ‘fixed price’ or ‘time and materials’ basis.

The vast majority reported that they did not have ‘normal working hours’ (although we accept that, given modern working practices, the notion of normal hours is becoming increasingly nebulous). Of those that did, trips were mainly made within normal working hours. Very few made the trip entirely outside normal working hours.

Table 6.1: Key characteristics of the employer and employee

	Employer SP Car	Employer SP Rail	Employee SP Car Briefcase	Employee SP Car Blue Collar	Employee SP Rail	Employee SP 'Other PT'	Employee RP Rail
DECISION MAKING AND BUYING TIME SAVINGS	[a]	[a]	[a]	[a]	[a]	[a]	
Employer allows decision in company's best interests	100 (43%)	67 (36%)	1075 (46%)	630 (46%)	1765 (44%)	350 (42%)	498 (53%)
Justify decision with line/travel manager	102 (43%)	83 (44%)	500 (21%)	235 (17%)	1180 (29%)	205 (24%)	180 (19%)
Employers makes all decisions and no choice	11 (5%)	15 (8%)	165 (7%)	180 (13%)	305 (8%)	25 (3%)	86 (9%)
Employee allowed to do what want and charge to expenses	19 (8%)	21 (11%)	405 (17%)	155 (11%)	445 (11%)	100 (12%)	161 (17%)
Don't know	3 (1%)	0 (0%)	155 (7%)	140 (10%)	200 (5%)	145 (17%)	20 (2%)
n/a			45 (2%)	40 (3%)	115 (3%)	15 (2%)	
HOW SURE OF DECISION MAKING							
Very sure	n/a	n/a	1500 (64%)	780 (57%)	2470 (62%)	470 (56%)	695 (74%)
Quite sure			615 (26.2%)	390 (28%)	1115 (28%)	250 (30%)	219 (23%)
Unsure			185 (7.9%)	170 (12%)	310 (8%)	105 (12%)	31 (3%)
n/a			45 (1.9%)	40 (3%)	115 (3%)	15 (2%)	
COMPANY POSITION ON BUYING TIME SAVINGS							
Employer pay if benefits exceed costs	27 (13%)	24 (17%)	1355 (58%)	710 (51%)	2355 (59%)	470 (56%)	627 (66%)
Employer would pay come what may	160 (76%)[b]	92 (67%)[b]	305 (13%)	175 (13%)	330 (8%)	70 (8.%)	142 (15%)
Employer not interested in buying time savings	23 (11%)	22 (16%)	640 (27%)	455 (33%)	1210 (30%)	285 (34%)	176 (19%)
n/a			45 (2%)	40 (3%)	115 (3%)	15 (2%)	
TRIP TIME							
Mean trip time	104.12 (95.3)	121.04 (64.1)	159.37	158.54	123.6	36.5	75.59 (24.2)
Less than 30 minutes	32 (15%)	5 (4%)	130 (6%)	105 (8%)	165 (4%)	420 (50%)	-
30-59 minutes	48 (23%)	21 (15%)	300 (13%)	220 (16%)	435 (11%)	290 (35%)	345 (37%)
60-89 minutes	37 (18%)	11 (8%)	300 (13%)	160 (12%)	555 (14%)	70 (8%)	312 (33%)
90-119 minutes	13 (6%)	19 (14%)	210 (9%)	105 (8%)	870 (22%)	20 (2%)	224 (24%)
120 minutes and over	80 (38%)	82 (59%)	1405 (60%)	790 (57%)	1985 (50%)	40 (5%)	64 (7%)
% WHO WORK WHILE TRAVELLING			[c]	[c]	[c]	[c]	
Outward	16.2 (36.9)	78.9 (40.9)					78.5 (41.1)
Return	15.7 (36.5)	65.2 (47.8)					61.7 (48.6)
% OF TIME SPENT WORKING WHILE TRAVELLING			[c]	[c]	[c]	[c]	
Outward	7.3 (21.6)	44.7 (33.5)					50.3 (35.1)
Return	6.9 (21.4)	33.4 (33.2)					34.9 (35.1)
% SOME OF TRIP OUTSIDE NORMAL HOURS			[c]	[c]	[c]	[c]	
Outward	24.8 (43.2)	56.5 (49.7)					51.3 (50.0)
Return	20.5 (40.4)	55.1 (49.1)					62.0 (48.6)

	Employer SP Car	Employer SP Rail	Employee SP Car Briefcase	Employee SP Car Blue Collar	Employee SP Rail	Employee SP ‘Other PT’	Employee RP Rail
CATEGORIES FOR EMPLOYEES DATA:							
Partly outside normal working hours							
Mainly outside normal working hours							
Entirely outside normal working hours							
Within normal working hours							
No answer							
% TIME OUTSIDE NORMAL WORKING HOURS			c	c	c	c	
Outward	11.6 (26.7)	28.7 (37.8)					31.0 (39.7)
Return	9.5 (25.0)	29.7 (38.6)					43.2 (43.8)
PERSONAL INCOME							
Mean wage rate	20.23 (12.3)	28.18 (16.9)	22.30 (23.93)	14.57 (7.87)	24.93 (16.41)	30.63 (72.90)	26.93 (15.8)
10th, 25th,50th,75th, and 90th Percentile Wage Rate	8,12,16,24,36	12,16,22,35,50	9,13,19,27,37	7,9,13,19,24	11,15,21,31,44	8,13,20,30,50	13,17,23,33
<£10k	1 (1%)	1 (1%)	65 (3%)	65 (5%)	95 (2%)	30 (4%)	,46
£10-20k	14 (7%)	6 (4%)	205 (9%)	320 (23%)	295 (7%)	115 (14%)	19 (2%)
£20-30k	24 (11%)	11 (8%)	425 (18%)	430 (31%)	705 (18%)	120 (14%)	54(6%)
£30-40k	19 (9%)	13 (9%)	525 (22%)	275 (20%)	835 (21%)	180 (21%)	141 (15%)
£40-50k	14 (7%)	16 (12%)	375 (16%)	140 (10%)	660 (17%)	90 (11%)	176 (19%)
£50-75k	10 (5%)	11 (8%)	410 (17%)	110 (8%)	695 (17%)	145 (17%)	153 (16%)
£75-100k	8 (4%)	9 (7%)	185 (8%)	10 (1%)	355 (9%)	65 (8%)	193 (20%)
>£100k	2 (1%)	9 (7%)	105 (5%)	0 (0%)	250 (6%)	70 (8%)	107 (11%)
Don’t know/Refused	118 (56%)	62 (45%)	50 (2%)	30 (2%)	120 (3%)	25 (3%)	86 (9%)
							16 (2%)

Notes: Figures in () brackets are standard deviations unless otherwise denoted as %. RP is for employed only. a. Multiple responses were allowed but the %s relate to the total. b. We suspect here that when asked whether the company would pay come what may that this has been interpreted as they would pay an expense claim. c. There were mistakes in asking this question and hence the data is not reliable.

Table 6.2: Key characteristics of the self-employed

	Car		Rail		'Other PT'	
	Briefcase	**Blue Collar**	**Briefcase**	**Blue Collar**	**Briefcase**	**Blue Collar**
Wage Rate	22.26 (19.14) [111]	12.16 (13.94) [61]	31.03 (51.67) [116]	10.49 (6.13) [27]	30.49 (20.22) [31]	21.51 (35.31) [7]
Trip Time	133.97 (110.95) [111]	118.55 (119.69) [61]	101.66 (70.40) [116]	93.52 (62.50) [27]	47.74 (42.55) [31]	44.29 (33.92) [7]
HOW TRAVEL COSTS COVERED?						
Directly charged/reimbursed by client	19 (17%)	11 (18%)	39 (34%)	5 (19%)	7 (23%)	-
Directly charged/reimbursed by customer	3 (3%)	3 (5%)	6 (5%)	1 (4%)	2 (7%)	1 (14%)
Included in call-out charge	6 (5%)	5 (8%)	-	-	1 (3%)	-
Included in bill	36 (32%)	16 (26%)	16 (14%)	3 (11%)	4 (13%)	2 (29%)
Not covered	44 (40%)	25 (41%)	52 (45%)	17 (63%)	17 (55%)	4 (57%)
Missing	3 (3%)	1 (2%)	3 (3%)	1 (4%)	-	-
COULD YOU CHARGE FOR TRAVEL TIME?						
Yes	31 (28%)	15 (25%)	17 (15%)	2 (7%)	10 (32%)	-
No	77 (69%)	45 (74%)	96 (83%)	24 (89%)	21 (68%)	7 (100%)
Missing	3 (3%)	1 (2%)	3 (3%)	1 (4%)	-	-
CHARGING BASIS[a]						
Fixed Price	14 (45%)	10 (67%)	11 (65%)	1 (50%)	5 (50%)	-
Time and Materials	17 (55%)	5 (33%)	6 (35%)	1 (50%)	5(50%)	-
TRIP MADE OUTSIDE NORMAL HOURS[b]						
Yes, partly outside office hours	13 (39%)	2 (13%)	10 (30%)	1 (33%)	-	1 (50%)
Yes, mainly outside office hours	1 (4%)	-	2 (6%)	-	-	1 (50%)
Yes, entirely outside office hours	3 (9%)	-	4 (12%)	-	-	-
No	16 (48%)	13 (87%)	17 (52%)	2 (67%)	9 (100%)	-

Note: **a** asked only of those who said they could charge for travel time. **b** asked only of those who reported normal working hours. Figures in () brackets are standard deviations unless otherwise denoted as %. Figures in [] brackets are numbers of observations.

6.4 Employers' business

In this and the following two sections, we will summarise and compare estimates of VTT from the employer SP, employee SP and employee RP surveys. The present section begins with the employer SP. As noted above, the values reported here were generated by a sample enumeration process on the behavioural dataset.

6.4.1 Choice context

Company representatives were recruited who were in a position to respond authoritatively to issues surrounding company travel and the benefits to the company of time savings. Note that the survey was restricted to 'white collar' business travellers, often termed 'briefcase' travellers. These were defined to be *"trips made by office-based staff travelling to attend meetings, attend conferences and similar business activities but not to provide trade services"*.

This was because it was felt that the Cost Savings Approach (CSA) might be appropriate for 'blue-collar' workers, on the grounds that for these workers travel is generally a means to the end of undertaking productive work at the destination. Hence it would not make sense to dilute the relatively small target sample (of 400 individuals, from **Table 3.3**) with a mix of the two. This turned out to be a sensible decision, since the consequences of limited sample size are very apparent even though the focus is entirely upon briefcase travellers. However, we should be mindful of the implication that if, in due course, the view is taken that the employer values represent the 'truth' to which other values should be rescaled, then the data collected here will inform such rescaling only for the briefcase segment.

As outlined in the previous chapter, the employer representative was presented with three SP exercises, depending upon whether they reported a car or rail trip by a relevant employee:

- Choices between time and cost only (SP1).
- Choices involving usual time, cost and a distribution of travel times (SP2).
- Choices involving travel time in various degrees of congestion and cost for car and travel time in various degrees of crowding and cost for rail (SP3).

Prior to the SP exercise, the company representative was asked a series of questions that would help in interpreting their responses, in terms of how reliably they were able to reflect company willingness-to-pay, as well as more general questions potentially useful in explaining variations in business VTT. The key questions are summarised below.

6.4.2 Segmentation information

Interpretive questions

The questions asked primarily to understand the reliability of the responses provided were:

- Responsibility for decision making relating to company travel, informed by the prior qualitative research (Q1b).

- Familiarity with the reference employee trip (Q12).
- Familiarity with what the employee would be expected to do while travelling (Q25b).
- The SP debrief questions relating to degree of understanding, perceived realism, the ability to make choices that would reflect real-life, and how difficult the choice task was.

Explanatory questions

The questions asked primarily to understand variations in business VTT were:

- Seniority of staff (Q2a).
- A whole series of questions on company policy (Q15s).
- Expected productive use of specific employee's time on outward and return leg (Q11b, Q11c), the efficiency of work undertaken while travelling (Q11d), and proportion of time spent working while travelling in general by that category of staff (Q11j).
- Productive use of time that would have occurred in time saved (Q11e).
- How much of the trip would be spent in employee's own time on each leg (Q11f, Q11g).
- In general, what proportion of business trips are spent in employees' own time (Q11i).
- Whether there is recompense to employees for undertaking business trips in their own time (Q11h).
- Trip distance, duration and cost (Q2, Q3, Q6, Q7, Q10, Q10b, Q11).
- Solus or group travel (Q8b, Q8c).
- Mode and, for rail, class of travel (Q9).
- Whether trip started/ended at home or work (Q3a, Q3c).
- Typical weekly work hours (Q13) and personal income (Q14).
- Type of organisation (Q35) and its size (Q38, Q38b, Q28).

The results for models estimated jointly to the SP1, SP2 and SP3 data for car and rail travellers were presented earlier in **Table 5.10**.

6.4.3 Implied values of time for employers' business SP

The implied VTTs for employers' business based on SP are reported in **Table 6.3** for car and rail. As with the values for non-work reported in **Chapter 4**, these are based on Δt of 10 minutes. As was noted earlier, it is important to stress that these values are sample enumerated on the behavioural dataset, and make no correction for representativeness against NTS.

Given the low VTTs for time spent in free-flow and light traffic, we have to conclude that SP3 for car has not been entirely successful. However, it is not immediately apparent why this is the case, particularly as employers seem able to distinguish between different amounts of crowding on rail, and the least we might expect would be for them to value all aspects of car time similarly and in line with their SP1 responses.

The income category was reported in only 43% of cases for car and 55% of cases for rail. The average wage for those reporting was £20.20 per hour for car and £28.20 for rail data. These imply CSA valuations of £23.90 and £33.40.

The VTTs for car are broadly similar in SP1 and SP2, and fall somewhat short of the CSA. The relationship between the rail values and the CSA are different but not entirely consistent. The SP1 value is somewhat less than the CSA value, to a similar extent as for car, but the SP2 value and the SP3 values for seating in less than crowded conditions provide strong support for the values implied by the CSA, with a premium to be attached to more crowded conditions and standing. Whilst for personal travel we might suspect that strategic bias could influence crowding related valuations, since respondents have an incentive to protest against the contentious issue of crowding, we do not see the same incentive for employers to exaggerate their responses.

Table 6.3: Employers' business SP values of time £/hr

Car		Rail	
VTT_SP1	14.05	VTT_SP1	21.31
VTT_SP2	17.80	VTT_SP2	34.36
VTTR_SP2	3.39	VTTR_SP2_early	-77.25
VTT_FF_SP3	1.78	VTTR_SP2_late	61.84
VTT_LC_SP3	4.43	VTT_SP3_seating1	33.05
VTT_HC_SP3	15.58	VTT_SP3_seating2	32.66
		VTT_SP3_seating3	40.09
		VTT_SP3_seating4	43.00
		VTT_SP3_seating5	49.29
		VTT_SP3_standing1	51.09
		VTT_SP3_standing2	43.14
		VTT_SP3_standing3	61.05
		VTT_SP3_standing4	65.90
		VTT_SP3_standing5	89.35
Wage Rate	20.2		28.2
CSA	23.9		33.4

Despite the significant challenges of conducting SP research on employers, the relatively small sample and the more restricted modelling than for employees, credible estimates of VTT have been obtained. These values indicate that companies seem to be prepared to pay somewhat more to save travel time than employees would themselves, and that their implied valuations are not markedly different to what would be obtained using the CSA, with a premium for crowded conditions.

6.5 Employees' business

6.5.1 Choice context

Employees making business trips were offered the same SP exercises as non-business travellers with one important difference. In the introduction to the SP exercises they were told:

> *"For each of the following pairs of options, carefully compare the two options, bearing in mind the company related decision making process that we asked you about earlier"*

For those who had previously declared in response to Q59Y that their employer 'would not be interested in buying a time saving for me on this journey', the instruction was instead:

> *"You previously said that the company would not be interested in paying to save time, therefore please answer as if you were paying yourself"*

Note therefore that we are not trying to educate respondents or inform them that, say, the company would pay. Nor is there the ambiguity – present in some studies – as to whether the company or individual is paying because there were no clear instructions. Instead we are simply operating with business travellers' perceptions of the constraints of company policy and how sure they are of that policy.

6.5.2 Segmentation information

Interpretive questions

A key issue here is the extent to which respondents understand, and indeed are able to answer in accordance with, what their company would permit, and hence the degree of confidence that can be placed in their SP responses accurately reflecting company valuations. The questions asked primarily to understand the reliability of the responses provided were:

- How business travellers would decide on paying for a possible time saving (Q59e) and the degree of certainty with how that decision would be made (Q59x).
- Perceptions of the company's position on paying for time savings (Q59y) and the degree of certainty regarding how much the company would pay (Q59e).
- Debrief questions relating to whether the SP exercise could be understood and its difficulty, and importantly how sure the respondent was that the choices made would have been allowed under company policy.

Questions regarding company policy not only provided a means of interpreting the results of modelling the SP data, but are important in encouraging proper consideration of company policy when making SP choices.

Explanatory questions

The questions asked primarily to understand variations in business VTT were:

- Trip distance, duration and cost (Q1b, Q6, Q18, Q33, Q37, Q19, Q20, Q23b, Q34, Q38b, Q39, Q40).

- Leg of travel (Q1), nights away (Q11), departure time (Q16) and time spent at destination (Q12).
- Group size (Q13).
- Trip frequency (Q15).
- Degree of crowding on train (Q26, Q27, Q28) and congestion on roads (Q30).
- Preferred departure and arrival time (Q44, Q46, Q47).
- Self-employment or paid employment[62] (Q50), whether blue or white collar (Q69) and status in company (Q71B).
- How travel costs (Q55) and travel time (Q55b, Q55c) were covered for self-employed travellers and whether employees travelled in their own time (Q55e) and the degree of compensation for out-of-hours travel (Q56).
- Age (Q64) and gender (Q65).
- Company size (Q68b), company type (Q68c) and industry area (Q71).
- Personal income (Q73) and typical weekly working hours (Q70).

Model results

The results of the employees' SP models were reported in **Section 4.8**; we do not repeat those results here, but instead focus upon the features of the models pertinent to business travel time savings.

As explained earlier in **Chapter 4**, pooled models were estimated across the three SP exercises. Focussing on the business purpose specifically (recall that the earlier analysis covered the three purposes of business, commute and other non-work), the number of achieved interviews of car, rail and 'other PT' travellers following cleaning[63] were 917, 945 and 206 respectively, yielding across the three SP exercises large datasets of 13755, 13390 and 2515 observations for estimating the business VTT. We here summarise the key features of the models.

Referring back to **Section 4.8**, the rail and car models included elasticities to time, cost and distance, and give rise to higher valuations for longer trips. The 'other PT' model included only a time elasticity, but this reflects the large number of intra-London trips sampled that generally involve shorter distances.

The income elasticities were all in the range 0.3 to 0.4, contrasting with an income elasticity of one under the CSA. The other covariates varied across modes and only a few significant effects were detected. For car, trips within London had a 75% higher VTT, all else equal, with almost a doubling for within London rail trips. Other effects for rail were a 30% lower VTT for one-way trips and 22% lower for trips which involve two or more nights away. For 'other PT', multipliers of 0.2 (i.e. the VTT was 20% that of the 'base' category) and 0.35 were obtained for trips involving two or more nights away and where there were two or more passengers whilst, in line with conventional wisdom, the multiplier for an access trip was around 2.

[62] Some might deem this of use in interpreting the data rather than just as an explanatory covariate.

[63] The cleaning was the 'basic cleaning' and additionally some who stated that they were neither self-employed nor in paid employment.

As for the variables that assist us in interpreting the nature of the SP responses and models, the one that was most often significant concerned the company's position on paying for a time saving (Q59Y). Note that there were no significant variations in the VTT according to company decision making policy (Q59b), how sure respondents were of that company policy (Q59x), and the SP debrief question enquiring how sure the respondent was that the expressed choices would be allowed under company policy.

Adopting the position where the company 'would pay if the benefits exceeded the costs' as the base, the VTT was 30% higher for car users who stated that the company 'would pay come what may', and 56% lower for those who stated that their company 'would not be interesting in buying a time saving' (in this case, respondents were instructed to assume they would pay themselves). Encouragingly, the latter figure is broadly in line with the difference between the VTTs for work and non-work trips that we have estimated in the course of the present study.

For rail, the corresponding figures were 47% higher and 65% lower, with additionally a category representing those who did not reply having values 44% lower. For 'other PT', the VTT was found to be 45% lower where it was felt that the company would not pay.

These figures all relate to the employed. For the self-employed, car users who stated that their travel costs would not be covered (Q55) had VTTs 44% lower, with a 51% reduction in the case of 'other PT'. We would expect these travellers to be more sensitive to purchasing time savings, all else equal, and hence to have lower VTTs.

Car users who were self-employed had a multiplier of 0.68 (i.e. reflecting a VTT 68% that of employed travellers, all else equal), whilst self-employed briefcase and blue collar rail users had multipliers of 0.55 and 0.36 respectively. It may be that these travellers had lower values on account of the time saved not being converted to income generation, or that a time saving would actually reduce their income in cases where they were being paid to travel.

Finally, it should be recalled that **Chapter 4** tested the influence of a wide range of factors on VTT and, in this regard, time use (i.e. the traveller's ability to do something else whilst travelling, to work or surf the net) was found to have little influence. Whereas the CSA fails to account for time use, an attraction of WTP is that it should in principle reflect the productivity of travel time, given current travelling conditions and opportunities to use that time. Whilst the results show that VTT did not vary with time use, this is not to say that time use is unimportant – the results could have been different if the opportunities to use travel time productively had been significantly different.

6.5.3 Implied values of time for employees' business SP

The values of time for employees – based on the SP sample, and not corrected for NTS – are reported in Table 6.4 for car, rail and 'other PT'. The mean hourly wage rates for car, rail and 'other PT' were £19.80, £25.40 and £30.20, which correspond to current CSA valuations of £23.50, £30.10 and £35.80.

Table 6.4: Employees' business SP values of time £/hr

Car		Rail		'Other PT'	
VTT_SP1	17.85	VTT_SP1	26.36	VTT_SP1	9.55
VTT_SP2	27.18	VTT_SP2	45.18	VTT_SP2	9.05
VTTR_SP2	11.51	VTTR_SP2_early	-69.92	VTTR_SP2_early	-15.05
VTT_FF_SP3	7.51	VTTR_SP2_late	124.81	VTTR_SP2_late	17.65
VTT_LC_SP3	12.09	VTT_SP3_seating1	21.07	VTT_SP3_crowding1	7.16
VTT_HC_SP3	22.42	VTT_SP3_seating2	21.34	VTT_SP3_crowding2	7.16
		VTT_SP3_seating3	27.92	VTT_SP3_crowding3	7.16
		VTT_SP3_seating4	31.66	VTT_SP3_crowding4	8.40
		VTT_SP3_seating5	37.85	VTT_SP3_crowding5	12.74
		VTT_SP3_standing1	35.91		
		VTT_SP3_standing2	38.65		
		VTT_SP3_standing3	43.67		
		VTT_SP3_standing4	44.93		
		VTT_SP3_standing5	56.64		
Wage Rate	19.8		25.4		30.2
CSA	23.5		30.1		35.8

The VTTs for car employees do not tell a consistent story. The split of time between free-flow, light traffic and heavy traffic across car employees was 34%, 40% and 26% respectively. This would imply a weighted valuation of £13.21 from the SP3 values of free-flow, light traffic and heavy traffic, which is not entirely consistent with the SP1 value and a lot less than the SP2 value. Moreover, the ratio of the VTT spent in heavy traffic and in free-flow traffic somewhat exceeds what might be expected. As with employers, some doubts surround the SP3 exercise and, as we shall see, another similarity with the employer results is that there does seem to have been an ability to distinguish clearly between different crowding levels.

Focussing upon the SP1 and SP2 results for car, the results are consistent with those for employers in indicating a premium valuation compared to personal values, and for those valuations to be not greatly different from the CSA.

For rail travellers, as with car, the SP2 value is somewhat larger than the SP1 value. However, the SP1 value of £26.36 is consistent with the VTT from SP3, given that rail conditions rarely involve standing but can involve high load factors. As with car, there is clearly a premium VTT for employees compared to what would be their personal values, and these values are not greatly different to the CSA.

The findings for car and rail contrast with those for 'other PT'. In the latter case, average wages are the highest of all modes, presumably due to surveying business travellers in London, but we observe the lowest VTTs across the three modes by some considerable margin. Indeed, values are 25% or less of the CSA, and might therefore be considered low even in terms of personal values.

There must be reasons for these low values. As was suggested earlier in **Chapter 2**, it may be the case that for 'short trips about town', expenses are not always claimed, or are covered by pre-purchased travel cards. Moreover, there may be a 'blurring' of purposes here. A trip say to a client might be made as part of the commute to work, whilst what might have started out as a business trip could, on the return leg, merge into personal business (e.g. shopping) and/or the commute home.

6.5.4 Implied employee SP values by employment type

Whilst **Table 6.4** offers useful insight into the headline values by mode, these values mask some quite marked variations across different segments of the business travellers surveyed. Such variations are captured by **Table 6.5**, which separates out the car valuations, according to whether the business traveller was in paid employment or self-employed, and whether they were a blue or white collar worker.

Table 6.5: Employees' business SP car values by employment type £/hr

	Employee Briefcase	Employee Blue Collar	Self Emp Briefcase	Self Emp Blue Collar
VTT_SP1	20.43	17.45	13.03	8.55
VTT_SP2	31.11	26.58	19.85	13.03
VTTR_SP2	13.17	11.25	8.40	5.51
VTT_FF_SP3	8.59	7.34	5.48	3.60
VTT_LC_SP3	13.83	11.82	8.83	5.79
VTT_HC_SP3	25.66	21.93	16.37	10.74
Wage Rate	22.77	14.57	23.48	13.63
CSA	26.98	17.27	27.82	16.15

What is immediately apparent is that the self-employed, and particularly those who are briefcase travellers, have values somewhat less than the CSA. This may be because, as has already been argued, the time saved is not converted into income, but is instead taken as leisure time or indeed implies a loss of paid travel time. Unfortunately, due to a routing error in the web survey, we do not know what would have been done with the time saved, although from **Table 6.2** we observe that the largest category of cost recovery was that travel costs were 'not covered' and most 'could not charge' for travel time.

We find it encouraging that employed blue collar business travellers – who account for a key market segment – have values that support the CSA. In principle we see the CSA as an appropriate valuation for this segment, and find it reassuring that this is borne out empirically.

The remaining category of employed briefcase travellers has VTTs that seem to be lower than the CSA, although comparison is not helped by the variability in the values across SP exercises, with SP2 yielding somewhat higher values than the CSA, but SP1 and particularly SP3 yielding lower values.

Following on from the discussion of car, **Table 6.6** reports segmented VTTs for rail business travellers. Again we see the self-employed and in particular briefcase travellers to have VTTs somewhat lower than those implied currently by the CSA. For paid employees, there is strong support for a VTT in line with the CSA, with a premium for crowded conditions.

Table 6.6: Employees' business SP rail values by employment type £/hr

	Employee Briefcase	Employee Blue Collar	Self Emp Briefcase	Self Emp Blue Collar
VTT_SP1	29.93	19.48	15.28	5.91
VTT_SP2	51.29	33.38	26.19	10.13
VTTR_SP2_early	-79.38	-51.67	-40.54	-15.67
VTTR_SP2_late	141.70	92.23	72.35	27.99
VTT_SP3_seating1	23.92	15.57	12.21	4.72
VTT_SP3_seating2	24.22	15.76	12.37	4.78
VTT_SP3_seating3	31.70	20.64	16.19	6.26
VTT_SP3_seating4	35.95	23.40	18.36	7.10
VTT_SP3_seating5	42.97	27.97	21.94	8.49
VTT_SP3_standing1	40.77	26.54	20.82	8.05
VTT_SP3_standing2	43.88	28.56	22.41	8.67
VTT_SP3_standing3	49.58	32.27	25.32	9.79
VTT_SP3_standing4	51.00	33.20	26.05	10.07
VTT_SP3_standing5	64.30	41.85	32.83	12.70
Wage Rate	25.98	16.12	32.38	10.49
CSA	30.79	19.10	38.37	12.43

Table 6.7, which segments the 'other PT' values by employment category, does not add to what was reported in **Table 6.4**. The values are all implausibly low as a reflection of business travel.

Table 6.7: Employees' business SP 'other PT' values by employment type £/hr

	Employee Briefcase	Employee Blue Collar	Self Emp Briefcase	Self Emp Blue Collar
VTT_SP1	10.33	8.22	8.72	5.89
VTT_SP2	9.79	7.79	8.26	5.59
VTTR_SP2_early	-16.29	-12.95	-13.75	-9.29
VTTR_SP2_late	19.11	15.19	16.12	10.90
VTT_SP3_crowding1	7.75	6.16	6.54	4.42
VTT_SP3_crowding2	7.75	6.16	6.54	4.42
VTT_SP3_crowding3	7.75	6.16	6.54	4.42
VTT_SP3_crowding4	9.10	7.24	7.68	5.19
VTT_SP3_crowding5	13.79	10.97	11.64	7.87
Wage Rate	33.21	13.16	30.49	24.59
CSA	39.35	15.59	36.13	29.14

An important qualification concerning the above valuations is that some respondents felt that their company would be unwilling to pay for a time saving in respect of the trip captured by the SP. In this case, respondents were instructed to answer the SP exercise as if they were paying themselves. Comparing travellers whose company was expected to pay, against travellers whose company was *not* expected to pay, we found that the former segment elicited VTTs around three times larger for car and rail, and around double for 'other PT'.

We might argue that it is irrational for companies to refuse to pay for a reduction in their employees' travel time. We might also argue that the respondent was wrong to conclude this, such that their (personal) valuation which we collected should be discounted. Alternatively, it may be that the company is acting rationally in refusing to pay, on the grounds that the traveller would simply go home or was working whilst travelling.

Whilst it is difficult to know which of the above scenarios best explains the reality, it is useful to examine the impact of removing those who are effectively reporting personal values. In **Table 6.5** and **Table 6.6** above, the ratio of the SP1 value to the CSA values is 0.75 for car employee briefcase travellers and 0.97 for rail employee briefcase travellers. After omitting those who were reporting personal values, these ratios are revised to 0.88 and 1.26 respectively. In other words, this segment of individuals has a marked impact on the headline VTT.

6.6 Revealed Preference

The third and final piece of survey evidence informing our synthesis of business VTTs is that from the employee RP models.

6.6.1 Choice context

As outlined in **Chapter 2**, we identified rail operator choice as the most appropriate context, on the grounds that there are locations where rail travellers are faced with two or more operators providing services, where there are the time-cost trade-offs essential to value of time estimation, and where we can expect there to be a high level of familiarity with the operator choice. These were:

- The services offered by Virgin West Coast, Chiltern Railways and London Midland between Birmingham (New St or Snow Hill/Moor St) and London (Euston or Marylebone).
- The services offered by East Coast and Great Northern between Peterborough and London.
- The services offered by Virgin West Coast and London Midland from Stoke-on-Trent, Stafford and Rugby to London.

Between them, these different routes in principle provide a wide range of time vs. cost trade-offs. For example, Virgin West Coast can be an hour quicker than London Midland from Stoke but with less than half that saving at Rugby, whilst Chiltern Railways can be competitive with Virgin West Coast for trips to some parts of London (although it is to varying degrees slower on the trunk haul).

Indeed, the recruitment screening was based around a question eliciting awareness not only of the other operator, but also of the possibility to save time (money) at the expense of a higher fare (longer travel time).

Prior analysis of ticket sales data indicated that there was genuine competition between the different operators on these three routes.

6.6.2 Segmentation information

Interpretive Questions

The questions asked primarily to understand the reliability of the responses provided were:

- Whether prior to making the trip, the fares of the different operators were compared (Q6B).
- Whether prior to making the trip, the trip times of the different operators were compared (Q6C).
- The degree of certainty with the fares of other operators (Q12B, Q15B, Q18B, Q21B).
- The degree of certainty with the times of the other operators (Q27b, Q27c).
- Who paid the fare (Q24).
- How decisions regarding choices between faster/slower and dearer/cheaper are made (Q67b2) and the degree of certainty with this decision making (Q67x).
- Company policy on paying for time savings (Q67y) and degree of certainty regarding how much the company would pay (Q67e).

Explanatory Questions

Employer representatives were asked a wide range of background questions along the lines of the employee questionnaire. However, it was never the intention to segment by such factors, given that the relatively limited sample size would not support meaningful and robust analysis. Instead, the aim was to obtain an overall valuation for briefcase travellers making inter-urban rail trips to London.

6.6.3 Implied values of time for RP

The estimates from the RP model were reported in **Section 5.3**, and will not be repeated in full here. **Table 6.8** instead summarises the headline VTT from the RP, which was £54.36 per hour for those respondents who were 'very sure' of the times and costs of the competing operator(s). As was reasoned in **Section 5.3**, we believe that this value has been amplified because of confounding with operator preferences, and indeed we can see that this value is considerably higher than both the corresponding wage rate and CSA values. In this case, it should be qualified that this valuation is based on a 'raw' VTT estimate from the behavioural model and – given the limited specification of this model – has not been subject to any form of sample enumeration exercise.

Table 6.8: Employees' business RP values of time £/hr

	Coeff	t ratio
Very Sure	54.36	3.49
Quite Sure	145.47	4.16
Wage Rate	26.93	
CSA	31.91	

6.7 Synthesis

In this section we bring together results previously discussed separately in order to draw out some conclusions. We do this on a modal basis, on the grounds that there is ample evidence that (consistent with current WebTAG guidance) there are differences in business travel valuations by mode.

Our overall view is that the employer and employee SP values are sufficiently close that we can take the briefcase valuations obtained from employees to be reliable. This finding is in contrast to much previous empirical evidence, where SP values from employees can be claimed to be on the low side. A contributory factor in the latter is the ambiguity as to who is paying and company policy, and we believe that our efforts to improve on this here have led to employee values that can be reconciled with what employers are prepared to pay. Unfortunately though, the RP evidence does not provide the same degree of support, although it certainly does not support values of business travel time savings that are less than the wage rate for the types of trips represented in the RP survey.

We also find it reassuring, with regard to the validity of the SP employee-based approach, that the VTTs for employed blue collar workers are broadly in line with the CSA. The business scoping study concluded that there seemed to be little justification in principle for not using the CSA for this segment, and we think that this result can be confirmed by the evidence here. **The caveat should however be issued that the observed correspondence between the WTP and CSA is based on enumeration of the behavioural sample, and it should not be presumed that this will continue to hold once we move to enumeration of the NTS sample (Chapter 7).**

The rail VTTs are collated in **Table 6.10**. Whilst the rail context brings the RP evidence into play, we have already raised concerns concerning the reliability of these values, and they should therefore be treated with caution. Focussing on the SP evidence, we are reassured that the employer and employee values are broadly consistent – although it should be borne in mind that the employer sample was relatively modest in size, and that the employee sample is (statistically speaking) rather more robust. The evidence again reassuringly suggests that the CSA is appropriate for blue collar employees. The VTTs for the self-employed again appear somewhat lower than employees for a given income level.

Table 6.9 brings together the car values, which are based entirely on SP. The car values for briefcase travellers seem overall to be in line with the wage rate, and would be more in line with the CSA if those who have reported personal values were removed. The self-employed have values somewhat less than the CSA, presumably because the time saved is not put to productive use. The employee blue collar workers have a VTT closely linked to the CSA.

The rail VTTs are collated in **Table 6.10**. Whilst the rail context brings the RP evidence into play, we have already raised concerns concerning the reliability of these values, and they should therefore be treated with caution. Focussing on the SP evidence, we are reassured that the employer and employee values are broadly consistent – although it should be borne in mind that the employer sample was relatively modest in size, and that the employee sample is (statistically speaking) rather more robust. The evidence again reassuringly suggests that the CSA is appropriate for blue collar employees. The VTTs for the self-employed again appear somewhat lower than employees for a given income level.

Table 6.9: Car values £/hr

	Employer SP Briefcase	Employee SP Briefcase	Employee SP Blue Collar	Self Emp SP Briefcase	Self Emp SP Blue Collar
VTT_SP1	14.05	20.43	17.45	13.03	8.55
VTT_SP2	17.80	31.11	26.58	19.85	13.03
VTTR_SP2	3.39	13.17	11.25	8.40	5.51
VTT_FF_SP3	1.78	8.59	7.34	5.48	3.60
VTT_LC_SP3	4.43	13.83	11.82	8.83	5.79
VTT_HC_SP3	15.58	25.66	21.93	16.37	10.74
Wage Rate	20.20	22.77	14.57	23.48	13.63
CSA	23.94	26.98	17.27	27.82	16.15

Table 6.10: Rail values £/hr

	Employer SP Briefcase	Employee SP Briefcase	Employee SP Blue Collar	Self Emp SP Briefcase	Self Emp SP Blue Collar	Employee RP
VTT_SP1	21.31	29.93	19.48	15.28	5.91	54.36
VTT_SP2	34.36	51.29	33.38	26.19	10.13	-
VTTR_SP2_early	-77.25	-79.38	-51.67	-40.54	-15.67	-
VTTR_SP2_late	61.84	141.70	92.23	72.35	27.99	-
VTT_SP3_seating1	33.05	23.92	15.57	12.21	4.72	-
VTT_SP3_seating2	32.66	24.22	15.76	12.37	4.78	-
VTT_SP3_seating3	40.09	31.70	20.64	16.19	6.26	-
VTT_SP3_seating4	43.00	35.95	23.40	18.36	7.10	-
VTT_SP3_seating5	49.29	42.97	27.97	21.94	8.49	-
VTT_SP3_standing 1	51.09	40.77	26.54	20.82	8.05	-
VTT_SP3_standing 2	43.14	43.88	28.56	22.41	8.67	-
VTT_SP3_standing 3	61.05	49.58	32.27	25.32	9.79	-
VTT_SP3_standing 4	65.90	51.00	33.20	26.05	10.07	-

	Employer SP Briefcase	Employee SP Briefcase	Employee SP Blue Collar	Self Emp SP Briefcase	Self Emp SP Blue Collar	Employee RP
VTT_SP3_standing 5	89.35	64.30	41.85	32.83	12.70	-
Wage Rate	27.90	25.98	16.12	32.38	10.49	-
CSA	33.06	30.79	19.10	38.37	12.43	-

The 'other PT' values are reported in **Table 6.11**, although since they relate entirely to employees they have been discussed in **Section 6.6** above. Here the values are all somewhat lower than the CSA and very much in line with personal values. We cannot be certain why this is. We cannot simply conclude that the SP exercises have not worked here when they have clearly worked elsewhere with respondents who in many respects will be very similar. We speculate that there are background issues here that we cannot identify, such as blending of trips for different purposes. Similarly, we speculate that there may be issues surrounding the claiming of 'minor' travel expenses, such that trips are essentially being covered out of personal income or through travel cards. It is interesting to note that these inferences are supported by earlier findings reported in **Section 2.5.4**.

We would therefore not wish to conclude that the value of time savings for 'other PT' is somewhat lower than the CSA without detailed convincing evidence why this should be so.

Table 6.11: 'Other PT' values £/hr

	Employee SP Briefcase	Employee SP Blue Collar	Self Emp SP Briefcase	Self Emp SP Blue Collar
VTT_SP1	10.33	8.22	8.72	5.89
VTT_SP2	9.79	7.79	8.26	5.59
VTTR_SP2_early	-16.29	-12.95	-13.75	-9.29
VTTR_SP2_late	19.11	15.19	16.12	10.90
VTT_SP3_crowding1	7.75	6.16	6.54	4.42
VTT_SP3_crowding2	7.75	6.16	6.54	4.42
VTT_SP3_crowding3	7.75	6.16	6.54	4.42
VTT_SP3_crowding4	9.10	7.24	7.68	5.19
VTT_SP3_crowding5	13.79	10.97	11.64	7.87
Wage Rate	33.21	13.16	30.49	24.59
CSA	39.35	15.59	36.13	29.14

6.8 Recommendations on business VTT

The CSA is, across the world, the dominant approach to valuing business travel time savings where the appraisal of time savings is practised. Indeed, the Department has been one of its longest standing customers.

The business scoping study raised significant doubts as to whether it provides an appropriate basis for valuing business travel time savings for briefcase travellers, doubts that were based on a considerable amount of evidence, both theoretical and empirical. Nonetheless, we expressed the view that empirical research could be used to test those concerns.

We should firstly point out that our view is that the appropriate method for valuing business travel time savings is one that can be identified by suitable research, using both qualitative and quantitative research methods, and that in this study we have largely conducted such research within the constraints of time and money facing us.

We should secondly point out that, as is apparent from the previous comments, we have no predisposition to a particular approach, be it the CSA, the Hensher approach, employer-based WTP or employee-based WTP.

Thirdly, we should point out that there is no reason why values obtained by any of the methods used here should bear any resemblance to the 'wage rate plus' valuations of the dominant CSA. We can understand that studies might sometimes have 'target' values that it would be desirable to recover, such as VTTs not far from WebTAG recommendations, or values of walk and wait time not greatly different to twice in-vehicle time. Instead, our approach has been for the modellers to provide what they regard to be the best explanation of the data. Guidance has been provided to them on the key issues to explore, but apart from this direction, the results we have reported here are essentially independent of the conventions of the business travel market. In other words, the modelling of the business travel data has essentially followed along similar lines to the modelling of the non-work data.

Whilst the values presented in this chapter have **not** been subject to adjustment for representativeness against NTS, we have sought to conduct an objective empirical comparison of behavioural estimates of business VTT from the various approaches. With this qualification, we conclude that behavioural values based on employees' SP responses would seem to provide a suitable basis for taking forward to appraisal. Our reasons for concluding this are as follows:

- Unless there is strong evidence to support the CSA, we feel that it is sensible to move towards a (reliable) WTP approach given previous concerns with the CSA. After all, the CSA can be seen as a special case of the WTP approach, and WTP responses that are a reliable reflection of company policy should (productivity issues aside) lend support to the CSA. Indeed, as with non-work trips, a great advantage of the WTP approach is that we do not need to formulate exactly how the company arrives at its valuation of travel time savings for its employees (whether CSA, Hensher or some other basis), provided the expressed WTP from whatever source is an accurate guide to what the company would pay.
- Although we have reservations about the reliability of the RP valuations, the employee SP valuations seem to be generally credible. Our view is that in large part this is due to the efforts we made both to make people think about company policy as they perceived it and to enquire about that policy. It seems clear to us that these are not in general personal values, and indeed we have isolated such values in the dataset.

- Whilst we have recognised that the employer SP models have their shortcomings, with the limited sample sizes, issues of unfamiliarity, and heterogeneous employers as contributory factors, the employer values are broadly supportive of the employee values. We feel that such correspondence does not occur by chance, but instead indicates that the WTP valuations are genuinely reflecting what companies would be prepared to pay. Although we would have been prepared to consider rescaling employee valuations if they were somewhat out of line with employer values, this would not have been without controversy given the challenges of obtaining employer values and the limited samples sizes involved. It is therefore encouraging that – using a common approach of enumerating the behavioural sample – the SP values obtained from employers and employees are broadly in line.
- Whilst some elements of the WTP evidence elicit values which are in line with the CSA, this does not mean that the CSA is itself supported as a basis for time valuation. We should point out that the relationships implied by the CSA, such as the absence of distance effects, an income elasticity of one and the same valuations for different types of time, are not supported in our results. Also, it should be remembered that the observed correspondence between WTP and CSA is based on the behavioural sample and may not translate to the NTS sample.
- The RP evidence, although not without its problems, would if anything suggest a premium on the CSA, and is not markedly different to the employee SP evidence. However, it should be qualified that the RP valuations presented in this chapter are 'raw' model estimates and have not been subject to any form of sample enumeration. These inferences should therefore be treated with caution.
- The employee SP values provide more detail than the other approaches and to the extent they are plausible should be allowed to influence official guidance.
- Based on enumeration of the behavioural sample, the employee SP values obtained for blue collar are broadly in line with the CSA, which is reassuring. From a theoretical perspective, and as rehearsed in the business scoping study, this is what we would expect. However, it remains to be seen how this finding translates to the NTS sample.
- We find the correspondence of the employee blue collar worker valuations to the CSA to provide further evidence that the SP approach applied to employees is robust. Again, this finding is based on the behavioural sample.
- The self-employed values are lower than for employees. Whilst we cannot demonstrate this empirically, we find this to be plausible if the time saved is taken as leisure. We suggest that the self-employed acting in a rational manner is an appropriate segment for implementation of the Hensher equation.
- There would seem to be a premium valuation for crowding on rail although the evidence regarding congestion was not convincing. Moving to a WTP approach would require decisions on the appropriate valuations to place upon other elements of travel time.

To reiterate, this chapter has reconciled evidence on **behavioural** estimates of business VTT from different approaches, and recommended that employee-based SP should be adopted as the definitive source of evidence. The following chapter (and specifically **Section 7.6.6**) will now examine the implications for **appraisal** values of business VTT.

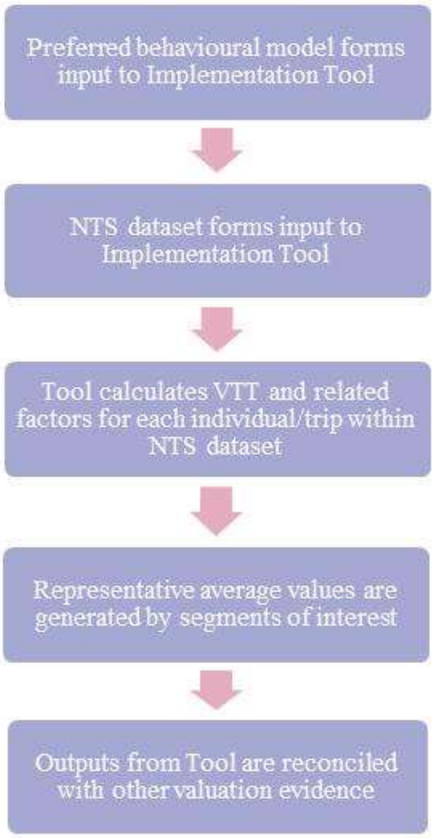
7 Generating National Averages

7.1 Overview of the conversion from modelled values to appraisal values

In the ITT for this study, the Department stipulated that: *"The collected data and evidence must be able to produce robust representative, national average values and have sufficient coverage of different segments and trip characteristics for analysis of key factors that affect the values, which can be used for segmentation"*. The ITT further stipulated that: *"The requirement also includes sensitivity tests around the central recommended values to represent the uncertainty around them. Recommendations for sensitivity tests may be informed by the wider evidence base on values of travel time savings, as well as the level of statistical uncertainty (confidence intervals) around the recommended average and segmented values. The recommended sensitivity tests should be practically implementable in transport appraisal and also consider how the values increase over time with income and the impact this has on uncertainty (e.g. through the estimated confidence intervals)"*.

Against this background, the present chapter sets out the process for the translation of modelled values of travel time savings and related factors from the 'core' modelling of **Chapter 4** into appraisal values suitable for inclusion in WebTAG and other industry guidance. The process for undertaking this translation is summarised in **Figure 7.1** below, and described more fully in **Section 7.2** to follow.

Figure 7.1: Implementation process



7.2 Calculation of appraisal values

The models described in **Chapter 4**, which supply the formulae for the VTT as a function of covariates, are estimated from a sample which is not nationally representative. Whilst we do not expect this to give rise to any bias in the coefficients, the models cannot, without further information, provide appropriate values for selected aggregations of the travelling population, as would be required for establishing recommended values for appraisal.

In the 2003 study[64], the only covariates were distance[65] and income, while separate models were derived for the commuting and other non-work purposes, with business travel having been excluded from that study. This meant that representative values could be calculated by providing a representative matrix of trips for each 'cell' representing a combination of distance and income, applying the formula to each cell, and calculating a weighted average. In the current study, the scope of the model is much wider in that a) it contains many more covariates and b) valuations are provided for a number of quantities in addition to travel 'time', and a matrix-based approach would, therefore, be unwieldy.

While the principles are essentially the same, it is much more convenient to make use of a 'sample enumeration' approach. In this context, this involves the calculation of appropriate valuations (of time, etc.) for each observation in the sample, making use of the relevant covariates, and then calculating weighted averages over the sample to ensure national representativity. We can represent this mathematically as follows:

$$\overline{VTT} = \frac{\sum_n w_n VTT(z_n)}{\sum_n w_n} \qquad (7.1)$$

where $\overline{VTT}$ is the weighted average, n represents an observation in a sample with w_n the necessary weighting to obtain representativity, $VTT(.)$ is the time value formula derived from the model as a function of a vector of covariates $\mathbf{z}$, and $\mathbf{z}_n$ is the set of covariates relating to observation n.

As the best way of ensuring national representativity, it was agreed with the Department that we would use the NTS sample of trips made by persons over 16 by all motorised modes collected during the years 2010-12. It was judged that a 3 year period was appropriate for giving a representative picture of current (as opposed to historical) travel behaviour, and 2012 was the most recent year at our disposal[66]. At the time of undertaking this work, a further update of NTS to 2013 was pending, but it was not anticipated that this update would introduce significant differences. While NTS is – in common with the RP and SP data – a sample, it contains a set of weights aimed at achieving a representative picture of national travel.

---

64 Mackie, P.J., Wardman, M.R., Fowkes, A.S., Whelan, G.A., Nellthorp, J., and Bates, J.J. (2003) 'Value of Travel Time Savings in the UK'. Report to Department for Transport. Available at: http://eprints.whiterose.ac.uk/2079/2/Value_of_travel_time_savings_in_the_UK_protected.pdf

65 In fact, trip distance was not collected in the 1994 data, so that the variable appearing under this name is a construct derived from the reported cost.

66 The NTS data available in the 'special licence version' has been used.

For each trip in the NTS sample, the recommended behavioural model is used to calculate the appropriate valuations, taking account of the covariates of the NTS record. This calculation makes use of the same 'code' used in the model estimation procedure, to ensure complete compatibility. In addition, the estimated standard errors (etc.) are transferred so that each NTS record also carries information about the statistical reliability of its valuations, obviating the need for a special subsequent step to calculate the confidence intervals associated with the recommended values.

Note that it is possible to restrict the calculation of the quantity to summations over the NTS sample observations with particular characteristics, whether or not these characteristics are within the set of covariates defining the valuation formula. In this way, it is possible to derive separate valuations for e.g. geographical breakdowns or for income bands, as well as mode, purpose etc.

In order to provide maximum flexibility, an 'Implementation Tool' has been constructed, which permits the calculation of valuations for different segments and based on a variety of weighting options. Since it offers the necessary functionality and is available as 'freeware', the Tool is programmed in the 'R' language.

Before describing the Implementation Tool in more detail, we reflect upon two critical issues for deriving mean values: these are the question of how to weight the data, and the treatment of income.

7.3 Distance vs. trip weighting

Our primary interest in this chapter is in appraisal and an accurate representation of total travel time benefits. Through covariate analysis, the behavioural models presented in **Chapter 4** identified the characteristics of the traveller and the trip that affect the VTT.

Invariably, however, demand modelling does not in practice occur at a level of disaggregation consistent with the covariates used in the behavioural models, because of constraints in the data underlying the demand model. This has implications for appraisal. Given the need to make simplifications in the model, there is a requirement to use some form of average VTT. The current WebTAG values are distance-weighted averages. In support of this, it was argued in Section 7.3 of Mackie *et al* (2003)[67] that if individuals j have different values of time v_j and travel time t_j, it would seem sensible to try to ensure that $\overline{VTT}\sum_j t_j = \sum_j t_j VTT$, so that the value of the total time disutility is correct. Since time is reasonably correlated with distance, a distance weighting would be a way of achieving this. Nonetheless, it could also be argued that from an appraisal point of view, the criterion should be the weighted **changes** in time rather than total trip time. These changes in time may also be correlated with distance.

A further, and more telling, justification for distance weighting is that the probability that a trip will experience a travel time change is a function of trip distance. When transport interventions are targeted on specific links in the

[67] Mackie, P.J., Wardman, M.R., Fowkes, A.S., Whelan, G.A., Nellthorp, J., and Bates, J.J. (2003) 'Value of Travel Time Savings in the UK. Report to Department for Transport'. Available at: http://eprints.whiterose.ac.uk/2079/2/Value_of_travel_time_savings_in_the_UK_protected.pdf

network, long trips have more chance of experiencing a travel time improvement than short trips – as they travel over more 'links'.

As discussed in **Chapter 4**, the pure distance effect in the behavioural models estimated in this study varies between models but is in the main low or statistically insignificant (in which case it has been excluded). However, as the reference time and cost do contribute to the variation in VTT, it can be observed that VTT will vary with trip distance – since, on average, longer trips are associated with higher reference times and costs.

Table 7.1 shows the ratio between average distance-weighted VTT and average trip-weighted VTT. In the 2003 data (middle column) we can see that the difference for non-work trips between trip- and distance-weighted VTT is very large – with the distance-weighted VTT for other non-work being 78% higher than the trip-weighted value. For the present study we have similarly large proportions (right hand column).

For business we can see that the ratio in WebTAG is much smaller than the non-work ratio. This is because the business VTT in WebTAG is estimated using the CSA, which is unrelated to distance *per se* (though there will still be some correlation between wage rates and trip length). In contrast, the ratio between distance and trip weighting for business in the current study is as large as the non-work ratios. This arises because the business models estimated here (which are based on the employee SP) are also sensitive to reference time and cost. We can therefore see that the decision to weight by trip distance can make a substantial difference to the average 'appraisal' VTT.

Table 7.1: Ratio between distance- and trip-weighted VTT by trip purpose

Trip purpose	**Ratio between distance: trip VTT**	
	2003 study/ webTAG	**2014 study**[68]
Commute (all modes)	1.52	1.33
Other non-work (all modes)	1.78	1.98
Business – car driver	1.19	1.83
Business – rail passenger	1.04	1.61

Source for 2003 data: Own analysis of data supplied by DfT as part of the VTT peer review (Laird, Bates and Mackie, 2013)[69].

Quite apart from the question of how to carry out the weighting to derive an appropriate national average, there is the issue as to whether it would be appropriate to segment the values by distance within the appraisal. The prevalence of network wide models has increased markedly over the last two decades, with area-wide models of most city regions now in existence. Within these models the origins/destinations of most trips are known (aside from/to the external zones) and, it follows, the trip distance profile is also known. Hence, it is conceptually

[68] The values for the 2014 study have been obtained using the Implementation Tool which is described later in this chapter.

[69] Laird, J., Bates, J. and Mackie, P. (2013) 'Peer review of proposals for updated values of travel time savings'. Report to DfT.

possible to use trip-specific VTT in an appraisal. However, the modelling and appraisal effort in undertaking this would be large and it is important to understand what added value such an effort would bring.

As part of our preliminary investigation into this issue, we undertook some synthetic analysis using the income model developed by Zhang and Laird (2014)[70] and the overall NTS distance profile. This analysis was conducted using the 2003 VTT model for other non-work trips, focussing upon a 5 minute time saving for all trips affected by the scheme. The likelihood of a trip receiving a time saving is assumed proportional to the trip distance. We calculated the actual time saving benefit under three different distance profiles, as shown in **Table 7.2**.

Table 7.2: Synthetic analysis – Average VTTs for different distance profiles (other non-work, 2003 VTT model, 5 minute time saving for all trips)

Distance profile	Average VTT per trip
National distance profile	**£6.05**
Example short distance profile (all trips < 25 miles)	£4.02
Example long distance profile (all trips > 10 miles)	£6.67

Notes: (1) National distance profile sourced from the NTS. (2) Short distance profile is based on the national distance profile, but truncated at an upper bound of 25 miles. (3) Long distance profile is based on the national distance profile but truncated with a lower bound of 10 miles. (4) Income distribution synthesised following Zhang and Laird (2014). (5) Probability of trip benefitting from the scheme being appraised (5 minute time reduction) is proportional to trip length (that is within any distance profile longer distance trips are more likely to receive the benefit than shorter distance trips).

Using a national distance profile with distance weighting (the first row in **Table 7.2**) gives us a value of £6.05 which is close to the WebTAG value – thereby giving us re-assurance that our modelling approximations are realistic. Against this value we can compare the average VTT value for a trip with a 'short distance' profile (row 2 of the table), which is £4.02 – some 30% below the WebTAG value. Similarly we can compare the VTT value for a trip with a 'long distance' profile which is £6.67 – some 10% above the WebTAG value.

We have also calculated that the trip-weighted average VTT from a national trip profile is £3.44. This is less than the VTT from the example short distance profile (row 2) and the almost half of the VTT from example long distance profile (row 3). We can see that the trip-weighted VTT value (£3.44) is a better approximation for the short distance example than the distance-weighted value (£6.05). Conversely the distance-weighted value is a better approximation for schemes with a longer trip distance.

What we draw from this analysis is that both distance-weighted and trip-weighted averages (based on a national distance profile) are approximations – to varying degrees – of the 'real' VTT associated with the trips affected by the scheme in the appraisal. This is primarily due to the different distance profiles associated with different schemes. To some extent, approximations will always exist. We would anticipate that varying VTT by distance band and possibly by geography type (urban, inter-urban) – as geography types will be correlated with distance – would

[70] Zhang, G. and Laird, J.J. (2014) 'Local values of travel time savings – should we forget about them?' The 19th International Conference of Hong Kong Society for Transportation Studies, Hong Kong 13-15 December, 2014.

reduce these discrepancies. This is because the ratio between distance-weighted and trip-weighted VTTs becomes much smaller when VTT values are disaggregated by distance.

Following from this conceptual discussion, **Section 7.6.3** below presents trip-weighted and distance-weighted VTTs for all distances and by distance category, and discusses the issue further.

7.4 Treatment of income

7.4.1 Principles and alternative approaches

A long standing and contentious question is whether the VTT used in appraisal should be a pure willingness-to-pay value or should be adjusted somehow to a standard value. In part this is a decision based on economic theory, but it is also a policy decision based on the practical infeasibility of collecting robust data for every scheme appraisal. With regards to the latter, we have assumed that policy is unchanged: that is, standard values will continue to be required in one form or another. Then there is a second issue; how exactly should variations of VTT with income be allowed for when computing the standard value?

Discussions with the Department identified four possible ways of dealing with this within the present study:

Option (1): Include the observed variations in income by person and trip in the calculation of average VTT. We take this to represent the Department's current approach. The advantage of this method is that it represents the 'true' average across the population. One issue that may arise with the method is that if VTTs vary geographically/regionally then spatial variations in income will influence the regional VTTs unless this is controlled for in other ways; see for example option (4) and the discussion on distributional weights below.

Option (2): When calculating average values across the NTS data, treat all respondents as being at 'average' income. To some extent, this method avoids the issue of addressing regional differences in income through some form of distributional analysis. There is though the issue of which 'average' to use: e.g. mean versus median, national vs NTS sample, and also the issue that the values estimated do not represent 'true' averages.

Option (3): Remove the income covariate from the behavioural model, letting the effect be picked up by other covariates. However, this option may introduce model misspecification.

Option (4): Include variations with income as in option (1), but then re-weight according to the process set out in the Green Book. The attraction of this approach is that it is most closely aligned with both theory on distributional weights and the overarching guidance of the Green Book. It is however a challenging undertaking.

In progressing our own analysis, we implemented options (1), (2) and (4) within the Implementation Tool. We do not consider option (3) to be an appropriate method: effects identified in the model should be explicitly dealt with in implementation (unless they can be rejected on other grounds), otherwise their exclusion will bias other coefficients in the model. That is to say, excluding income from the model would be a misspecification. Option (4) has some attraction, in terms of following the general principles of the Green Book, but requires careful treatment as we discuss further below.

The Treasury Green Book approach can be seen as the ideal, in which behavioural values are carried through directly, and a set of distributional weights are used to re-weight the resulting benefits/costs[71]. The arguments for and against this position were considered in the 2003 study (Mackie *et al*, 2003 pp79-83[67]). In principle, Mackie *et al* were in agreement with the Green Book approach, however, they identified a number of practical problems with implementation:

(i) Obtaining the relevant data on the pattern of usage by income and social group;

(ii) Defining the final incidence of costs and benefits;

(iii) Treating the non-monetised elements in the appraisal consistently with the monetised elements; and

(iv) Agreeing a set of weights.

Balancing principle and practice they proposed three types of appraisal, with the values used varying by type:

Level 1: Routine appraisal of small and medium sized schemes: standard values for commuting and other non-work are appropriate.

Level 2: Major schemes and strategies: the use of VTT by three income levels benchmarked against a standard VTT case is suggested, backed up by an analysis of the distributive implications of the scheme.

Level 3: Special cases (toll roads, user charging schemes, metros and other 'user pays' interventions): it is proposed that VTTs should be taken from specific market research exercises benchmarked against standard values and with appropriate quality controls.

For the main part, WebTAG presents standard values of time by mode (for business) and by trip purpose for non-work. This constitutes values for the Level 1 type analysis. However, in the modelling part of WebTAG (Databook M2.1 and M2.2), VTTs for three different income bands are also presented (<£20,000, £20,000-£40,000 and >£40,000). The guidance states that they are primarily for modelling purposes (e.g. for toll roads), but they can be used in appraisal for comparison with results based on standard values. It is our understanding that values varying by income are very rarely, if ever, used in appraisal – in part this will be due to problems with identifying the incomes of the beneficiaries, but problems also arise in reconciling the different outcomes of the standard and the income-varying approaches. Thus effectively only Level 1 appraisals have been undertaken since the 2003 study.

In our view, the modelling and appraisal landscape has changed very little since 2003 in relation to the treatment of income. It is true that the Treasury has now

[71] The Green Book approach to deriving the weights uses the concept of an underlying social welfare function that links personal utility (or satisfaction) to changes in income. Alternative approaches can have a more social element that would include the value society places on say a more equitable distribution of income.

defined a set of weights that can be used in appraisal, and it is also true that modelling methods that may help to identify the final beneficiaries, such as land use-transport models, are more prevalent than they were in 2003. Thus some progress has been made on two of the four practical challenges identified by Mackie *et al*.

From an appraisal perspective, however, we remain a long way from attempting a full distributional appraisal based on the output from, say, a land use-transport model. There has been almost no research effort in this direction over the last decade or more. We also do not know under what conditions a more simplified distributive analysis – that of adjusting VTT by Green Book weights (i.e. option (4) above) – would be a suitable proxy for a full distributive analysis based on final beneficiaries. Additionally, there remains the data gap on travellers' incomes. Such income data are rarely collected in roadside interview surveys and are clearly not present in rail or bus ticket sales data[72] – both of which are the current bedrock for demand modelling and, therefore, appraisal.

The lack of progress in advancing this element of appraisal over the last decade contrasts with increasing interest in using VTT more reflective of local characteristics in appraisals. This is most evident in London where TfL uses an uplifted WebTAG VTT to reflect higher earnings in London. Overseas we also find disaggregation by short and long distance (or urban and inter-urban trips) and occasionally by trips to/from the capital and the rest of the country (e.g. France). There is, therefore, an interest at a policy level in disaggregating VTT further than just by trip purpose – for example by mode, by distance, by geography and by time period. Zhang and Laird (2014) also argued that the lack of disaggregation by distance in UK VTT is the main source of bias introduced by the use of standard values disaggregated by trip purpose – as opposed to the lack of variation in VTT by income *per se*. This is because standard values of time contain an implicit weighting which may vary by scheme, creating the potential for inconsistencies between appraisals. If however we allow standard values to vary with distance, then this reduces such bias. This provokes a conceptual argument for segmenting VTT by more than just trip purpose (non-work) or mode (work).

Given this context and the state of progress since 2003, we feel that the main thrust of Mackie *et al*'s recommendations with respect to income is still relevant today. That is to say, there are:

- Three levels of appraisals.
- At Level 1, income can be standardised.
- At Levels 2 and 3, variations in income should be taken into account.

Our next interest is in how far the Level 1 appraisal VTT should be standardised (for any particular trip purpose, mode or geography segment). Drawing reference to earlier discussion, the two principal options are: (1) to average over income but not to segment by income; and (2) to calculate values at 'average' income.

[72] Zhang and Laird (2014) had to synthesise household incomes from zonal deprivation levels in order to estimate VTT of commuting trips.

7.4.2 Non-work

Without pre-judging our final recommendations regarding VTT, we envisage that the Department may wish to adopt either:

- VTTs disaggregated by trip purpose only (this is the current position); or
- VTTs disaggregated by trip purpose and/or mode, distance and geography.

In our view, these different degrees of disaggregation may warrant different treatments of income.

Level 1 appraisal (routine appraisal of small and medium sized schemes)

Our recommendations vary according to the degree of disaggregation the Department chooses to adopt.

- If the Department chooses to disaggregate VTT only by trip purpose (as with current WebTAG) then the option (1) income treatment should be used (basing the average value on observed variations in income by person and trip). This would give a representative VTT without overly compromising the appraisal given the absence of distributional weights.
- If the Department chooses to disaggregate VTT by trip purpose as well as by one or more of mode (e.g. bus, rail, car), distance (e.g. <30kms, >30kms) and geography (e.g. London, other cities, rural) then, given the systematic variation in income by these levels of disaggregation, we are concerned that the omission of distributional weights would overly bias the appraisal. Our preference, therefore, is for income option (2), such that we treat all respondents as being at 'average' income (note that this is the approach adopted in Sweden[73]). The average income used for the calculation should be the average income of travellers – in this case motorised travellers, since the present report is concerned with the VTT of motorised trips.

Level 2 appraisal (major schemes and strategies)

For Level 2 appraisal, income option (1) (using the observed variations in income by person and trip to calculate an average value) should be used irrespective of the level of disaggregation.

A comparison of the different VTT values under the different treatments of income is presented in **Table 7.16** later in this chapter.

7.4.3 Business

Values of business travel time savings are different from non-work time in one key respect. Whereas VTT for commuting and other non-work purposes can be assumed to accrue to the person making the trip, for business the benefit accrues in the first place to the business enterprise on whose behalf the trip is made, and ultimately to consumers through changes in prices and output via the economic process. The result is changes in real income for households which are distributed in some way across the population. Tracing these changes through is impractical

[73] Eliasson, J. (2013) 'International comparisons of transport appraisal practice: Annex 1 Sweden country report', report prepared for the Department of Transport, Institute for Transport Studies, University of Leeds.

within the context of most scheme appraisals for the reasons set out earlier and, as also discussed earlier, this issue has not been explored by researchers in any depth.

We therefore face an awkward question, namely how should the social value of business travel time savings be determined? In principle, we think the best option is to use the willingness-to-pay (WTP) values presented later in this chapter, at whatever level of spatial or modal disaggregation is required. It will be seen that the results suggest significant differences in business VTT by mode, and we would recommend treating these as real differences associated with variations in the income, journey length and other characteristics of business trips by mode. There are of course practical limitations to the extent that local/regional variations can be incorporated into an appraisal, and our recommendations reflect this.

Level 1 appraisal (routine appraisal of small and medium sized schemes)

For Level 1 appraisal, it would be impractical to collect local data on the incomes of business travellers. Our recommendation is that national averages by mode should be used. For professional drivers, these should be based on the CSA (as now), whilst for all other business trips they should, as we will recommend later, be based on WTP but differentiated by mode. These would use income option (1) (using the observed variations in income by person and trip to calculate an average value) for the calculation of a single national value by mode. This is consistent with existing guidance.

Should the Department wish to disaggregate business VTTs further, for example by distance or geography (London, rural, etc.), then we would again recommend that income option (1) is used.

Level 2 appraisal (major schemes and strategies)

Here again, average values for professional drivers should be based on the CSA (as now). Values for all other business trips should be differentiated by region as well as mode and distance, if statistically reliable in the context for which they are to be used. Income option (1) should again be used.

7.5 Implementation Tool

7.5.1 Overview

The aim of the Implementation Tool is to create a sample enumeration system based on the National Travel Survey (NTS), which can estimate mean values of time, reliability and crowding etc. for selected aggregations of the travelling population, for any chosen model, as well as the confidence intervals associated with those values. The Tool is described in detail in **Appendix I**.

The NTS is an established series of household surveys of personal travel in Great Britain, designed to track long-term development of trends in travel. NTS data is collected via interviews with people in their homes, and a diary that they keep for a week to record their travel. The NTS covers travel by all age groups, including children. In each year, diary data was collected from over 7,700 households, covering over 18,000 individuals.

The necessary NTS data available in the ‘special licence version’ has been obtained for the years 2002-2012, though only the years 2010-2012 have been used for the purposes of ensuring representative results. The required data for implementation are the ‘trip’ records from the diary, with additional data from the household and individual records appended. This dataset has been reduced so that only those variables relevant to the VTT modelling are retained.

The NTS covariates have two roles: to allow for identified variation in VTT (and associated valuations) and to define appropriate aggregations. The VTT for each NTS record is calculated using the same formulae as used to estimate the behavioural model. Where the model involves dependency on covariates not in the NTS file, the default assumption has been that the estimation sample is representative in relation to those covariates: however, the Tool offers the scope to adjust this assumption on the basis of external data.

The output generated for each identified aggregation is as follows:

- The base VTT for SP1 and SP2 (£/hr)
- Value of the standard deviation of VTT (car only) (£/hr)
- Values of lateness (for public transport) (£/hr)
- VTT for different traffic conditions for car (free-flow, light traffic and high congestion) (£/hr)
- VTT for different levels of crowding (for public transport) (£/hr)
- and, for all these quantities, the 95% confidence intervals are reported.

Fares and incomes have been adjusted to 2014 prices and values using CPI for fares and CPI and real income growth for incomes. All values output by the Tool are therefore in 2014 **perceived** prices.

The confidence intervals integrate the model-related error (covariance matrix of the coefficients) with the sampling variation implicit in the NTS, the details of which are provided in **Appendix I**.

According to the NTS User Guide[74]: *“Analysis of travel data is based on the diary sample. This comprises all 'fully co-operating households', defined as households for which the following information is available: a household interview, an individual interview for each household member, a seven day travel diary for each individual and, where applicable, at least one completed vehicle section. Weights were produced to adjust for non-response and, at the trip-level, they were also produced for drop-off in recording observed during the seven day travel week.*

The weighting methodology produces weights at the household, LDJ and trip level. The household weights apply to all individuals and vehicles within the household, and they have therefore been attached to the individual and vehicle files for ease of use. Similarly, the trip level weights apply to all stages within trips and have therefore been attached to the stage-level records for ease of use.”

[74] Department for Transport, National Travel Survey: Data Extract User Guide, 2002-2012.

The appropriate weight to use as the basis for the Tool, which is trip-based, is W5. A key point here is that we account for error in the model estimation and (unlike previous VTT work) for error in the sample.

7.5.2 Behavioural models

Eleven behavioural models developed earlier in **Chapter 4** have been implemented in the Tool. These are:

- Car: commute, other non-work, business.
- Bus: commute, other non-work.
- ‘Other PT’: commute, other non-work, business.
- Rail: commute, other non-work, business.

These models are all based on the SP data. Supported by the arguments presented in **Chapter 6**, the business models are all based on the employee SP survey; we therefore refer to “employees’ business” in the tables reported subsequently.

Whilst a concessionary fare model was estimated in **Chapter 5**, it was considered that the model did not perform satisfactorily enough to act as a basis for the estimation of VTT for concessionary fare travellers. Other behavioural models from **Chapter 4** are therefore applied to concessionary fare travellers.

7.5.3 Size effect options

The modelling analysis in **Chapter 4** indicated the clear presence of ‘design effects’ in the SP data. For the most part, the modelling has been able to neutralise these, but for one particular effect – the ‘size’ effect which relates to the changes in time (and other variables) offered relative to the ‘reference value’ – this has not been possible. This means that, for example, the value of travel time (VTT) is dependent on a particular assumption about the size of the time change Δt. Hence, it is necessary to operate the Tool making specific assumptions about Δt.

7.5.4 Income treatment options

With reference to earlier discussion in **Section 7.4.1**, the Tool implements the three viable approaches for the treatment of income within the calculation of appraisal values of time, namely:

Option (1): Averaging over income, but not segmenting by income

This is the existing option where the effect of income is included when calculating the value for each trip VTT. These are averaged to give single, ‘standard’ values for commuting and other non-work (and other levels of segmentation).

Option (2): Calculating values at ‘average’ income

This approach is similar to option (1) but treats all trips in the weighting process as having ‘average’ income. The user specifies household income (for non-work VTT) and personal income (for business VTT).

Option (4): Applying Green Book distributional weights

Values are again calculated for each trip, using the same parameter values as under option (1). The resulting values are then weighted, according to the income quintile which the income band falls in, using the weights given in the Green Book. This applies to non-work only.

7.5.5 Public transport cost imputation

The formulation of the behavioural models in **Chapter 4** was such that any observations with a zero cost (e.g. concessionary fare travellers) were in effect assigned a zero VTT. One version of the Tool was developed which adhered to this convention. However, since it is highly unlikely that such travellers really hold a zero VTT, the final version of the Tool replaces any zero cost entries with an imputed cost. The imputed cost was based on regression models for each mode/purpose, estimated by using the NTS sample to explain public transport cost as a function of distance, time and geographical area.

7.5.6 Tool validation

Prior to use, the Tool has been validated in various ways, as described in **Appendix I**, and we can therefore be confident regarding the robustness of the translation process from behavioural to appraisal values of travel time savings (and related factors).

7.6 Appraisal values

7.6.1 Introduction

In this section we report the results generated by the Implementation Tool and discuss what recommendations to issue. There is a large amount of detail in what follows: in terms of the different valuations covered by the three SP experiments; variations by modes and purposes; as well as the key options for running the Tool (namely how to deal with variation by Δt, treatment of income, choice of weighting, and cost imputation). Because of the number and combinatorial nature of these issues, we feel it necessary to make initial decisions about some issues, before returning to these later in the section. We begin by summarising the key considerations.

Size effects

It was shown in **Chapter 4** that VTT was generally sensitive to the value of Δt presented in the SP, requiring a multiplier of $\Delta t^{\kappa-1}$ (though this was largely restricted to SP1). For the record, the estimated values of this exponent in the various models were as shown in **Table 7.3**.

Thus the effects are not especially large, but a value of, say, 0.25 for the exponent $(\kappa - 1)$ is still sufficiently high to make a difference of a factor of 1.8 between values for Δt of 2 minutes and 20 minutes. Such size effects have generally been found in the analysis of SP data for VTT estimation, and the question of how to deal with them always arises. It remains unclear to what extent they represent reality, as opposed to being a response induced by the SP exercise.

Table 7.3: Δt parameters for SP1 models

Mode	Purpose	κ–1
Car	Commuting	0.400
	Employees' business	0.240
	Other non-work	0.267
Rail	Commuting	0.214
	Employees' business	0.230
	Other non-work	0.133
Bus	Commuting	0
	Other non-work	0
'Other PT'	Commuting	0.164
	Employees' business	0.397
	Other non-work	0.117

It is certainly plausible that people may consider very small time savings (measured in seconds) as of essentially zero value, but over time even very small time changes may accumulate into something of more use. This was discussed in Mackie *et al* (2003), at the outset of Chapter 4, and in more detail in Section 4.2: the approach taken there was effectively to reduce the weight given to observations with time changes in the range ± 11 minutes around the current trip time, and given the time changes actually offered in the SP designs, the results were dominated by those with gains or losses of 15 or 20 minutes. It should be noted that in the SP designs the majority of the time changes offered were 10 minutes or less.

An overriding consideration is that we should obey basic laws of 'adding up' in scheme appraisal: the incremental value of (A minus B) plus the incremental value of (B minus C) should equal the incremental value of (A minus C). Thus there is a strong appraisal argument for the use of a constant unit value[75] and we would need equally strong counter arguments to move away from this position. Nevertheless, given the dependence of VTT on Δt, there is still the need to take a view on what value to use.

Welch and Williams (1997)[76] noted that for many road schemes the time savings are small (they suggest that anecdotally the average time savings may be in the range of 1-3 minutes), but there are of course other cases (such as high speed rail) where very large savings are possible. The general view is that discounting the value of small time savings of < 5mins can reduce user benefits by up to 50% (see

[75] WebTAG currently uses a constant unit value and requires the distribution of time savings to be illustrated.

[76] Welch, M. and Williams, H.C. (1997) 'The Sensitivity of Transport Investment Benefits to the Evaluation of Small Time Savings'. Journal of Transport Economics and Policy vol. 31, pp231-254.

Welch and Williams for an urban case study and Yakeen and Laird (2014)[77] for an inter-urban case study). Nonetheless, the 'adding up' arguments need to be kept in mind, so that the proportion of time savings of whole investment strategies will be larger than the scheme-by-scheme median time saving.

The Danish study[78] suggested different threshold values of Δt for different modes, though for most modes it was stable between 10-20 min. Based on this they decided that 10 minutes was 'reasonable' for all modes. The Swedish study[79] also found varying thresholds, especially between long and short distances: in the event they decided to use 15 minutes for regional trips and 20 minutes for long distance, though these were essentially arbitrary decisions based on what was felt to be 'reasonable'. In Norway the VTT for short distance modes turned out to be rather stable at Δt =10 min, but not for long distance modes. They used the threshold 10 minutes for short distance travel and 15 minutes for all long distance modes, following the reasoning in the UK, Danish and Swedish studies.

Good practice would suggest that the estimated behavioural models should only be applied within the range of covariate values over which they have been estimated. The range of Δt presented in the SP data tends to be on the low side (particularly for bus and 'other PT', where more than 80% of the values are less than 10 minutes), but for car and rail a reasonable proportion (more than 15%) are above 20 minutes.

Based on all this, we consider that a value of Δt = 10 minutes is defensible for basing appraisal values on. Hence, the values reported in later tables are all based on the assumption that Δt = 10 minutes, though the effect of different Δt on the appraisal values can be calculated using the factors in **Table 7.3**. The relevant variations by mode and purpose with Δt for SP1 are presented later in this section in **Table 7.9**.

The treatment of income

We have discussed the conceptual arguments in some depth in **Section 7.4**. We first present the main results using option (1), which takes account of the actual income (by mode and purpose, etc.) in the travelling population (though of course this is only relevant when the model identifies an effect due to income – we note that this was not the case for bus commuters). This corresponds to the methodology adopted by Mackie *et al* (2003). We then indicate how the results will change as a result of options (2) and (4) – though we do not repeat the discussion from **Section 7.4**. In general, **all** time values within a particular model will change in the same way as a result of these alternatives, so the level of detail can be reduced. For commuting and other non-work, the results are based on household income, while the results for business are based on personal income.

---

[77] Yakeen, F. and J.J. Laird (2014) 'Small time savings, loss aversion and transport cost benefit analysis'. European Transport Conference, Frankfurt, 29th September to 1st October 2014.

[78] Fosgerau, M., Hjorth, K. and Lyk-Jensen, S.V. (2007) 'The Danish Value of Time Study, Results for Experiment 1, Note 5'.

[79] Börjesson, M. and Eliasson, J. (2014) 'Experiences from the Swedish Value of Time study', Transportation Research Part A, 59, pp144-158.

Weighting

Based on the discussion in **Section 7.3**, we will present the main results using distance weighting – this again is in line with existing methodology. We then indicate the impact of using trip-based weighting instead. Again, this impacts in the same way on all time values within a particular model.

Imputed cost

Because the base cost was found to be an important model variable for the calculation of VTT, we need to decide what can be done when the NTS data used in the Tool does not contain the trip cost. For all car journeys, NTS records only the time and distance: we have used WebTAG methodology to impute the fuel costs, assuming a standard vehicle and using the parameters given in WebTAG Unit A1.3[80]. The absence of trip costs for public transport is more problematic. The incidence of zero costs for public transport in the NTS sample is shown in **Table 7.4**.

Table 7.4: Proportion of trips with zero cost in the NTS dataset by mode and purpose

	Commute	Business	Other non-work
Car	-	-	-
Bus	13%	-	54%
'Other PT'	10%	14%	24%
Rail	5%	5%	12%

The majority of these records relate to concessionary travellers, but there are probably other reasons for some of the zero reported costs. Recall that in the model estimation anyone with zero cost was assigned a zero VTT. The only model where zero cost was allowed was for concessionary travellers (**Section 5.2.2**) – but this model is not being implemented here.

Faced with this position, there appeared to be three possible courses of action:

a) Accept the fact that the model would assign a zero VTT for any public transport trip where the fare was zero.

b) Drop all such zero records from the Tool.

c) Impute the costs based on other characteristics of the trip.

Option a) would be incompatible with the model estimation and would involve an unreasonable extrapolation. Option b) runs the danger of operating with a biased sample. Hence, option c) was applied, as described at the end of **Section 7.5**.

Business values

It is important to recognise that in previous SP work examining employees' business values, it has not always been clear exactly what was being valued: business values, personal values or some mixture of both (Wardman *et al*, 2013). The problem stems from the phrasing of the questions posed to the interviewees. As discussed in **Chapter 6,** to avoid this ambiguity, respondents in this study were asked if they thought their company would be willing to pay for a time

[80] DfT (2014) TAG UNIT A1.3. User and Provider Impacts. November 2014.

saving. Those who responded that their company would **not** pay for the time saving were instructed to respond to the SP as though they would be paying for the time saving out of their own personal money. Clearly, for these respondents, we unambiguously have a personal valuation of a travel time saving (though in the course of work). For those who responded that their company would pay for a time saving, we unambiguously have a value for the business time saving (subject to the interviewee being able to respond appropriately). From an appraisal perspective, we require the **business** valuation (that is, the latter valuation). Therefore, in implementing the employees' business model within the Tool, we need to neutralise the effect of those who were unambiguously responding in a personal capacity. **Section 7.6.6** to follow will report estimates of business VTT on this basis.

7.6.2 Detailed results by mode and purpose

In this section we present the results in line with the different models that have been estimated – i.e. by mode and purpose. No further segmentations are imposed at this stage, and our chief interest is in a) the variation between modes and purposes, and b) the interpretation of the different values for different kinds of time, emanating from different SP games. As has already been noted, we use distance weighting, imputed costs where necessary, income option (1), and adopt a value of Δt = 10 minutes. All other 'design effects' are neutralised in the model estimation.

On this basis, **Table 7.5**, **Table 7.6**, **Table 7.7** and **Table 7.8** present results from the Tool for the modes car, rail, bus and 'other PT' respectively. All values are in £/hr. Based on the arguments outlined in **Chapter 6**, the business values are taken from the employees' business SP.

Table 7.5: Car VTT values from preferred behavioural models by type of time, and trip purpose (Income option 1, 2014 perceived prices, £/hr)

			Commute	Employees' business	Other non-work
Car	VTT	SP1	11.70	16.74	4.91
	Average travel time	SP2	15.36	25.49	10.67
	Value of sd travel time	SP2sd	5.00	10.79	3.77
	Free-flow	SP3ff	6.00	7.04	2.32
	Light traffic	SP3lc	8.41	11.34	4.09
	Heavy traffic	SP3hc	15.98	21.03	9.26

Notes: Distance-weighted, income option 1, based on sample enumeration from NTS 2010-2012 motorised trips; Δt=10; Tool version 1.1 (fuel costs imputed and employers paying for EB trips).

Table 7.6: Rail VTT values from preferred behavioural models by type of time, and trip purpose (Income option 1, 2014 perceived prices, £/hr)

			Commute	Employees' business	Other non-work
Rail	VTT	SP1	12.42	27.61	8.68
	Scheduled travel time	SP2	16.60	47.32	10.07
	Early	SP2early	-29.45	-73.24	-23.60
	Late	SP2late	47.50	130.73	32.38
	seated 50% load	SP3c1	8.90	22.07	6.77
	seated 75% load	SP3c2	9.65	22.35	6.77
	seated 100% load	SP3c3	12.27	29.25	9.46
	seated 1 pass per m^2	SP3c4	13.34	33.17	10.75
	seated 3 pass per m^2	SP3c5	16.08	39.65	13.18
	standing 0.5 pass per m^2	SP3c6	14.20	37.62	11.46
	standing 1 pass per m^2	SP3c7	14.65	40.48	11.97
	standing 2 pass per m^2	SP3c8	16.14	45.75	14.82
	standing 3 pass per m^2	SP3c9	19.30	47.06	16.89
	standing 4 pass per m^2	SP3c10	22.81	59.32	20.49

Notes: Distance-weighted, income option 1, based on sample enumeration from NTS 2010-2012 motorised trips; Δt=10; Tool version 1.1 (cost imputed for a trip with a zero cost and employers paying for EB trips).

Table 7.7: Bus VTT values from preferred behavioural models by type of time, and trip purpose (Income option 1, 2014 perceived prices, £/hr)

			Commute	Employees' business	Other non-work
Bus	VTT	SP1	3.15	-	3.26
	Scheduled travel time	SP2	4.62	-	3.43
	Early	SP2early	-12.45	-	-10.95
	Value of late	SP2late	13.31	-	8.64
	Plenty of seats free and did not have to sit next to anyone.	SP3c1	2.96	-	2.52
	A few seats free but had to sit next to someone/could not sit with people travelling with.	SP3c2	3.10	-	2.55
	A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	SP3c3	3.47	-	3.03
	No seats free – a few others standing.	SP3c4	4.28	-	3.95
	No seats free – densely packed.	SP3c5	7.42	-	7.01
	Value of free-flow	SP3ff	3.43	-	3.69
	Value of slow down	SP3sd	4.83	-	4.13
	Value of dwell time	SP3dw	2.35	-	4.74
	Value of headway	SP3hw	5.84	-	4.83

Notes: Distance-weighted, income option 1, based on sample enumeration from NTS 2010-2012 motorised trips; Δt=10; Tool version 1.1 (cost imputed for a trip with a zero cost).

Table 7.8: 'Other PT' VTT values from preferred behavioural models by type of time, and trip purpose (Income option 1, 2014 perceived prices, £/hr)

			Commute	Employees' business	Other non-work
'Other PT'	VTT	SP1	6.35	8.33	5.23
	Scheduled travel time	SP2	6.32	7.89	5.17
	Early	SP2early	-15.18	-13.13	-15.38
	Late	SP2late	11.09	15.40	11.61
	Plenty of seats free and did not have to sit next to anyone.	SP3c1	4.41	6.24	4.22
	A few seats free but had to sit next to someone/could not sit with people travelling with.	SP3c2	4.47	6.24	4.22
	A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	SP3c3	4.63	6.24	4.22
	No seats free – a few others standing.	SP3c4	5.24	7.33	4.65
	No seats free – densely packed.	SP3c5	7.89	11.12	7.90

Notes: Distance-weighted, income option 1, based on sample enumeration from NTS 2010-2012 motorised trips; Δ*t*=10; Tool version 1.1 (cost impute for a trip with a zero cost and employers paying for EB trips).

With reference to the earlier discussion of size effects (**Section 7.5.3**), the variation in VTT by mode by trip by Δ*t* is presented in **Table 7.9** below. Multiplying these ratios by the values in the above VTT tables will give the VTTs for different Δ*t*. Across the purposes, Δ*t* effects are largest for commute and smallest for other non-work. Across the modes, Δ*t* effects are largest for car; there are no Δ*t* effects for bus.

Table 7.9: Variation in SP1 VTT by Δ*t*, mode and purpose (Δ*t* = 10 is the base)

Mode	Δ*t* (mins)	Commute	Employees' business	Other non-work
Car	1	0.40	0.58	0.54
	5	0.76	0.85	0.83
	10	1.00	1.00	1.00
	15	1.18	1.10	1.11
	20	1.32	1.18	1.20
Bus	1	1.00	1.00	1.00
	5	1.00	1.00	1.00
	10	1.00	1.00	1.00
	15	1.00	1.00	1.00
	20	1.00	1.00	1.00
'Other PT'	1	0.69	0.40	0.76
	5	0.89	0.76	0.92
	10	1.00	1.00	1.00
	15	1.07	1.17	1.05
	20	1.12	1.32	1.08

Mode	Δt (mins)	Commute	Employees' business	Other non-work
Rail	1	0.61	0.59	0.74
	5	0.86	0.85	0.91
	10	1.00	1.00	1.00
	15	1.09	1.10	1.06
	20	1.16	1.17	1.10
All modes	1	0.45	0.58	0.58
	5	0.78	0.85	0.85
	10	1.00	1.00	1.00
	15	1.16	1.10	1.10
	20	1.29	1.18	1.18

Notes: Modelling used imputation for zero reported PT costs and employers paying for EB trips. Tool version 1.1. SP1 with different Δt.

There are a large number of issues to discuss, but the most problematic relate to the car results.

Car

The results from the simplest time/cost trade-off (SP1) – which is, of course, most comparable to the results on which the existing methodology, using the 1994 AHCG data, is based – give a value of £11.70 for commuters, £16.74 for employees' business and £4.91 for other non-work. In both cases, the values lie between the values from SP3 for light and heavy traffic, and this is in line with the underlying model, since the relativities have not been changed as a result of the NTS re-weighting.

We have to form a judgment as to whether these are compatible, and if not, which values to use. To do this, we need to reflect on how the time values were presented in the SP tasks. For SP1, the instructions were

"Please imagine that each situation is exactly the same as for #LEG2# actual car journey at the time you made the journey, except ….The one way travel time may be different because of changes in congestion". Variations in cost were suggested in terms of changed fuel cost.

It is noteworthy that in the 1994 survey (which of course was for car only) the corresponding instructions were:

"Please imagine that each situation is exactly the same as for your actual journey at the time you were surveyed except….the travel time … can be different from the actual situation at that time because there is, for example, more or less congestion." Variations in cost were suggested in terms of changes in petrol price or parking charges.

Hence in terms of the background to the SP1 experiment, it seems fair to conclude that they were identical in both surveys, with a clear suggestion to relate the changes to the conditions of the reference trip. Any changes in results should therefore be attributed to:

a) changes in the SP design (this should largely improve the accuracy rather than lead to different results *per se*; see **Chapter 2**)

b) changes in preferences and behaviour etc. (including the underlying NTS travel characteristics; see **Chapter 3**)

c) changes in the method of analysis (in particular the switch from additive to multiplicative model specifications; see **Chapter 4**)

Turning to the results from SP3, again there is a relation to the reference trip. Respondents had previously been asked 'about how much' of their time was spent in each of the three following conditions, presented in words and pictures, as:

"Heavy traffic: *Your speed is noticeably restricted and frequent gear changes are required".*

"Light traffic: *You can travel close to the speed limit most of the time, but you have to slow down every so often".*

"Free-flowing: *You can travel at your own speed with no problems over-taking".*

The implied proportions have been used as covariates in the SP1 model, but the effects are generally weak[81] and far lower than we find for SP3. And, of course, as the NTS data does not contain any information about the time spent in different conditions, they cannot be used in the Tool. For covariates such as these, the Tool uses averages from the SP sample. By contrast, the SP3 values of time for commuters display ratios relative to free-flow time (ff) of 1.4 for light traffic (lc) (interestingly, this is the same as the multiplier using the actual proportion) and 2.66 for heavy traffic (hc). For business the ratios were 1.61 and 2.99 and for other non-work they were 1.76 and 3.98.

This latter value does seem extremely high. For example, if we assume a free-flow speed of 70 mph and a heavy traffic speed of 25 mph (with an inter-urban trip in mind), then the heavy traffic/free-flow ratio of about 4 for other non-work implies that to avoid a stretch of heavy traffic drivers would be willing to take a free-flow diversionary route 11 times as long, which seems unreasonable, even if one also considers that the user would receive a more reliable trip travelling at free-flow speeds than in heavy traffic conditions. For commute it is less extreme but still about 7.5 times the distance. This might work for very short stretches, but it does not seem generally plausible. It would require some very good route choice evidence to support such a range.

As noted, NTS does not carry information on conditions for individual trips. The Tool has been set up to calculate an average of the three types of time in SP3, using the average shares in the SP sample for different purposes, which are as shown in **Table 7.10**.

---

[81] The multipliers (relative to free-flow) were reported for commuting as 1.40 for light traffic and 1.56 for heavy traffic: however, neither of these were significantly different from 1 at the 95% level. The corresponding values for other non-work were 1.36 and 1.46, again with neither being significantly different from 1.

Table 7.10: Average shares of congested driving conditions in the SP sample

	Car		
	Commute	**Employees' business**	**Other non-work**
Free-flow	0.33	0.34	0.36
Light traffic	0.36	0.40	0.41
Heavy traffic	0.31	0.26	0.23

Using these overall proportions, we obtain SP3-based values of £9.98 for commute, £12.44 for business and £4.62 for other non-work. These are lower than the SP1-based values: the commute value is 15% lower, the business value 26% lower and the other non-work value 6% lower. But it is arguable as to whether this is appropriate: the proportions are likely to be related to the kind of trip (urban vs. inter-urban etc.), and the SP sample is certainly not representative in terms of distance travelled.

The fact that, at least on average, respondents seem to have allocated the three types of time in relatively equal proportions casts some light on how the question may have been interpreted. For example, the description of speed under high congestion conditions being 'noticeably restricted', and the relatively high proportions of time spent in these conditions, suggest that this has not been interpreted as absolute gridlock. Given this, the assumption in the example of 25mph in relation to heavy traffic may be too pessimistic. However, even assuming 50mph would only halve the length of implied diversionary route.

As far as the implied variation in valuations for the three levels is concerned, we can note that the levels presented are three possible positions within a continuum related to the 'volume/capacity' ratio, and would be very hard to apply in practice, since they would need to be aligned with actual traffic conditions in relation to volume-delay functions. Within the SP, no explicit reason is given for the changes in traffic conditions presented. There is, therefore, a possibility that respondents, having been asked explicitly to focus on the three 'levels' of congestion, may over-react to these.

On balance, it seems safer to rely on the SP1 values for the immediate future, while treating the SP3 results as providing some indication of the sensitivity to higher congestion. The future availability of data on differing levels of congestion may permit the use of SP3 results in the medium to longer term, though there would still be the need to resolve how model output (in terms of volume/capacity ratios) relates to the levels of congestion presented in the SP, and how to interpolate between the free-flow, light and heavy traffic points.

Turning to the SP2 results, we find higher valuations based on the average time presented in the reliability experiment relative to the SP1 values, by a factor of 1.31 for commute, 1.52 for business and 2.17 for other non-work. Now it might be argued that by implying the possibility of unreliability, there is some suggestion of (greater) congestion. However, the questionnaire says that the situation is the same as the reference trip, while the reasons for variation in overall travel time could be due to 'improvements in traffic control', and the variation (unreliability) could be due to 'breakdowns, unplanned roadworks, or general traffic'. It is not obvious that this has to imply that SP2 values > SP1 values, particularly not at the scale seen for other non-work, where the value is well in excess of that for heavy traffic.

This also presents a problem for the ‘reliability ratio’, as we have to decide whether to take the value of the standard deviation relative to the SP2 VTT or the SP1 VTT. If we do the former, we get values of 0.33 (commuting), 0.42 (business) and 0.35 (other non-work): for the latter, the values are 0.43, 0.64 and 0.77 respectively. The former values are low by ‘received wisdom’ (though the evidence base for that is not especially strong), while at least for other non-work, the SP1-based result is close to the current WebTAG value of 0.8.

On balance, it seems more reasonable to interpret them relative to the SP2 time multiplier, on grounds of internal consistency within the SP2 experiment. This is also in line with the way reliability ratios have been derived in other work (e.g. Black and Towriss (1993)[82]). However, this is not a strongly-based recommendation, and we are still left with the conundrum of explaining the high SP2 time multiplier.

The final consideration is the relationship between the overall commuting and other non-work values (we discuss the level of the business values in more detail in **Section 7.6.6**). Unlike the 2003 study, where the distinction was marginal, with commuting only higher by about 9%, here they (the SP1 values) differ by a factor of 2.4 (the result is lower, at 1.44, for the SP2 values but, as noted, these values are harder to justify, especially in terms of the SP3 results). This also compares to the average 12% difference between commuting and other non-work found by Abrantes and Wardman (2011)[83] in their meta-analysis. Referring back to the 2003 study, we also note that the increase in the ratio between commuting and other non-work arises because the commuting values have increased and the other non-work values have decreased (against existing WebTAG values). Some of these differences will be due to income and distance, as we discuss later, but for the car mode at least the difference remains substantial, and, on the face of it, requires additional research to be fully defensible.

Rail

We now turn to the other modes, where the results are more generally consistent between the different SP games, dealing firstly with rail.

The results from the time/cost trade-off (SP1) give a value of £12.42 for commuters, £27.61 for business and £8.68 for other non-work. For commuters, this is very slightly above the ‘seated 100% load’ value from SP3, while for business and other non-work, it is in the range between 75% and 100% load. Once again, this is in line with the underlying model, since the relativities have not been changed as a result of the re-weighting. These results seem reasonable.

For SP1, the instructions were:

“Please imagine that each situation is exactly the same as for #LEG2# actual rail journey at the time you made the journey, except …. The one-way travel time may

[82] Black, I. G. and Towriss, J. G. (1993) ‘Demand Effects of Travel Time Reliability’, Centre for Logistics and Transportation, Cranfield Institute of Technology, Great Britain

[83] Abrantes, P.A.L. and Wardman, M.R. (2011) ‘Meta-analysis of UK values of travel time: An update’. Transportation Research A, 45 (1), pp1–17.

be different because, for example, there are slower or faster trains". Variations in cost were simply suggested in terms of changed fares.

Turning to the results from SP3, again there is a relation to the reference trip. Passengers were asked:

"When you boarded the train, how crowded was it?

(for travellers on their own:)

- *Plenty of seats free and did not have to sit next to anyone*
- *A few seats free but had to sit next to someone. No one standing*
- *A few seats free but had to sit next to someone. Some people were standing*
- *No seats free – a few others standing*
- *No seats free – densely packed*

(for travellers with others, the first three levels were substituted by):

- *Could sit with people travelling with me*
- *Could not sit with people travelling with me*
- *A few seats free but could not sit with people travelling with me. No one standing*
- *A few seats free but could not sit with people travelling with me. Some people were standing"*

Travellers were also asked if the level of crowding changed as the trip progressed and whether they had to stand for any part of their trip (how much?). As described in the SP, the crowding level relates to 'when you boarded'.

In the SP, a visual presentation was given of 10 crowding levels, with some verbal description (**Section 2.12.2**). Whilst there is precedent for using this form of presentation, it is not clear how easy it would be to relate these to the passenger's actual trip: the description in fact relates to the number of persons standing round the door and (level 10 only) in the aisles, though this is an attempt to convey the level of passengers/metre^2. The visual presentation is not train-specific.

The crowding multipliers, relative to the lowest level (seated 50% load), are as shown in **Table 7.11**. These are slightly higher than current PDFH recommendations (which also vary by London & SE, Regional, and Intercity services), but generally show considerable compatibility. Note that the PDFH multipliers are recommended for use in appraisal in WebTAG (Unit A5.3), and given the weight of evidence behind these, we suggest that the current results be treated as corroborative rather than definitive.

Table 7.11: Rail crowding multipliers

	Commute	Employees' business	Other non-work
Seated			
seated 50% load	1	1	1
seated 75% load	1.084	1.013	1.000
seated 100% load	1.378	1.326	1.396
seated 1 pass per m2	1.499	1.503	1.587
seated 3 pass per m2	1.806	1.797	1.947
Standing			
standing 0.5 pass per m2	1.595	1.705	1.693
standing 1 pass per m2	1.646	1.835	1.767
standing 2 pass per m2	1.813	2.073	2.189
standing 3 pass per m2	2.168	2.133	2.494
standing 4 pass per m2	2.563	2.689	3.026

Notes: Distance-weighted, income option 1, based on sample enumeration from NTS 2010-2012 motorised trips.

Turning to the SP2 results, once again we find higher valuations based on the average time presented in the reliability experiment, by a factor of 1.34 for commute and 1.16 for other non-work (in the latter case the multiplier was not significantly different from 1 in the estimation): the factor for business seems unreasonably high, at 1.71. The rubric used in the questionnaire for SP2 is identical to that for SP1, but possible variations (reliability) are attributable to 'service disruptions or incidents on the line'. There has been some uncertainty in the modelling as to whether the average times presented in this experiment could be treated as scheduled times (which is what would be desirable to give an appropriate interpretation to the 'early and late' multipliers). In spite of this, the early and late effects appear to be generally plausible, with substantially higher values for unplanned shifts in arrival time and a general tendency for equal and opposite effects, but with late values somewhat higher. Given the reasonable agreement between SP1 and SP2, these conclusions are not greatly affected by whether we adopt SP1 or SP2 as the 'base', and, in line with the car mode, we will adopt SP1.

Current WebTAG guidance on valuing rail reliability again directs users to PDFH, wherein multipliers vary by flow type and purposes. For non-rail PT, WebTAG advocates a multiplier of 3 (based on a wait multiplier of 2.5 multiplied by a factor of 1.2 to reflect the variability of lateness). The multipliers implied by **Table 7.6**, specifically 2.86 (commuting), 2.76 (business) and 3.21 (other non-work), show close correspondence to existing guidance. The other measure of public transport reliability – the 'PT reliability ratio' (value of SD of lateness / value of average lateness) – was not included in the SP2 model for rail.

Note that A1.3 also recommends use of waiting time multipliers (again 2.5): for rail, however, the more usual convention is to include interval penalties from PDFH in 'generalised journey time' (GJT) and then value changes in GJT. The rail SP experiments did not in general include a headway variable. Headway was considered in the mode and operator choice models reported in **Chapter 5**, but these results were judged to be less satisfactory than the core models reported in **Chapter 4** and were not taken forward to the Tool.

As with the car mode, but to a lesser extent, the commuting (SP1) values are higher than the other non-work values – here by 43% (if we use the SP2 values, the ratio is 65% higher). It turns out that about half of this effect can be ascribed to the income distribution. Again, we discuss the level of the business values in more detail in **Section 7.6.6**.

Bus

For **bus**, the results from the time/cost trade-off (SP1) give a value of £3.15 for commuters and £3.26 for other non-work: thus, substantially lower than the previous two modes. For commuters, this is very slightly above the level 2 crowding value ('A few seats free but had to sit next to someone/could not sit with people travelling with') from SP3, while for other non-work, it is in the range between levels 3 and 4 (see **Table 7.12** for definitions). Once again, this is in line with the underlying model, since the relativities have not been changed as a result of the NTS re-weighting. These results seem reasonable, though the position of the two purposes with respect to crowding is slightly unexpected.

For SP1, the instructions were identical to those for rail. For bus, SP3 is divided between games SP3a and SP3b: the former varies travel times split by different conditions (normal speed, slowing down/congested, dwell time at stops) as well as frequency, while the latter deals with crowding. For SP3a, the rubric noted:

"In addition to variations in fare and service frequency between the two options, we also vary the amount of time in the following situations as a result of:

- *Differences in the amount of time the bus spends at bus stops*
- *Time spent slowing down or in congested traffic*
- *The amount of time spent at normal speeds*

These changes could come about because of the introduction of bus priority schemes."

The results for SP3a are more difficult to interpret (and there are few studies with which to compare). On the face of it, normal speed time is valued slightly more highly than the SP1 value (for both purposes), but oddly, for commuters a lower value is ascribed to dwell time. For both purposes, time spent slowing down is valued more highly than normal speed time. Relative to normal speed time, headway has a multiplier of 1.70 for commuters and 1.31 for other non-work. If the services are frequent, it is reasonable to assume that waiting time is on average half the headway, so these translate into waiting time multipliers of 3.40 and 2.62 respectively, compared with the recommended value (TAG A1.3) of 2.5.

For SP3b, the same rubric as for rail is used for the reference level, but the levels offered in the SP are different: they relate directly to the levels used to describe the reference journey (again, at the time of boarding). The crowding multipliers, relative to the lowest level ('plenty of seats free and did not have to sit next to anyone') are shown in **Table 7.12**.

Table 7.12: Bus crowding multipliers

	Commute	Employees' business	Other non-work
Plenty of seats free and did not have to sit next to anyone	1.000	-	1.000
A few seats free but had to sit next to someone/could not sit with people travelling with.	1.047	-	1.012
A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	1.172	-	1.202
No seats free – a few others standing.	1.446	-	1.567
No seats free – densely packed.	2.507	-	2.782

Notes: Distance-weighted, income option 1, based on sample enumeration from NTS 2010-2012 motorised trips.

These are similar between the two purposes and generally plausible: the most crowded multipliers are high, but not unreasonable, and can be compared with the corresponding rail results.

Turning to the SP2 results, once again we find higher valuations based on the average time presented in the reliability experiment, by a factor of 1.47 for commute and 1.05 for other non-work (in the latter case the multiplier was not significantly different from 1 in the estimation). The rubric used in the questionnaire for SP2 is identical to that for SP1, but the reasons for possible variations (reliability) given are the same as those for car. The early and late effects appear to be generally plausible, with substantially higher values for unplanned shifts in arrival time and a general tendency for equal and opposite effects, but with late values somewhat higher.

Unlike the two previous modes, the commuting SP1 values are more or less the same as the other non-work values – though if the SP2 values are taken then they are higher by 35%.

'Other PT'

Finally, we consider 'other PT'. The results from the time/cost trade-off (SP1) give a value of £6.35 for commuters, £8.33 for business and £5.23 for other non-work: thus, intermediate between bus on the one hand and car and rail on the other. For all three purposes, the SP1 values lie between crowding levels 4 and 5, which might suggest relatively crowded conditions prevailing (this is unsurprising given the focus of the data collection on London). Once again, this is in line with the underlying model, since the relativities have not been changed as a result of the NTS re-weighting. These results seem reasonable, though the position of the three purposes with respect to crowding is slightly unexpected.

For SP1, the instructions were identical to those for rail. There is no equivalent of SP3a for bus, but for SP3b, the same rubric as for rail is used for the reference level, while the levels offered in the SP are the same as for bus, relating directly to the levels used to describe the reference trip (again, at the time of boarding).

The crowding multipliers, relative to the lowest level ('plenty of seats free and did not have to sit next to anyone') are shown in **Table 7.13**.

Table 7.13: 'Other PT' crowding multipliers

	Commute	Employees' business	Other non-work
Plenty of seats free and did not have to sit next to anyone	1.000	-	1.000
A few seats free but had to sit next to someone/could not sit with people travelling with.	1.014	1.000	1.000
A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	1.050	1.000	1.000
No seats free – a few others standing.	1.188	1.175	1.102
No seats free – densely packed.	1.789	1.781	1.872

Notes: Distance-weighted, income option 1, based on sample enumeration from NTS 2010-2012 motorised trips.

These are similar between the three purposes and are generally plausible: the most crowded multipliers are high, but not unreasonable, and can be compared with the corresponding rail results. It is noticeable that there is no substantial effect until very high levels of crowding.

Turning to the SP2 results, in this case we find no major differences in the valuations based on the average time presented in the reliability experiment. The rubric used in the questionnaire for SP2 is identical to that for SP1, with the reasons for possible variations (reliability) being the same as those for rail. The early and late effects appear to be generally plausible, with substantially higher values for unplanned shifts in arrival time and a general tendency for equal and opposite effects, but, somewhat oddly, with early values rather higher in this case.

As generally with the other modes, the commuting values are higher than the other non-work values – here by 21%. In this case there are no substantial differences in the income distributions for the two purposes.

On the basis of the discussion so far, we can note some provisional conclusions. In relation to the issue of which of the three SP experiments should have primacy for general recommendations, generally we consider that the SP1 values are preferred, at least for the immediate future. When deriving ratios/multipliers, we generally recommend deriving these 'within game' for consistency, while recognising that there remain issues for application. As far as variations by purpose are concerned, the differences in values between commuting and other non-work are larger than we would expect *a priori*, and by no means all of them can be explained by income variation or distance (as we shall see). The business values are largest, as we would expect, though we postpone discussion on their levels until **Section 7.6.6**.

7.6.3 Weighting

Section 7.3 presented a conceptual discussion regarding the preference between distance or trip weighting, and concluded that both were approximations, primarily due to the differing distance profiles associated with different schemes being appraised. It also suggested that varying VTT by distance band and possibly by geography type (urban, inter-urban), since geography types will be correlated with distance, would reduce these discrepancies. This is because the ratio between distance-weighted and trip-weighted VTTs becomes much smaller when values are disaggregated by distance.

We now present ratios of VTTs showing how the differences between trip-weighted and distance-weighted values change when we disaggregate by distance. As can be seen from **Table 7.14**, distance-weighted VTT is up to double the trip-weighted VTT if we do not disaggregate by distance. However, if we disaggregate into three trip bands (<20 miles, between 20 and 100 miles, and more than 100 miles), then this gives rise to ratios between distance-weighted values and trip-weighted values that only exceed 1.2 for commute trips <20 miles. **Table 7.14** also shows that the very different 'distance' markets served by different modes (e.g. bus vs. rail) mean that the ratio of distance- to trip-weighted VTT is mode-dependent, and it is really only when one controls for distance that the ratio becomes more similar across modes. We therefore recommend that some disaggregation of VTT by distance be adopted. For the conceptual reason that the probability of a trip benefitting from the scheme being appraised is proportional to trip length within each distance band, the VTT should be distance-weighted.

Table 7.14: Ratio of distance-weighted to trip-weighted averages by mode and trip purpose (2014 study)

		Car	Bus	'Other PT'	Rail
All distances	Commute	1.28	0.99	1.01	1.37
	Employee's business	1.83	-	0.99	1.60
	Other non-work	1.85	1.04	0.97	1.19
Trips 0 to 20 miles	Commute	1.14	1.22	1.25	1.30
	Employee's business	0.98	-	1.03	1.01
	Other non-work	1.00	0.99	0.97	1.00
Trips 20 to 100 miles	Commute	1.04	1.01	1.02	1.07
	Employees' business	1.09	-	1.01	1.11
	Other non-work	1.09	1.01	1.01	1.03
Trips > 100miles	Commute	1.00	-	-	1.06
	Employees' business	1.03	-	-	1.02
	Other non-work	1.04	0.95	-	1.00

Note: Values estimated using traveller income and $\Delta t = 10$. All coefficients by mode, purpose and distance have same or similar ratio to SP1. Distance is distance travelled by main mode.

Further work, outside the scope of the current study, involving distance profiles from real scheme appraisals would be needed to identify the full implications of the approximations to the 'real' VTT of the distance-weighted VTT under different distance disaggregation. This research would also determine the exact distance banding to be used in appraisal.

7.6.4 Further considerations relating to income and distance

As was noted in **Section 7.2** above, the main variation taken account of in the 2003 study was that due to income and distance. The current study has identified income elasticities of various sizes for most combinations of mode and purpose: in the case of distance, it has not generally been found that there is any **independent** effect (with the exception of the bus mode), but quite strong effects have been found due to the reference time and cost, and these are both substantially correlated with distance. Thus in practice we may expect some variation in VTT due to distance.

Ignoring size effects etc., the essential utility function for SP1 can be interpreted as having the form:

$$U = \mu \cdot \left[V_{ref} \cdot \frac{C^{\lambda_C}}{T^{-\lambda_T}} \cdot \Delta T + \Delta C \right] \qquad (7.2)$$

where by introducing the minus sign we have taken note of the fact that the models generally produce $\lambda_C > 0$ and $\lambda_T < 0$, and where T and C refer to the base values of time and cost for a particular traveller.

This can be re-scaled to give:

$$U = \theta_{T\,ref} \cdot \Delta T \cdot \left(\frac{1}{T} \right)^{-\lambda_T} + \theta_{C\,ref} \cdot \Delta C \cdot \left(\frac{1}{C} \right)^{\lambda_C} \qquad (7.3)$$

In this sense, it can be seen as a test of the 'proportionate' response – i.e. that a given change (both for time and cost) has a lesser impact on utility the smaller the proportion it is of the reference value (T and C). A completely proportionate effect would produce values for the exponents of 1, and an absence of a proportionate effect would produce values of 0: both exponents in the equation are positive. The model results imply that the proportionate effect is present for both variables.

While mainstream economic theory does not support the proportionate approach (certainly not for cost, where £1 should have the same value regardless of where it is saved or spent), it is commonly observed in empirical work. In practice, of course, we are not observing the same individual with different VTT for different levels of cost and time: the variation is between different travellers. The data is silent as to whether the proportionate effect would also apply to a single traveller. Thus the reference time and cost elasticities pick up some self-selectivity effects we would expect in cross-sectional data.

In all cases, the elasticities to time and cost are more or less equal and opposite: they are substantially higher than was found in the 2003 work (where the variation related to cost only), suggesting a large proportionate response. However, typically these two coefficients are very highly (negatively) correlated: e.g. for car the correlations were -0.928 for commute and -0.936 for other. This suggests that the absolute levels are less reliable than the difference between the absolute values.

We can get some idea of the direct effects of income on VTT by examining the elasticities for both income and distance from the model estimations presented in

Chapter 4. These are reproduced in the third and fifth columns of **Table 7.15**. If we make the simplifying assumption that both cost and time are reasonably proportional to distance, then the implied distance elasticity is obtained by adding the two coefficients (and including the distance coefficient where it has been found to be significant). This implied distance elasticity is presented in the sixth column of **Table 7.15**.

Table 7.15: Modelled and implied income and distance elasticities (by mode and purpose)

Mode	Purpose	Income elasticity		Distance elasticity		
		Model (see Chapter 4)	Estimated from regression analysis of VTT output by the Tool	Model (see Chapter 4)	Estimated by summing distance, time and cost elasticities	Estimated from regression analysis of VTT output by the Tool
Car	Commuting	0.580	0.601	0	0.055	0.179
	Employees' business	0.300	0.320	0.239	0.236	0.340
	Other non-work	0.682	0.532	0	0.122	0.298
Rail	Commuting	0.298	0.339	0	0.389	0.306
	Employees' business	0.357	0.428	0.059**	0.453	0.370
	Other non-work	0.294	0.314	0	0.058	0.088
Bus	Commuting	0	0.024	0.153**	0.099	-0.037
	Other non-work	0.457	0.461	0.076**	0.290	0.063
'Other PT'	Commuting	0.117	0.134	0	0.141	0.043
	Employees' business	0.395	0.357	0	-0.488	0.0154**
	Other non-work	0.271	0.379	0	-0.007	-0.072

** Insignificantly different from zero at the 5% level

However, given that this simplifying assumption does not in fact hold in the NTS and that some covariates (other than time and cost) vary with distance, it was felt that a more appropriate way of understanding the practical impact of distance and income was to use the Implementation Tool to calculate the VTT for each record in the NTS data, unweighted by distance, income, or other weighting configuration. Then a regression analysis of the form $ln(VTT) = constant + \omega_D.ln(trip\_distance) + \omega_Y.ln(income)$ was conducted, thus directly estimating the distance and income elasticities. These runs use household income for commute and other non-work, and personal income for employees' business. Sensitivity tests that weighted the trips in accordance with the within-record error in VTT did not give significantly different results. These results are presented in columns 4 and 7 of **Table 7.15**. We believe they are plausible, and they also take due account of the observed correlation between income and distance (as well as other covariates).

Generally, the 'implied' income elasticities (column 4) are of the same order as the direct model estimates (column 3), but with some variations. There are more differences in the 'implied' distance estimates (between columns 6 and 7), and we conclude that the estimates from the regression model (column 7) are more reliable (especially as they do not require the assumption that time and cost are linear with distance, and account for how other covariates vary with time and cost). We also note that the income elasticities for business travel are much lower than 1 – as would be implied by the CSA. We will return to this point in **Section 7.6.6** below.

The values for the distance elasticity taken from the 2003 study (NB for car only), as reported in WebTAG (Unit M2, variable demand modelling, p64, para. C.3.4) were 0.421 (commute) and 0.315 (other non-work), so the new car values are lower. Of course, it must be remembered that the values in WebTAG are really giving a cost elasticity (not distance) and are based on models that exclude several other covariates (but include income). Arguably the distance elasticities from the 2003 study were high, as the evidence base (e.g. Abrantes and Wardman, 2011) is reasonably well aligned to the implied elasticities in **Table 7.15**. With the exception of rail commute and business, the implied 'distance elasticity' is low for the public transport modes. Note that WebTAG does not give a distance elasticity for business (nor elasticities for other modes).

By contrast, the income elasticities for car values are higher than found in the 2003 work (0.36 and 0.16 respectively), though they are not unreasonable. The general experience is that cross-sectional income elasticities are lower than those found from temporal analysis. Given that there was some feeling from the 2003 work that the income elasticity was low and the distance elasticity high, these results are more plausible on both counts. There remains some concern about the low income elasticity for business, though this derives from the estimated models: this is further discussed in **Section 7.6.6**.

With this in mind, we can now consider the impact of income option (2), rather than option (1). In this method, all non-work trips in the NTS sample are treated as if they had the mean household income (this is a trip-weighted average across non-work trips of £49,684), and all business trips as if they had the mean personal income of £35,070 for car, £20,219 for bus, £45,019 for 'other PT' and £55,319 for rail (these are trip-weighted income averages across the business sample of the NTS by mode).

Generally, it is noteworthy how little difference it makes to the 'All modes' average values (right hand column of **Table 7.16**). *A priori* one would expect income option (2) to increase the VTT of modes and/or trip purposes that have users with lower than average incomes, and to decrease the VTT for those with higher than average incomes. This is clear from **Table 7.16** where bus VTT increases and rail decreases, whilst commuting typically decreases. The other effect of income option (2), not seen in **Table 7.16**, is that with all households having the same income, the distribution of the VTT is less skewed than it is with income option (1). Thus income option (2) not only alters the mean VTT but also the distribution of the VTT (by mode and purpose).f

Table 7.16: Ratio of VTT by income treatment (all distances)

		Car	Bus	'Other PT'	Rail	All modes
Fixed income (2)/ trip specific income (1)	Commute	0.89	1.00	0.99	0.87	0.89
	Employees' business	0.97	-	1.13	1.01	0.98
	Other non-work	1.07	1.43	1.02	1.02	1.07
Green Book (4)/ trip specific income (1)	Commute	0.85	0.99	0.94	0.88	0.85
	Employees' business	1.00	-	1.00	1.00	1.00
	Other non-work	0.80	0.86	0.81	0.85	0.80

Notes: Fixed income trip-weighted (option 2) is £49,684 for household (non-work) income. For personal income (EB) it is £35,070 for Car, £20,219 for Bus, £45,019 for 'Other PT', £55,319 for Rail. Ratio of distance-weighted VTT. Modelling used imputation for zero reported PT costs and employers paying for EB trips. Δt=10. Tool version 1.1.

We can also see this effect when we disaggregate by distance (as in **Table 7.17**). Typically those with lower incomes travel shorter distances – therefore we expect that the VTT for shorter distance trips increases under income option (2), whilst that for longer distance trips decreases. This is indeed what we find.

Table 7.17: Ratio of VTT by income treatment for short and long distance trips (all purposes, all modes)

	< 20 miles	=> 20 miles
Fixed income (2)/ Trip specific income (1)	1.06	0.96
Green Book (4)/ Trip specific income (1)	0.82	0.88

Notes: Fixed income trip-weighted (option 2) is £49,684 for household (non-work) income. For personal income (EB) it is £35,070 for Car, £20,219 for Bus, £45,019 for 'Other PT', £55,319 for Rail. Ratio of distance-weighted VTT. Modelling used imputation for zero reported PT costs and employers paying for EB trips. Δt=10. Tool version 1.1.

Turning now to the Green Book weights, given the higher incomes of travellers on average compared with the national population, we would expect the Green Book weights to adjust the VTTs downward. This is very clear from **Table 7.16**. For reasons discussed above, however, we do not recommend the use of Green Book weighted VTT for use in current appraisals.

Values by distance and area type

All the trip purposes are sensitive to distance (as previously presented in **Table 7.15**), but to varying degrees. Commute is the least sensitive, whilst business is the most sensitive. As previously discussed, these variations are due to the way socio-economic and trip characteristics systematically vary with distance, as well as to the pure distance elasticities. **Table 7.18** presents VTT values by distance, mode and purpose for a range of distances. The 'implied' distance elasticities for business are substantially larger than those for commuting, leading to business values just less than commuting values at short distances and substantially greater than commuting values at long distances.

Table 7.18: Values by mode and by distance (£/hr, perceived prices)

		Commute	Other non-work	Employees' business
All modes	All distances	11.21	5.12	18.23
	<5 miles	6.61	2.30	5.39
	<20miles	8.70	3.11	8.31
	5-20 miles	9.15	3.47	8.84
	20 to 100 miles	13.61	6.14	16.05
	>=20 miles	14.24	7.24	21.14
	>=50 miles	17.55	8.48	24.55
	>=100 miles	n/a	9.25	28.62
Car	All distances	11.70	4.91	16.74
	<5 miles	7.25	2.15	5.27
	<20miles	9.61	2.99	8.21
	5-20 miles	10.12	3.36	8.79
	20 to 100 miles	13.98	5.97	15.85
	>=20 miles	14.35	7.01	19.51
	>=50 miles	17.18	8.30	22.53
	>=100 miles	n/a	9.08	25.74
Bus	All distances	3.15	3.26	-
	<5 miles	3.22	3.10	-
	<20miles	3.10	3.20	-
	5-20 miles	3.05	3.27	-
	>=20 miles	n/a	3.67	-
	20 to 100 miles	n/a	3.71	-
	>=50 miles	n/a	n/a	-
	>=100 miles	n/a	n/a	-
'Other' PT	All distances	6.35	5.23	8.33
	<5 miles	6.41	5.62	8.33
	<20miles	6.21	5.24	8.29
	5-20 miles	6.19	5.15	8.28
	20 to 100 miles	n/a	n/a	n/a
	>=20 miles	n/a	n/a	n/a
	>=50 miles	n/a	n/a	n/a
	>=100 miles	n/a	n/a	n/a
Rail	All distances	12.42	8.68	27.61
	<5 miles	5.88	6.53	n/a
	<20miles	6.98	6.45	10.11
	5-20 miles	7.02	6.44	10.19
	20 to 100 miles	13.00	8.06	17.55
	>=20 miles	14.55	9.02	28.99
	>=50 miles	18.41	9.53	32.56
	>=100 miles	n/a	10.01	n/a

Notes: Distance-weighted, Income option 1. n/a = insufficient NTS sample size (<300 records), VTT is imputed for PT trips with zero cost, Δt=10. Employers paying for EB trips. Tool version 1.1.

Table 7.19 presents VTT values by geography.

Table 7.19: VTT by geography (all modes, £/hr, perceived prices).

Mode	Area/Trip type	Commute	Employ-ees' business	Other non-work
All modes	All trips	11.21	18.23	5.12
	Urban: London to London	6.91	12.35	3.82
	Urban: Urban to urban <20 miles (excluding London)	8.65	7.85	2.99
	Non-urban: Long distance: London to other	14.90	24.66	8.67
	Non-urban: Urban to urban > =20 miles (excluding London)	13.53	20.39	7.00
	Non-urban: Rural to urban or urban (excluding London)	12.10	15.96	4.92
Car	All trips	11.70	16.74	4.91
	Urban: London to London	9.23	13.54	3.07
	Urban: Urban to urban <20 miles (excluding London)	9.36	7.89	2.91
	Non-urban: Long distance: London to other	14.32	19.27	7.89
	Non-urban: Urban to urban > =20 miles (excluding London)	13.94	19.79	7.00
	Non-urban: Rural to urban or urban (excluding London)	12.32	15.48	4.86
Bus	All trips	3.15	-	3.26
	Urban: London to London	2.86	-	3.15
	Urban: Urban to urban <20 miles (excluding London)	3.17	-	3.15
	Non-urban: Long distance: London to other	n/a	-	n/a
	Non-urban: Urban to urban > =20 miles (excluding London)	n/a	-	3.54
	Non-urban: Rural to urban or urban (excluding London)	n/a	-	3.82
'Other PT'	All trips	6.35	8.33	5.23
	Urban: London to London	6.32	8.04	5.45
	Urban: Urban to urban <20 miles (excluding London)	n/a	n/a	3.97
	Non-urban: Long distance: London to other	n/a	n/a	n/a
	Non-urban: Urban to urban > =20 miles (excluding London)	n/a	n/a	n/a
	Non-urban: Rural to urban or urban (excluding London)	n/a	n/a	n/a
Rail	All trips	12.42	27.61	8.68
	Urban: London to London	7.43	14.83	6.60
	Urban: Urban to urban <20 miles (excluding London)	5.82	n/a	6.27
	Non-urban: Long distance: London to other	15.65	29.74	10.15
	Non-urban: Urban to urban > =20 miles (excluding London)	10.67	n/a	7.59
	Non-urban: Rural to urban or urban (excluding London)	n/a	n/a	8.74

Notes: Distance-weighted, Income option 1. n/a = insufficient NTS sample size (<300 records), VTT is imputed for PT trips with zero cost, Δt=10. Employers paying for EB trips. Δt=10. Tool version 1.1.

The behavioural models estimated and implemented in the Tool also included controls for five types of geography:

i) London to London

ii) London to other

iii) urban to urban with distances <= 20miles

iv) urban to urban with distances => 20 miles, and

v) rural trips (other than to London).

In the main, these controls were not significant. With five area types there are four potential controls in each of the 11 mode/trip purpose models. Of the potential 44 geography controls across all the models, only four were statistically significant. Therefore, variations in VTT by geography are generally explained by variations in trip and socio-economic characteristics.

When one considers that London incomes are higher than elsewhere, and that intra-London trips and intra-urban (outside of London) trips are shorter than other trip types, one can see that the variation in VTT by geography type is dominated by the effects of distance and income. We can also see that trip length effects (that is, the way socio-economic and trip characteristics vary with trip length) are more relevant in determining VTT than a 'London effect' – whether that be from income or other factors unique to London.

This is in line with the 'implied' distance and income elasticities presented in **Table 7.15** earlier.

Commuting and non-work

Returning to the relative values for commuting and other non-work, we find that even after we neutralise the effect of income we still get large differences – although the level of disparity is somewhat reduced (see **Table 7.20** below). There also appears to be a lack of consistency between modes in terms of the ratios.

To analyse these changes further we have undertaken an equivalent trip/person analysis – looking at the VTT for the different mode/trip purpose combinations for the 'same trip/person'. For all the covariates required by the Tool, we have calculated the overall averages in the NTS dataset, and used these values to define a 'typical' trip. The results of this are presented in **Table 7.21** below. Here we see that, all else being equal, rail and 'other PT' commuting values are between 91% and 94% of other non-work, whilst bus commute values are about 70% of other non-work. Car commute on the other hand is valued at almost 3 times the other non-work value.

We have investigated this car commute effect further and have found that, relative to the 2.382 ratio in **Table 7.20**, the ratio in the SP sample is much lower at 1.4. We also note from **Table 7.21** that the ratio of VTT for employees' business to VTT for non-work is less than one for PT, but greater than one for car – probably reflecting the different productivities of travel time between modes.

Table 7.20: Impact of income options on the relative VTTs of commuting and other non-work

<table>
<tr><th rowspan="2">Mode</th><th rowspan="2">Purpose</th><th colspan="2">Ratio of commute VTT to other non-work VTT (SP1)</th></tr>
<tr><th>Income option 1</th><th>Income option 2</th></tr>
<tr><td rowspan="2">Car</td><td>Commuting</td><td rowspan="2">2.382</td><td rowspan="2">1.995</td></tr>
<tr><td>Other non-work</td></tr>
<tr><td rowspan="2">Rail</td><td>Commuting</td><td rowspan="2">1.432</td><td rowspan="2">1.222</td></tr>
<tr><td>Other non-work</td></tr>
<tr><td rowspan="2">Bus</td><td>Commuting</td><td rowspan="2">0.967</td><td rowspan="2">0.677</td></tr>
<tr><td>Other non-work</td></tr>
<tr><td rowspan="2">‘Other PT’</td><td>Commuting</td><td rowspan="2">1.213</td><td rowspan="2">1.168</td></tr>
<tr><td>Other non-work</td></tr>
</table>

Notes: Fixed income trip-weighted (option 2) is £49,684 for household (non-work) income. Ratio based on distance-weighted VTT. Modelling used imputation for zero reported PT costs, Δt=10

Table 7.21: Relative values by trip purpose (other non-work is the base)

	Commute	Employees’ business	Other non-work
Car	2.95	2.70	1.00
Bus	0.70	-	1.00
‘Other PT’	0.91	0.58	1.00
Rail	0.94	0.46	1.00

The implication of this analysis that the observed differences between VTTs by trip purpose (within mode) arise due to differences in both the model parameters and also – importantly – the covariates (the latter of course vary between different samples, e.g. between the SP and NTS samples in the present case). As has been acknowledged throughout this report, it was fully expected that our intercept survey would not collect a representative sample, hence the need to correct for representativeness *ex post* using NTS.

Comparing the new estimates of VTT reported here against the 2003 study values and meta-analysis evidence, the indication is that the other non-work value has decreased, whilst the commuting value has increased. This could be due to changing preferences and conditions of travel between the studies. We will draw further comparison with meta-analysis evidence in **Section 7.8**, when discussing changes in VTT over time.

7.6.5 Modal differences

The issues

Even with income option (2), where the effect of income has been neutralised, substantial differences by mode remain. For commute, VTTs for car and rail are more or less equal, though for other non-work the rail value is higher. Values for bus are, for both purposes, not more than 1/3 of the corresponding car and rail values, while for ‘other PT’, the values are more or less double those for bus. There are further issues as to whether the car values should be treated as relating to the driver only or to all the occupants of the vehicle.

Mackie *et al* (2003) in Sections 7.3 and 8.3 noted a number of reasons why values might vary by mode:

(i) The income and socio-economic characteristics of travellers might vary systematically by mode. Low income users with low average VTT might gravitate to mode A while high income users with high average VTT might tend to choose mode B.

(ii) The composition of trips and purposes might vary systematically by mode. Mode A might have a strong market share in short distance trips, while mode B might be stronger at longer distances[84].

(iii) A cross-section of people with given income and socio characteristics making a given trip will have a distribution of values of time (and individual values may vary according to the constraints faced). People with low VTT for that trip will self-select into relatively low cost/high time modes and vice versa.

(iv) For any individual, VTT by mode may vary due to the different characteristics of the modes in terms of comfort, cleanliness, reliability, level of personal control, and other quality attributes.

Mackie *et al* argued that, point (iv) aside, individuals should, from a theoretical perspective, have the same VTT for a given trip regardless of mode used, hence favouring an approach which picks up (i) to (iii) through the income, socio-economic characteristics and trip and purpose characteristics of the traffic modelled to the various sub-markets. Any remaining variation in VTT should then reflect ‘comfort’ effects (note that these include questions related to ‘time use’, i.e. the extent to which travel time can be used productively or otherwise enjoyably, as discussed in **Chapter 3** earlier). The ranking of the modal values in the meta-analysis reported in Mackie *et al*, was “essentially the opposite of the presumed ‘comfort’ effect”.

In the course of external review of this report, it has been put to us that item (iii), which might be described as the ‘self-selection’ effect, constitutes a valid reason (in addition to item (iv)) for segmenting the values by mode, and that aggregating valuations of time savings across modes could involve a risk of misallocation of public funds. But in our view the self-selection argument would need to be considerably expanded to make such a case. The relationship between the modes in terms of speed and cost per km is less clear-cut in practice than is implied by the standard description of rail as ‘fast/expensive’ and bus as ‘slow/cheap’. In any case, for the self-selection argument to work, there is the implication that, for a particular journey, alternative modes (faster and more expensive) actually exist. There are also quite a few practical difficulties of implementing the rule-of-a-half (RoH) formula consistently where VTTs vary across modes. Our view is thus that we should only take account of modal variation which unambiguously relates to ‘comfort’ effects.

As we have seen, income option (2) reduces the modal variation in VTT, but it does not remove all of the socio-economic variation or trip-related characteristics from consideration – only income, though this is likely to be the main effect (especially since, as we have seen, the distance effects are less prominent than in

[84] Given the much lower distance elasticities being found in the current study, this example has less force.

the 2003 model). To understand the modal variation better, we have also examined VTT for a ‘typical’ person/trip in the dataset by mode and trip purpose (as introduced above).

The variation in VTT by mode for such a typical trip is presented in **Table 7.22**, separately for each purpose. If the modal variation related solely to ‘comfort’, we would expect the lowest values for rail and the highest for bus, with car and ‘other PT’ intermediate. For commuting, we do find such an effect for car, rail and ‘other PT’, but the bus values are low rather than high. For business, where bus is not of relevance, there does appear to be some relation to comfort (especially in terms of the possibility of working on the train). For other non-work, bus values are much higher, in line with the ‘comfort’ hypothesis, but the values for rail and ‘other PT’ go in the opposite direction to what we would expect. All in all, we consider that these differences between modes certainly cannot be explained solely by comfort.

Table 7.22: Ratio of modal VTT by trip purpose for an average person (car = base)

	Commute	Employees’ business	Other non-work
Car	1.00	1.00	1.00
Bus	0.51	-	2.14
‘Other PT’	0.99	0.69	3.19
Rail	0.73	0.39	2.29

Non-work

Given the above considerations and evidence, and in line with the argument in Mackie *et al*, our preference for non-work is to retain mode-free values, by averaging the values over the sample of trips for all (motorised) modes, maintaining the distance weighting. If this is done, maintaining the approach of using the SP1 values, the values given in **Table 7.23** are obtained.

Table 7.23: All mode non-work VTT by treatment of income option (£/hr, perceived prices)

	Commute	Other non-work
Using income option (1)	11.21	5.12
Using income option (2)	10.03	5.49

Notes: Fixed income trip-weighted (option 2) is £49,684 for household (non-work) income. For personal income (EB) it is £35,070 for Car, £20,219 for Bus, £45,019 for ‘Other PT’, £55,319 for Rail. Ratio of distance-weighted VTT. Modelling used imputation for zero reported PT costs and employers paying for EB trips. Δt=10. Tool version 1.1.

It can be seen that the two income options do not produce very different results, but (as seen earlier), option (2) narrows the gap between commute and other non-work values. As might be expected, the car values dominate, given that well over 80% of distance travelled is by car (see **Table 7.24**).

Table 7.24: NTS car, bus, rail and 'other PT' mode split

	Mode split	
	Trips	Person-km
Car	87.0%	84.6%
Bus	8.6%	4.4%
Rail	2.8%	9.7%
'Other PT'	1.5%	1.3%
Total	100.0%	100.0%

Source: NTS 2010-2012

Employees' business

As discussed in the treatment of income section earlier (**Section 7.4**), for business trips we consider that the observed differences by mode reflect real differences. Correspondingly we recommend modal values based on income option (1). These have already been presented in earlier tables, but for completeness we present them again below for all distances (**Table 7.25**).

Table 7.25: Employees' business modal VTT (£/hr, perceived prices)

	Employees' business (all distances)
Car	16.74
Bus	-
'Other PT'	8.33
Rail	27.61

Notes: Distance-weighted, income option 1, based on sample enumeration from NTS 2010-2012; SP1 Δt=10; Tool version 1.1 (cost imputed for a PT trip with a zero cost and employers paying for EB trips).

Car VTT: individual and vehicle values

In the case of car, we have further to decide whether VTT relates to the driver only or whether he/she takes into account the wishes of the passengers. If the latter is the case, we would expect to find some effect on VTT related to the number of passengers. This has been a perennial problem for interpretation, and it is worth reproducing some of the text of the 1980s study (MVA *et al*, 1987: paragraphs 4.3.6-7).

"Whenever a driver is accompanied by passengers, any or all of the following may be true.

a. *The driver may 'share' out the cost of the journey among himself and the passengers (though not necessarily pro rata). In as far as he expresses his own preference in terms of time saving, this will tend to inflate his value of time. This we may interpret as 'the driver's value of time assuming passenger contributions'.*

b. *The driver may take into account, to a greater or lesser extent, the preferences of his passengers in terms of time saving, though he may not value their time as highly as they would. However, his expressions of preference would yield a* *value of time for the* ***vehicle****. This we may interpret as 'the vehicle's value of time as perceived by the driver'. It is likely to be an underestimate of the total vehicle value of time, though this very concept relies on a 'market' which in practice does not exist!*

c. *Regardless of whether he takes account of passenger time preferences, his own value of time may reflect the utility, or disutility, to him of having passengers in the car. This we may refer to as 'the driver's value of time reflecting the value of company'. The driver's value of time would be reduced in as far as the passengers were a source of comfort, stimulation etc., and increased in as far as they were not, as might be the case with younger children.*

There is no way in which we can discover the extent to which each of these separate factors is playing a part in the driver's response. All we can do is to examine the empirical evidence, vis-à-vis the unambiguous case of unaccompanied drivers".

In the 1980s study, one of the surveys suggested higher values (by up to 40%) for drivers with passengers, while another suggested (for commuters) slightly lower values. On balance the view was taken that *"the value for single-occupant vehicles [might be increased] by a maximum of 20% for each passenger."* However, this was not accepted by the Department, who chose to multiply the single-occupant value by the number of passengers to obtain a vehicle value. Essentially, this remains the current position. The analysis of the 1994 data suggested that passengers might have VTT lower by a factor of 20-25% but this was not considered sufficiently robust to be included in the recommendations by Mackie *et al* (2003). AHCG (1999) found no significant effects from occupancy for either business or commuting, but a reduction in VTT for other non-work.

The modelling results from the present study have revealed no effects from occupancy for the other non-work purpose, but for commuting they suggested a **reduction** of around one-third when there were passengers in the car. In line with the citation above, this implies a 'value of company'. Nevertheless, the proportion of car commuters who carry passengers is low (WebTAG suggests an average occupancy for commuting of about 1.15), so this is unlikely to play a major role.

On balance, while we still think there is some argument in favour of the 1980s recommendation, we must conclude that we have not produced sufficient evidence to undermine the current convention, which effectively assumes that the driver values are representative of all occupants in the car, so that they should be applied separately to each occupant. From this point of view, they may be considered on a compatible basis with the public transport VTTs, as referring to an individual. If there were to be further work on this issue, we would recommend focussing on the other non-work purpose where the vehicle occupancy effect on the final values is more significant.

7.6.6 Business values

Chapter 6 reconciled the results from the different strands of business VTT analysis. In the main, we see consistency between the two different SP analyses (employer and employee), and we also see some similarity between the values from the Cost Saving Approach (CSA) – which is the Department's current basis for estimating business VTT – and the SP-based values. This is particularly the case for blue collar workers, whom we would expect to have low productivities whilst travelling. Briefcase travellers, who may engage in work whilst travelling, seem to exhibit a lower VTT relative to the CSA.

Moreover, the broad similarity between the VTT from the SP sample and the CSA observed in **Chapter 6** is, as we will discuss below, in part a function of the trip length distribution within the SP sample, and does not hold all over all distances. This is to some extent expected, since **Section 4.8** and **Table 7.15** have highlighted income elasticities for business VTT which are significantly less than one.

Table 7.26: Business VTT by method of calculation, mode and distance (2014 perceived prices and values, £/hr)

Source/method	Distance	All modes	Car	Bus	'Other PT'	Rail
WebTAG (2014 prices and values)	All distances	25.47	24.43	15.64	24.72	30.07
CSA estimate from NTS 2010 to 2012 data (2014 prices and values)	All distances	28.27	27.05	13.13	26.33	36.46
Employees' business SP re-weighted to NTS 2010-2012 (2014 prices and values)	All distances	18.23	16.74	-	8.33	27.61
	<5 miles	5.39	5.27	-	8.33	n/a
	5-20 miles	8.84	8.79	-	8.28	10.19
	>=20 miles	21.14	19.51	-	n/a	28.99
	>=50 miles	24.55	22.53	-	n/a	32.56
	>=100 miles	28.62	25.74	-	n/a	n/a

Notes: All modes. Distance-weighted, income option 1, SP1 Δt=10; Tool version 1.1. PT cost is imputed for a trip with a zero cost, and employers paying for EB trips. WebTAG 'Other PT' is Underground passenger, WebTAG car EB is weighted average of driver and passenger (vehicle occupancy of 1.2)

As mentioned earlier in this chapter, the behavioural model implemented in the Tool is the employees' business SP model. The Tool then applies this model to business trips in the NTS to derive an average value over a specified segmentation.

The average distance-weighted personal income across the NTS is £46,615 (2014 prices and values). This would give a business VTT in 2014 perceived prices of £28.27 using the CSA – second row, third column of **Table 7.26**.

This compares to the CSA-based WebTAG values which have an all modes value of £25.47 (first row, third column).

If we compare these values to the SP-based values re-weighted for NTS (third row), then we see that the VTT for employees' business across all modes is £18.23. This is 72% of the WebTAG value, considerably lower than the proportion evidenced from the sample enumeration of the SP sample in **Chapter 6**. We also find substantial variation by mode with 'other PT' lowest at £8.33 and rail highest at £27.61 for the all distance values. As proportions of the WebTAG values, these range from 34% ('other PT') to 92% (rail).

We have already seen that the SP-based VTTs for business are sensitive to trip distance (see **Table 7.15** and **Table 7.18**). From **Table 7.26** we can also see that at low distances the SP-based values are substantially less than the current WebTAG values, but as trip distances increase the SP-based values increase such that they are close to or exceed the current WebTAG values at long distances (>50 miles). It is the long average distance of the SP sample therefore that gives rise to similarities in the business VTT for the SP sample relative to the CSA (see **Section 6.8**).

From a behavioural perspective, this variation with distance implies that the marginal time spent travelling on short distance trips is similarly productive (or unproductive as the case may be) as its alternative uses – this is the reason why the VTT is low. By contrast, the marginal time travelling on a long distance trip is quite unproductive relative to its alternative uses. We speculate that part of the reason for this is the existence of ‘slack’ time between business engagements that occur naturally in the course of business trips. These will have a disproportionate impact on short distance trips relative to long distance trips.

With an income elasticity much less than 1, we primarily attribute these differences between modes to differences in average modal trip lengths and the manner in which socio-economic and trip properties are correlated with trip length. Trip length distribution also explains the similarity between the CSA and SP-based values when no correction is made for NTS (**Chapter 6)** and the differences that emerge once the NTS correction is applied (**Table 7.26**). This is because the SP sample has a long average trip length – corresponding to the upper limits of the distance range presented in **Table 7.18**.

Indeed, the discrepancy between the SP-based and CSA values in **Table 7.26** is in line with other evidence on employee valuations. Wardman *et al* (2013, p49/50) in their meta-analysis found that SP employee valuations are approximately 64% of CSA values. Abrantes and Wardman (2011) found that employee valuations are approximately double other non-work. In our results, the employee valuations are three times larger than other non-work – but as noted earlier, other non-work VTT has seemingly fallen relative to current WebTAG values.

As discussed at the beginning of **Section 6.5**, the SP questionnaire unambiguously elicited a personal valuation of a travel time saving (though during working hours) from some respondents and for the majority of respondents unambiguously elicited the value to the business of a travel time saving (subject to the interviewee being able to respond appropriately). From an appraisal perspective, we require the **business’s** valuation (that is, the latter valuation): therefore, in applying the employees’ business model in the Tool we neutralised the effect of those who were unambiguously responding in a personal capacity. This is what has been done in all of the values presented in this section. Clearly, including those who have responded in a personal capacity lowers the business value of time compared to those previously presented. This can be seen in **Table 7.27** below, where if we include respondents for whom we elicited a personal valuation, the average business VTT decreases by between 15% (‘other PT’) and 24% (rail).

Table 7.27: Comparison of business VTTs based on employer valuations only and including those including personal valuations (2014 perceived prices, £/hr)

	Employer valuations only	Includes personal valuations
Car	16.74	13.76
'Other PT'	8.33	7.06
Rail	27.61	21.08
All	18.23	14.76

Notes: Distance-weighted, income option 1; Δt=10; Tool version 1.1 (fuel costs imputed).

7.6.7 VTT values and multipliers for use in appraisal

VTT values

Bringing the discussions presented in this chapter together, we can draw out some recommendations regarding the basis for VTT values for use in appraisal:

- VTT should continue to be distance-weighted, but should be disaggregated into distance bands to reduce the level of approximation between the standard VTT values and the 'real' scheme level VTT value. We note that further work is required to determine appropriate distance bands for use in appraisal.
- We should continue to distinguish between Level 1 appraisals of small and medium sized schemes, and Level 2 and 3 appraisals of major schemes and policies and significant 'user pays' initiatives. For a Level 1 appraisal, standard 'national' values of time can be used. For Level 2, the values may be amended to more accurately reflect local conditions. For Level 3, appraisal values derived from bespoke quality surveys would be appropriate.

For Level 1 and 2 appraisals, we make the additional recommendations:

- **Behavioural model:** Recommended VTT values should be based on SP1 with Δt = 10. In the longer term, SP3 values could be used but further research is necessary to utilise them.
- **Trip purposes:** There are differences between trip purposes and we should therefore continue to disaggregate VTT by trip purpose.
- **Modal values:** For non-work trips we should use an all modes value due to the non-work VTTs reflecting some self-selectivity between modes, but for business trips we should use modal values as we interpret differences between modes to be real differences.
- **Treatment of income:**
  - Non-work: For Level 1 appraisals with VTTs distance-banded, as recommended above, we should use income option (2) – that is, treating all non-work trips as having the same average household income (if however distance-banding is not implemented – at least in the short term – then income option (1) should instead be used). For Level 2 appraisals, we should use income option (1) applied at the appropriate regional level.
  - Business: For Level 1 we should use income option (1) using national data, whilst for Level 2 we should use income option (1) applied at the appropriate regional level.

We have illustrated these recommendations for a Level 1 appraisal in **Table 7.28**. At this stage it is only an illustration, as further research is needed to determine the distance bands that should be used. This table also presents, for the purpose of comparison, the existing WebTAG values converted to a comparable base (2014 perceived prices).

Table 7.28: Appraisal VTTs for a Level 1 appraisal (routine appraisal of small and medium sized schemes) with illustrative distance bands (2014 perceived prices, £/hr)

<table>
<tr><th rowspan="2">Mode</th><th rowspan="2">Distance</th><th>Commute</th><th>Other non-work</th><th colspan="5">Employees' business</th></tr>
<tr><th>All modes</th><th>All modes</th><th>All modes</th><th>Car</th><th>Bus</th><th>'Other PT'</th><th>Rail</th></tr>
<tr><td>WebTAG (2014 prices and values)</td><td>All</td><td>7.62</td><td>6.77</td><td>25.47</td><td>24.43</td><td>15.64</td><td>24.72</td><td>30.07</td></tr>
<tr><td rowspan="4"></td><td>All</td><td>11.21</td><td>5.12</td><td>18.23</td><td>16.74</td><td>-</td><td rowspan="4">8.33</td><td>27.61</td></tr>
<tr><td><20 miles</td><td>8.27</td><td>3.62</td><td>8.31</td><td>8.21</td><td>-</td><td>10.11</td></tr>
<tr><td>20 to 100 miles</td><td rowspan="2">12.15</td><td>6.49</td><td>16.05</td><td>15.85</td><td>-</td><td rowspan="2">28.99</td></tr>
<tr><td>>= 100 miles</td><td>9.27</td><td>28.62</td><td>25.74</td><td>-</td></tr>
</table>

Notes: Distance weighted, 'all distance' values based on income option 1, for distance-banded values: non-work based on income option 2 (household income = £49,684) and business on income option 1, VTT imputed for PT trips with zero cost, SP1 VTTs, $\Delta t=10$, employers paying for EB trips, Tool version 1.1.

VTT multipliers

In addition to the overall values, we also make some recommendations for adjustments for different types of time, and we present these as multipliers. In doing this, we have to have some consideration for the different averages coming from the different games, and our general approach of using SP1 values for the overall recommendations about VTT.

In the case of reliability (SP2), we think it is preferable to maintain the internal consistency of the valuations, and apply the implied ratios to the average values based on SP1. So, for example, the reliability ratio for car is taken from the ratio of the 'Value of sd travel time' to the value of 'Average travel time' from SP2, but would then be used (in order to get an absolute valuation of the standard deviation) by multiplying by the recommended value from **Table 7.5**, **Table 7.6**, **Table 7.7** and **Table 7.8**. The same approach is taken for the early and late multipliers. The fact that the SP2 VTTs are rather higher than those for SP1 in the case of car and rail does mean, of course, that the implied valuations of reliability will be lower. Without a clear understanding of the reason for the difference between SP1 and SP2 VTTs, this must remain an arbitrary judgment.

In addition to the 'headline' VTT, recommendations on the values of reliability were an important element in the ITT. The remaining multipliers may be considered of lesser importance in terms of the study output. More difficulties do indeed arise in the case of the SP3 effects.

For the public transport modes, we have chosen to align the results with the level of crowding closest to the SP1 value: for bus and other public transport this corresponds to the level *"a few seats free but had to sit next to someone/could not* *sit with people travelling with. Some standing"*, while for rail it corresponds to a load factor of 100% (i.e. *"all seats taken but no standing"*). Thus, in **Table 7.29** below, these multipliers are set at 1.0. With somewhat less justification, we have also used this base for the additional multipliers for the bus mode which derive from SP3.

Currently recommendations for most of these multipliers derive from PDFH research, and there is an issue as to how the work reported here might be integrated into WebTAG guidance. Our inclination is to allow these results to be added to the general corpus of research, rather than attempting to supplant the existing guidance.

In addition, on waiting time (headway), existing guidance (both the wait time multiplier of 2.5 in WebTAG and the approach to headway (without giving the penalties) in PDFH) is supported by other evidence (such as the meta-analysis review for the International Transport Forum (2015))[85]. Again, it seems preferable to use the findings of this study to supplement the existing corpus of evidence.

The more serious problems relate to the car mode, where we had previously queried the range of the multipliers. We note that in all cases the SP1 values fall inside the range between the light and heavy traffic values from SP3, though it is difficult to justify this apart from perhaps the commuting case. Since in any case we only consider that these ratios should be treated as indicative of the possible impact of congestion, we have divided the SP3 values for the three levels by the SP1 value to get the multipliers in **Table 7.29**.

Table 7.29: VTT multipliers

		Commute	Employees' business	Other non-work
Car	Reliability ratio	0.33	0.42	0.35
	Free-flow	0.51	0.42	0.47
	Light traffic	0.72	0.68	0.83
	Heavy traffic	1.37	1.26	1.89
Bus	Value of early	-2.69	-	-3.20
	Value of late	2.88	-	2.52
	Plenty of seats free and did not have to sit next to anyone.	0.85	-	0.83
	A few seats free but had to sit next to someone/could not sit with people travelling with.	0.89	-	0.84
	A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	1.00	-	1.00
	No seats free – a few others standing.	1.24	-	1.30
	No seats free – densely packed.	2.14	-	2.32

85 International Transport Forum (2015) 'Valuing Convenience in Public Transport'. Roundtable report 156. OECD, Paris, France.

		Commute	Employees' business	Other non-work
	Value of free-flow	0.99	-	1.22
	Value of slow down	1.39	-	1.36
	Value of dwell time	0.68	-	1.57
	Value of headway	1.68	-	1.60
'Other PT'	Value of early	-2.40	-1.66	-2.98
	Value of late	1.75	1.95	2.24
	Plenty of seats free and did not have to sit next to anyone.	0.95	1.00	1.00
	A few seats free but had to sit next to someone/could not sit with people travelling with.	0.97	1.00	1.00
	A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	1.00	1.00	1.00
	No seats free – a few others standing.	1.13	1.17	1.10
	No seats free – densely packed.	1.70	1.78	1.87
Rail	Value of Early	-1.77	-1.55	-2.34
	Value of Late	2.86	2.76	3.21
	seated 50% load	0.73	0.75	0.72
	seated 75% load	0.79	0.76	0.72
	seated 100% load	1.00	1.00	1.00
	seated 1 pass per m2	1.09	1.13	1.14
	seated 3 pass per m2	1.31	1.36	1.39
	standing 0.5 pass per m2	1.16	1.29	1.21
	standing 1 pass per m2	1.19	1.38	1.27
	standing 2 pass per m2	1.32	1.56	1.57
	standing 3 pass per m2	1.57	1.61	1.79
	standing 4 pass per m2	1.86	2.03	2.17

Note: SP2 VTT taken as base for reliability and early/lateness. SP1 VTT taken as base for car free-flow, light and heavy traffic. SP1 VTT taken to represent 100% occupancy of seats for PT.

International comparisons

With regard to the **range** of values presented, we have compared our recommendations with the most recent European studies. We have not attempted to provide a numerical comparison of the values – to do this in a meaningful way would require a detailed set of assumptions about the variations in currency and purchasing power, modal definitions and conditions (such as the level of congestion), and ensuring compatibility of purposes. We note that some results from previous studies are available in Mackie *et al* (2014)[86].

The Danish study[87] excluded business travel and for non-business travel concludes that *"Since differences between purposes are generally small there is little point in distinguishing between travel purposes*". On mode, it was noted that the observed differences in VTT between modes were the opposite of what would be expected based on the comfort of the mode (thus car values were higher than those for bus, even after controlling for income: rail values were intermediate), and argued that self-selection was the more likely explanation. On this basis, it was decided to use the overall average as the central value to be applied to all transport modes. A sample enumeration process was carried out using National Travel Survey data, with the intention that the average VTT be representative for the average kilometre travelled. In addition, the sample enumeration was carried out as if all segments had the same average income (corresponding with our option (2)).

In addition to the 'headline' value of in-vehicle time, the Danish study also provided multipliers for congested time, parking search time, access/egress, interchange waiting time (as well as a 'penalty' for interchange) and headway.

The Swedish study (see Börjesson and Eliasson, 2014[88]) also excluded business travel, but in contrast to the Danish study, made a distinction between commute and other non-work. In addition to this, a distinction is made between short and long distance travel: the boundary between these two was set at 100 km, though the purpose distinction is only maintained for 'short distance' trips. Separate values are given for the modes car, bus and train.

The modal variation has a consistent pattern by purpose and distance with car values being the highest and bus the lowest. This is so even when income option (2) is applied (the VTT model is evaluated at the 'mean income of the sample': it seems that the sample has been re-weighted by trip distance, though this is not quite clear). Thus, the outcome is comparable to that from the Danish study, where it was decided not to recommend separate values by mode.

The Dutch study (Significance, VU Amsterdam and John Bates Services, 2013[89]) presents mode-specific values (car, rail and 'other PT', as well as air) for three purposes (commute, business and other), though air values are not given for commute. The business values make use of the Hensher equation, and combine employee and employer values. The model values have been applied to the relevant samples of the Dutch National Travel Survey: as far as income treatment is concerned, this is in line with our income option (1).

---

[86] Mackie, P.J., T, Worsley and J. Eliasson (2014) 'Transport appraisal revisited'. Research in Transportation Economics 47 (2014) pp3-18.

[87] Fosgerau, M., K. Hjorth, and S. Vincent Lyk-Jensen (2007) 'The Danish Value of Time Study: Final Report: Report 5', Danish Transport Research Institute.

[88] Börjesson, M. and Eliasson, J. (2014) 'Experiences from the Swedish Value of Time study', Transportation Research Part A, 59, pp144–158.

[89] Significance, VU Amsterdam, and John Bates Services (2013) 'Values of time and reliability in passenger and freight transport in The Netherlands', Report for the Ministry of Infrastructure and the Environment.

In respect of the modal variation (disregarding air), there is little difference among the ‘other’ purpose values, but for commute, train values are the highest, and ‘other PT’ are the lowest, while for business car is the highest, with similar values for rail and ‘other PT’. The high business values for car appear to be driven by the employer’s contribution, reflecting the limited possibilities for productive use.

7.7 Variance and uncertainty around the results

The Implementation Tool has been developed in such a way that it outputs more than just the mean VTT. It also outputs standard errors of the estimates and confidence intervals. We present these below as a demonstration of the robustness of the VTT estimates. These 95% confidence intervals have been calculated for income option (1), distance-weighted, with costs for trips with a zero VTT imputed.

The 95% confidence intervals for all VTTs are presented in **Table 7.30**; it can be seen from this table that, broadly speaking, the models and associated VTTs are well-estimated. In general, the main headline VTT SP1 confidence intervals are just below +/-30% of the mean. We can also see that rail and ‘other PT’ have the tightest confidence intervals, whilst bus has the loosest. An exception is other non-work for car, where the confidence intervals are around +/-70%. Another observation is that the SP1 VTTs have slightly tighter confidence intervals (<30%) than SP2 and SP3. All of the valuations are significantly different from zero, apart from the value of bus dwell time.

Table 7.30: VTT 95% confidence intervals

			Commute	Employees’ business	Other non-work
Car	VTT	SP1	33.0%	21.5%	70.1%
	Average travel time	SP2	28.1%	38.9%	69.4%
	Value of sd travel time	SP2sd	34.3%	34.9%	71.0%
	Free-flow	SP3ff	41.5%	36.9%	77.6%
	Light traffic	SP3lc	36.6%	28.5%	71.1%
	Heavy traffic	SP3hc	34.9%	25.8%	68.9%
Bus	VTT	SP1	28.5%	-	24.7%
	Scheduled travel time	SP2	39.2%	-	33.6%
	Early	SP2early	-43.4%	-	-42.7%
	Value of late	SP2late	41.3%	-	36.7%
	Plenty of seats free and did not have to sit next to anyone.	SP3c1	46.7%	-	45.2%
	A few seats free but had to sit next to someone/could not sit with people travelling with.	SP3c2	46.4%	-	50.0%
	A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	SP3c3	43.4%	-	42.1%
	No seats free – a few others standing.	SP3c4	41.6%	-	41.3%

			Commute	Employees' business	Other non-work
	No seats free – densely packed.	SP3c5	38.0%	-	40.2%
	Value of free-flow	SP3ff	56.5%	-	58.3%
	Value of slow down	SP3sd	55.1%	-	52.6%
	Value of dwell time	SP3dw	173.9%	-	72.1%
	Value of headway	SP3hw	44.7%	-	41.9%
'Other PT'	VTT	SP1	27.0%	31.8%	20.7%
	Scheduled travel time	SP2	24.3%	28.6%	22.8%
	Early	SP2early	-29.0%	-51.8%	-27.7%
	Late	SP2late	27.6%	33.9%	27.8%
	Plenty of seats free and did not have to sit next to anyone.	SP3c1	27.1%	36.2%	27.0%
	A few seats free but had to sit next to someone/could not sit with people travelling with.	SP3c2	26.4%	36.2%	27.0%
	A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	SP3c3	27.4%	35.9%	27.0%
	No seats free – a few others standing.	SP3c4	25.8%	37.0%	27.3%
	No seats free – densely packed.	SP3c5	25.2%	34.1%	26.9%
Rail	VTT	SP1	14.2%	17.0%	14.0%
	Scheduled travel time	SP2	20.3%	26.5%	24.3%
	Early	SP2early	-38.3%	-42.9%	-36.3%
	Late	SP2late	28.7%	33.3%	23.3%
	seated 50% load	SP3c1	30.6%	32.4%	31.8%
	seated 75% load	SP3c2	28.9%	32.6%	31.8%
	seated 100% load	SP3c3	25.3%	28.5%	27.4%
	seated 1 pass per m2	SP3c4	24.2%	27.6%	26.2%
	seated 3 pass per m2	SP3c5	22.8%	25.4%	24.2%
	standing 0.5 pass per m2	SP3c6	25.8%	28.8%	28.0%
	standing 1 pass per m2	SP3c7	26.0%	28.8%	28.8%
	standing 2 pass per m2	SP3c8	26.1%	29.6%	26.7%
	standing 3 pass per m2	SP3c9	24.5%	31.7%	26.5%
	standing 4 pass per m2	SP3c10	23.2%	29.8%	24.9%

Notes: Distance-weighted, income option 1; Δt=10; Tool version 1.1 (fuel costs imputed and employers paying for EB trips).

There are potentially three factors that influence how narrow or wide the confidence intervals are, and how they vary by mode and purpose.

- **Differences in the estimation errors in the behavioural models by mode and purpose**

One of the primary reasons for confidence intervals to vary by mode/purpose combination is that a unique behavioural model has been estimated for each combination. The samples over which these models have been estimated vary in size and thereby directly affect the accuracy of the associated parameter estimates. This and the consistency by which respondents responded to the choice scenarios has a direct impact on the variance-covariance matrix Ψ which feeds into the Delta method for each mode-purpose.

Variations in standard errors between SP games relate solely to the accuracy by which the dBF and game-specific terms are estimated. For example, SP2 'early' for rail has a relatively low t-statistic which results in relatively wide confidence intervals. A related explanation for the variation in confidence intervals between SP games is that SP1 is solely concerned with time-cost trade-offs whereas SP2 and SP3 include additional trade-offs, making it harder to accurately identify the VTT from a statistical perspective.

Then there is the vector of first derivatives Φ, which account for the impact of parameters and associated covariates on overall uncertainty in the actual VTT value. Since all first-order derivatives are scaled by VTT, standard errors will increase with VTT (see **Appendix G**).

Those issues aside, it is difficult to identify why and how standard errors vary across the sub-samples due to covariates. The reason for this is that Ψ contains both positive and negative off-diagonal elements, and therefore the impact of covariates on Ψ can have both positive and negative impacts on the standard error due to correlation structures between covariates. Of all the mode/purpose combinations, car other non-work has the largest confidence intervals – and this is also the sub-sample where we find a large discrepancy in terms of the net effect of the covariates on mean VTT between the SP-estimation sample and NTS sample.

- **Differences in the representativeness of the NTS data by mode and purpose leading to different bootstrapping errors**

We have separated out the respective contributions to the overall standard error from the behavioural model and from the NTS. The contribution to the standard error from the NTS sample is marginal relative to that from parameter uncertainty in the model.

- **Differences in sample size affecting both of the above**

As we use a weighted average across records to get to Φ, larger sample sizes reduce the standard errors and therefore the confidence intervals.

With further regard to the sampling error in the NTS, discussions with the Department revealed that for NTS use they recommend a minimum sample size of 300 per segment, and ideally 1,000. For certain segments with high levels of disaggregation, it is quite easy to drop below the 300 threshold. Such segments include 'other PT', long distance trips, business, London to other, and urban to urban (>20miles). In terms of the VTTs presented in the previous section, all segment sizes that do not exceed 300 have been suppressed. Therefore further disaggregation of the above VTTs when using the Tool may not always be possible.

These confidence intervals are larger than those found by Wheat, Wardman and Bates (2012)[90] for the 2003 VTT model. The reason for this discrepancy is that Wheat *et al* used a simple MNL-base model with few covariates and thereby obtained an accurate, but potentially biased VTT estimate in the base year. The models presented here are more flexible in terms of preference heterogeneity (random parameters and additional covariates) and functional form (dBF parameters). These model advances come at the cost of less accurate VTT estimates.

As in their work, we would expect the confidence intervals to increase over time, as there is uncertainty in both the temporal VTT elasticity to GDP/capita growth and in the projected GDP/capita growth rate itself (see the following section, **Section 7.8**, for a discussion on changing VTT over time).

7.8 Changing VTT over time

Current WebTAG guidance advocates that both non-work and business VTTs should grow in line with GDP/capita. That is to say, a unit elasticity with respect to GDP/capita is applied to both non-work and business VTTs. This is a return to the position prior to Mackie *et al* (2003)[91], who for non-work VTT recommended an elasticity of 0.8.

Non-work time VTT

The theoretical considerations relating to how the non-work VTT might change over time were set out at some detail in section 6.5 of Mackie *et al* (2003), which in turn summarised the greater detail of WP566 (Wardman (2001)[92]). It noted that Beesley (1978)[93] had pointed out various sources of variation in the value of time over time as well as the uncertainty as to even the direction in which the values might vary, but that the first British national value of time study (MVA *et al*, 1987)[94], faced with apparent evidence that there did seem to have been an increase in the value of time over time, concluded that it was not possible to come to any firm conclusions. They recommended that the matter should remain on the agenda for further investigation.

Essentially, while increasing incomes are likely to make travellers less sensitive to variations in money costs, so that we can expect a positive (though not necessarily unit) elasticity to income, the disutility of travel can be expected to fall over time as quality, comfort and facilities improve (though congestion may act against

---

[90] Wheat, P., M.R. Wardman and J. Bates (2012) ‘Advice on Statistical Confidence of Appraisal Non-Work Values of Time’. Report to the Department for Transport.

[91] Mackie, P.J., Wardman, M.R., Fowkes, A.S., Whelan, G.A., Nellthorp, J., and Bates, J.J. (2003) ‘Value of Travel Time Savings in the UK. Report to Department for Transport’. Available at: http://eprints.whiterose.ac.uk/2079/2/Value_of_travel_time_savings_in_the_UK_protected.pdf

[92] Wardman, M (2001) ‘A Review of British Evidence on Time and Service Quality Valuations’. Transportation Research E, 37, pp107-128.

[93] Beesley, M.E. (1978) ‘Values of Time, Modal Split and Forecasting’. Ch21 (part 7) in Hensher, D.A. and Stopher, P.R. (eds), ‘Behavioural Travel Modelling’, Croom Helm.

[94] MVA, ITS and TSU (1987) ‘The Value of Travel Time Savings’. Policy Journals, 1987.

this). The overall trend over time in VTT will be the combination of the two effects.

Sections 6.6 and 6.7 of Mackie *et al* (2003) presented the case for, and the results of an extensive meta-analysis of VTT findings. A conclusion was that: *"The estimated elasticity, in the range 0.72 to 0.82 might be broadly consistent with an inter-temporal income elasticity of around unity and a negative time trend. Using such a range to predict future changes in the value of time requires the assumption that any downward trend in the marginal disutility of travel time is maintained into the future."* (p55). Nonetheless, it was noted that the two available 'repeat studies' (in the national Netherlands 1988 and 1998 studies[95] and for the Tyne Crossing in UK) both suggested a **reduction** over time. It was also noted that the results from the meta-analysis were higher than those obtained from cross-sectional analysis (though with the proviso that *"cross-sectional elasticities are not necessarily appropriate to variations in the value of time over time"*).

Mackie *et al* (2003) concluded: *"..that the evidence as a whole tends to support an intertemporal elasticity for non-working time of somewhat less than unity, probably in the range of 0.5 to 1."* (p69). In the event, a value of 0.8 was recommended and accepted by the Department. The elasticity for working time, where – of course – a different methodology (the Cost Saving Approach, or CSA) was used, remained at 1.0.

At the end of their report, they recommend: *"further targeted research on the following issues:*

- *variation in the marginal utility of time and cost with respect to the levels of time and cost, so as to provide a more secure foundation for variable VTTS with trip length. A mixture of RP, SP and experimental economics approaches may be useful;*
- *values of the non-time attributes of travel (comfort, security, information etc.). In principle we would like to see these introduced into mainstream cost-benefit analysis especially of public transport. Such values will need to be based securely against values of time;*
- *variations in VTTS between driver/passenger and for larger groups. We have found in this piece of work that larger groups should probably be assigned lower VTTS per person than solo drivers;*
- *the value of savings in congested time and in changes in reliability are increasingly important issues not considered in this report.*

We believe that each of the above could significantly affect the relative worth of different policies and projects and therefore merit pursuing further. More generally, VTTS remains a key parameter in transport modelling and appraisal, and its relationship with trip purpose, trip length and income, both cross-sectionally and particularly over time, need to be regularly revisited through review work, meta analysis and further bespoke studies."

In 2013, the Department decided to update their VTT appraisal values in line with the prevailing methodology, and to this end commissioned a Peer Review and

[95] Gunn, H. F. (2001) 'Spatial and Temporal Transferability of Relationships between Travel Demand, Trip Cost and Travel Time'. Transportation Research E, 37, pp163-190.

Audit of their updating work. This was carried out by Laird *et al* (2013)[96]. In addition to the calculations relating to the level of VTT, the Department had proposed an increase in the income elasticity for non-working time to 1.0, effectively putting it on the same basis as the growth in working time VTT. The basis for this was the updated meta-analysis in papers by Wardman and Wheat (2013)[97] and Abrantes and Wardman (2011). Laird *et al* concluded that: *"This suggests that the additional studies added to the meta-analysis data set since the 2004 work, from which the value of 0.82 was derived, has led to an increase to 1.04. The methodology appears to be comparable in both studies."*

In analysing the results further, it was noted that most emphasis had been placed on the car values, since these underlay the AHCG work on which the Mackie *et al* and subsequent Department methodology were based. It should be noted, of course, that the current study does not have this restriction. Wardman and Wheat also noted: *"Car travel, and indeed travel by other modes, will have become more comfortable over time with more opportunities to spend travel time usefully, and this would be expected to exert a downward influence on values of time over time, but the evidence does not support this. Presumably there are also countervailing influences which act to increase the value of time over time, such as more difficult and crowded travelling conditions. This suggests that either these factors are cancelling each other out or the GDP elasticity represents the net impact of a number of correlated factors, such as income growth, comfort and travel conditions."*

There was a strong suggestion that it was the data since 2003 which was driving the higher elasticity: over the period 1994 to 2003 there was in fact no indication of a relationship with GDP at all. In summary, Laird *et al* noted that this apparent instability implied reservations about the representativeness of the meta data (modes, purposes, trip length, urban/rural, London/non-London) to the market as a whole. They suggested that a comprehensive review of the meta-analysis value would be of value. They also remarked:

"We have given less attention to the level of the values, concentrating on the growth over time. There is an indication from Tables 7 and 8 of W[ardman] and W[heat] that for all modes combined, the official values have been on the high side, compared with the studies in the meta-analysis. This suggests that a re-assessment of non-car modes might be in order: as noted at the outset, the evidence at the time of the ITS/Bates report for modal variation in non-work VTTS was not particularly strong, but this might no longer be the case."

On this, we may comment that while we have been asked in the ITT for the present study to make *"recommendations for assumptions of how values increase over time"*, we have not been asked to carry out a comprehensive review of the (meta-analysis) values, nor to carry out new meta-analysis. But in relation to the

---

[96] Laird, J.J., Bates, J. and Mackie, P.J. (2013) 'Peer Review of Proposals for Updated Values of Travel Time Savings. Report to Department for Transport'. Report dated 20th October 2013. https://www.gov.uk/government/publications/peer-review-of-proposals-for-updated-values-of-travel-time-savings

[97] Wardman, M. and Wheat, P. (2013) 'Meta-analysis of post-1994 Values of Non-Work Travel Time Savings'. Prepared for the Department for Transport. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/252088/meta-analysis-vtts-dft-011.pdf

previous paragraph cited, we are now in a position to give more attention to the **level** of the (non-working) values.

The current study has made a sustained effort to derive separate values for all modes. In doing so, we had expected that some of the obvious modal differences relating to the potential use of time spent travelling might have emerged. While the crowding and congestion effects may be interpreted in this way, the general level of VTT between modes does not bear this out. Further, these differences remain when making allowance for the different income distributions appertaining to different mode users. Hence, we may certainly conclude that the differences in VTT by mode reflect factors other than the possibilities for time use.

On this basis, we do not feel comfortable in doing anything more than to recommend proceeding on the existing basis for non-work VTT but carrying out a review as soon as is practical. This would review the results of this study, and how they fit in with the meta-analysis data. It should also take into account further evidence on temporal stability from the recent studies in Denmark, Norway, Sweden and the Netherlands.

Business VTT

The peer review by Laird *et al* also considered changes in business VTT over time. As has already been noted above, the economic model underpinning current guidance on business VTT is the Cost Savings Approach (CSA). Such a model indicates that VTT should grow in line with the value of the marginal product of labour (VMPL). With competitive labour markets, the wage plus on-costs is regarded as a good approximation to the VMPL[98]. For appraisal purposes we therefore need to forecast how wages and on-costs will change over the 60 year appraisal period. Such forecasts are not however available, though forecasts in GDP/capita are available. Looking at historic changes we can see two things.

- On-costs as a proportion of wages have fallen over time. In the March 2001 TEN on-costs were cited as 24.1% of wages, in the 2000 Labour Cost Survey they were 21.2% of wages, and in the 2008 Labour Cost Survey they were 18.5% of wages (DfT, 2007; 2014)[99]. Further research would be needed to determine whether this trend has arisen because real wages have grown faster than real on-costs or whether real on-costs in absolute terms have decreased.
- Real wage growth is slightly higher than real growth in GDP/capita. ONS (2012)[100] indicated that the median real wage has increased by 62% over the 25 year period from 1986 to 2011. The higher income workers have also

---

[98] In a review of the evidence for labour market imperfections and their implication on business VTT Wardman *et al* (2013 pp79-80) considered that whilst there is some evidence that labour market imperfections for certain segments of the labour market it would seem that on the basis of what limited evidence there is the gross wage is a reasonable approximation for the marginal product of labour for business travellers. They take this view because “firstly a lot of business travel is conducted by ‘mobile’ workers, and secondly that what evidence there is suggests the rents accrue to workers”. The evidence base on this remains weak and further research is needed.

[99] DfT (2007) Values of Time and Operating Costs. TAG Unit 3.5.6. February 2007; DfT (2014) User and Provider Impacts. TAG Unit A1.3. January 2014.

[100] ONS (2012) Real wages up 62% on average over the past 25 years. Report dated 7th November 2012. http://www.ons.gov.uk/ons/dcp171776_286266.pdf

experienced the highest increases in real wages. Thus those earning wages at the 90th percentile have experienced a wage growth of 81%, whilst those at the 10th percentile have expressed a growth of 47%. As business/briefcase travellers (as opposed to professional drivers[101]) on average have higher incomes than the average worker, we would therefore expect the average wage of the business/briefcase traveller to have increased by more than 62% in real terms. Over the same period real growth in GDP/capita has been 57%[102].

Looking back over the last 25 years, on balance it appears that growth in GDP/capita is likely to be not too far removed from growth in the VMPL as measured by wages plus on-costs – it is likely to be slightly high for the professional driver and slightly low for the business/briefcase traveller. Looking forward over a 60 year time horizon therefore, in our view the best available estimate of the growth in VMPL is growth in GDP/capita – which is consistent with current WebTAG guidance.

The 2013 update by the Department to the VTT, using the 2008-2010 NTS and hours data from the Labour Force Survey (LFS) among other data sources, led to a decrease in the average business VTT of approximately 21%. The previous values had been estimated from 1999/2000 NTS data uprated to 2010 prices and values with a unit elasticity to GDP/capita. The Department analysis reviewed and reported by Laird *et al* found that half of the 21% reduction came about due to differences in travel patterns between 1999/2000 and 2008-2010. In broad terms, nominal GDP/capita increased by approximately 43% between 2000 and 2010 and if travel patterns had remained the same then business VTT would have increased by a similar amount (when adjusting from 2000 to 2010 prices and values). The Department analysis indicated that not all this growth actually occurred – broadly speaking only 65% of the expected growth in average business VTT occurred in practice.

There are three mechanisms at play here[103]. There is the inflationary change in prices, there is the real growth in GDP/capita and there are behavioural changes amongst business travellers. With regards to the latter, what is happening is that whilst business trips and the wages received by people continue to grow, the trip-making characteristics of individuals are changing as their incomes grow. This is most obvious in the higher income category, which should have seen a larger growth in trips than the evidence has indicated. Without further research it is unclear what the underlying causes of these changes in behaviour are. Maybe as incomes increase there is increased substitution to other forms of ‘higher speed’ travel – such as air. Maybe increases in incomes reflect changes in seniority within companies, and those who were previously the higher income travellers (in e.g. 2000) now have people coming to meet them rather than them going to meet others. Such behavioural changes reduce the average growth in business VTT below that expected from pure changes in GDP/capita.

Ideally, the demand modelling of an appraisal should account for the changing composition of the travelling component of the workforce (as remuneration

[101] Wage growth of professional drivers can be analysed using the Annual Survey of Hours and Earnings (ASHE), but this is outside the scope of the current study.

[102] Source: ONS GDP/capita data series at Chained Volume Measure - http://www.ons.gov.uk/ons/site-information/using-the-website/time-series/index.html

[103] A fourth mechanism, changes in hours worked, has been controlled for in the analysis.

grows), whilst the valuation methodology confines itself to how values at given income levels change over time. The issues however become confounded, particularly as it is extremely rare that data on the incomes of travellers forms part of the demand modelling process (see also discussion in the previous section regarding the treatment of income). In the absence of such modelling, ideally a composite elasticity of both changes in remuneration and changes in behaviour is needed.

Once again, without further research, in this case aimed at understanding these underlying changes, it is difficult to make a recommendation for changes in business VTT over time other than reverting to the status quo of a unit elasticity to GDP/capita. This is of course in the context of business VTT values being based on the CSA. As part of this study, we are also examining WTP as a basis for valuing time savings during the course of work. If the WTP approach is adopted, then we would consider that the growth in VTT should also be based on a meta-analysis of business WTP data. This is for similar reasons to non-work VTT; the productivity of time whilst travelling may alter over time and, related to this, the inconvenience and comfort whilst travelling may alter over time. In addition to these reasons, the nature of business travel in terms of how firms use travel as an input to the production process may alter – thereby placing different demands on the business traveller which could be reflected in valuations. However, there is relatively little direct evidence on this. We can draw some insights from the meta-analysis work on the value of employees' business travel time undertaken by Wardman *et al* (2013 pp51-52)[104], where they rejected any time trend in VTT differing from 1.0.

In the absence of a fuller research programme, and on the basis of the Wardman *et al* research, we fall back to the status quo position of a unit elasticity to GDP/capita growth rate for the VTT of business trips if based on WTP data. We recommend that consideration be given in the NTS to verifying further that business trips are funded by the employer, not the employee, since this appears to be a grey area.

[104] Wardman, M., Batley, R., Laird, J., Mackie, P., Fowkes, T., Lyons, G., Bates, J. and Eliasson, J. (2013) 'Valuation of Travel Time Savings for Business Travellers, Main Report'. Prepared for the Department for Transport. Report dated April 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/251997/vtts_for_business_main_report-dft-005.pdf

8 Conclusions and Recommendations

8.1 Summary of the research

This report has outlined the research undertaken and results produced in the course of a year by the study team. The team has worked to a very comprehensive brief encompassing:

- Estimation of VTT for both non-work and business purposes.
- Estimation of VTT for several modes, including car, rail, bus and 'other PT'.
- Estimation of multipliers of VTT covering a range of time-related factors, including reliability, traffic conditions and crowding.

The research has been reviewed and quality assured, through the following provisions:

- Internal review and quality assurance throughout the study.
- Review of key deliverables (namely the Phase 1 report, Emerging Values report and Phase 2 report), by the Analytical Challenge Team (ACT).
- Review of key deliverables by the Department's in-house Project Board.

The Department also commissioned SYSTRA/Imperial College/Danish Technical University to undertake an independent peer review and audit of the data collection and modelling work, and this is reported in a self-standing deliverable.

The research that we have undertaken in pursuit of the study aims (outlined in **Chapter 0**) can be summarised as follows.

- *Provide recommended, up-to-date national average values of in-vehicle travel time savings, covering business and non-work travel, and based on primary research using modern, innovative methods.*

Within the standard microeconomic framework of goods vs. leisure trade-off subject to money and time constraints, we have employed Stated Preference (SP) experimental methods to estimate willingness-to-pay (WTP) for time savings, for travellers in the course of business and non-work trips. The SP involved three games, which considered different trade-offs, namely: SP1 (time vs. money), SP2 (time vs. money vs. reliability), and SP3 (time vs. money vs. crowding/congestion).

The main field survey employed, primarily, an intercept-based recruitment approach at around 30 representative locations across England for each mode. Recruits were directed to a web-based questionnaire, and were paid a nominal reward on completion of the questionnaire. Potential gaps in this primary approach (e.g. infrequent travellers and/or those without computer access) were covered by a secondary approach of telephone recruitment and telephone and/or paper-based administration of the questionnaire. The field survey was completed successfully; most targets were achieved, and overall the pilot survey provided a good guide to the likely success, issues and response rates in the field survey.

As anticipated, given the primary use of an intercept-based approach, the samples were representative in socio-economic characteristics, but weighted towards higher than average trip lengths. This was recognised in the specifications used for the behavioural models, such that appropriate reweighting to representative levels was possible during the implementation phase. Data cleaning according to specified criteria resulted in the removal of an acceptably low proportion of the records for the analysis phase while helping us to avoid biased results. All cleaning processes were documented.

The modelling work carried out to derive the behavioural values made use of state-of-the-art approaches and included a number of innovations compared to existing methodology. In common with the most recent Scandinavian studies, we adopted discrete choice models that incorporate a multiplicative (as opposed to additive) error structure, giving a more flexible and direct approach for capturing the heteroskedasticity that is inherent within data of this type, especially as a result of the trip length mix. We also incorporated a flexible treatment of reference dependence, allowing for recognised behavioural phenomena such as size and sign effects. In common with most recent studies, we also recognised the inability of models to explain all heterogeneity in valuations on the basis of measured covariates, and thus allowed for additional random heterogeneity. As a further novel element, our work made use of a joint modelling approach across the three games presented to a respondent[105], greatly increasing our ability to produce robust estimates of covariate effects as well as the random heterogeneity in valuations across respondents.

The work has produced meaningful behavioural outputs for all estimated models. We have reported initial insights into key socio-demographic impacts, as well as levels of heterogeneity in the valuations, and differences across modes and purposes. The impact of covariates related to three categories, namely trip characteristics, traveller characteristics and design impacts. It is important to note that our modelling approach was developed in such a way that the design impacts, with the exception of size effects, can be factored out of the calculation of valuations. This allows the Department substantial flexibility in how these covariates are used for appraisal. Our work provides particular insights into the effects of different travel conditions, such as congestion and crowding, on valuations of travel time. The work highlighted that valuations are broadly consistent across the three games presented to each respondent (i.e. SP1-3), which validates our approach, though the full insights on these valuations become apparent only when the models are applied to a nationally representative sample (i.e. NTS).

Emanating from a multitude of academic and consultancy WTP studies, as well as numerous ‘national’ WTP studies sponsored by transport ministries and associated agencies across the world, there has accumulated a considerable body of knowledge on non-work VTT; this knowledge base underpins the research presented here. By comparison, WTP-based studies of business VTT have been few and far between. The convention in most countries, including the UK, has been to employ the cost-saving approach (CSA), or some variant thereof, to value business travel time savings. Against this background, the WTP-based approach employed in this study represents an important contribution to the limited evidence base on business VTT. In particular, we have ‘triangulated’ estimates of business VTT from several sources – not only our own WTP-based estimates from employees and employers – but also estimates derived from wage rates, and

[105] With the exception of a number of games presented to a subset of respondents, such as operator and mode choice, which were too distinct from the main experimental work to enable incorporation within a joint framework.

also from the CSA. Despite the challenging nature of this triangulation exercise, we feel that these multiple sources of evidence converge to tell a consistent story, which then justifies basing recommended values on the employee SP dataset on the grounds that this dataset is more robust than the employer dataset.

- *Investigate the factors which cause variation in the values (e.g. by mode, purpose, income, trip distance or duration, productive use of travel time etc.) and use this to inform recommended segmentation of the values.*

The modelling has been developed in a systematic fashion, whereby alternative model specifications have been tested in relation to:

- time and cost gains, losses and size effects;
- person characteristics such as age, gender, employment status, household composition and income; and
- trip characteristics such as mode, purpose, distance and geography.

All values have been reported by purpose and mode – given the importance of these variables as key policy segments – and as a function of income elasticity, as well as a function of elasticities with respect to the time and cost of the 'reference' trip. Where other covariates have been identified as significant, these have been retained in the preferred behavioural models and, in turn, feed into the process of eliciting values for implementation in appraisal.

Indeed, a key task has been the progression from the values derived from best specification models to recommended values for use in appraisal. This has entailed four steps:

- Review and modification of the behavioural models in the light of interpretation
- Reweighting of the estimation sample for national representativity using the National Travel Survey (NTS)
- Specification of standard values to be produced and weighting to produce those values
- Recommendations regarding changes in VTT over time

In order to allow the Department maximum flexibility, we have developed an 'Implementation Tool' which calculates VTT for different segments, based on a variety of weighting options. Rather than just apply weights from the NTS sample to the respondents in our estimation data, we make use of sample enumeration which applies the estimated models to the respondents in the NTS sample. A key benefit of the Tool is that it will enable the outputs of our study to be used with greater confidence over a longer time horizon, not least because it is possible to update the Tool as and when new NTS data becomes available.

- *Improve our understanding of the uncertainties around the values, including estimating confidence intervals around the recommended values.*

For any appraisal scenario of interest to the user, the Implementation Tool generates confidence intervals around the mean VTT. This estimation takes account of two potential sources of uncertainty in values, namely:

- Error in the mean VTT arising from the parameters estimated in the behavioural model.

- Error in the NTS sample, which arises from the fact that NTS is a sample drawn from the travelling population.

More generally, we have demonstrated the robustness of the estimated values through a number of exercises. First, we have validated our SP values against comparable RP values. This exercise involved the design and implementation of separate SP and RP experiments around a common choice context – namely the choice between competing rail operating companies for medium distance rail trips to London. Whilst values were somewhat high, due to confounding with operator preferences, the SP and RP delivered similar results. Second, we have validated our preferred behavioural model specification by testing against alternative specifications, including the recommended model from the 2003 national study. This demonstrated the greater functionality and superior statistical fit of the 2015 model, as compared with the 2003 model. Finally, it should be remembered that any decisions in relation to model structure have been based not just on best practice from past studies and more recent developments in the academic literature, but also on extensive testing on the data at hand, such as tests comparing additive and multiplicative model structures.

- *Consistently estimate values for other trip characteristics for which values are derived from the values of in-vehicle time savings.*

Whilst Stated Preference experiment 1 (SP1) offered the respondent trade-offs between time and cost, SP2 and SP3 offered trade-offs between travel time, cost, reliability and congestion/crowding. The objective of this exercise was three-fold: first, to ascertain whether, and to what extent, the ‘headline’ VTT from SP1 is confounded with the value of reliability and congestion/crowding, second, to estimate incremental multipliers for reliability and congestion/crowding suitable for implementation in appraisal, and third, to gain insights into what the values from SP1 represent in terms of type of time (i.e. trip conditions).

The use of the joint modelling framework across all games allowed us to examine the differences between individual valuations across games. This joint estimation yielded key insights, for example showing the extent of VTT increases with road congestion (for car and bus), as well as with the level of crowding (for all PT modes). It also showed differences across modes and across purposes as to what type of trip conditions in SP3 the values from SP1 relate to. The values coming out of SP2 diverge from those of SP1 and SP3, and this is likely to be a behavioural impact in terms of how respondents react to variability in trip times.

8.2 Recommended values and multipliers

The recommended estimates of VTT and associated VTT multipliers are given in **Table 8.1** and **Table 8.2** respectively[106].

[106] Note that these are a reproduction of Tables 7.28 and 7.29 presented earlier, and that the estimates of VTT are based on certain assumptions.

Table 8.1: Appraisal VTTs for a Level 1 appraisal (routine appraisal of small and medium sized schemes) with illustrative distance bands (2014 perceived prices, £/hr)

Mode	Distance	Commute	Other non-work	Employees' business				
		All modes	All modes	All modes	Car	Bus	'Other PT'	Rail
webTAG (2014 prices and values)	All	7.62	6.77	25.47	24.43	15.64	24.72	30.07
All modes	All	11.21	5.12	18.23	16.74	-	8.33	27.61
	<20 miles	8.27	3.62	8.31	8.21	-		10.11
	20 to 100 miles	12.15	6.49	16.05	15.85	-		28.99
	>= 100 miles		9.27	28.62	25.74	-		

Notes: Distance weighted, 'all distance' values based on income option 1, for distance-banded values non-work based on income option 2 (household income = £49,684) and business on income option 1, VTT imputed for PT trips with zero cost, SP1 VTTs, Δt=10, employers paying for EB trips, Tool version 1.1.

Table 8.2: VTT multipliers

		Commute	Employees' business	Other non-work
Car	Reliability ratio	0.33	0.42	0.35
	Free-flow	0.51	0.42	0.47
	Light traffic	0.72	0.68	0.83
	Heavy traffic	1.37	1.26	1.89
Bus	Value of early	-2.69	-	-3.20
	Value of late	2.88	-	2.52
	Plenty of seats free and did not have to sit next to anyone.	0.85	-	0.83
	A few seats free but had to sit next to someone/could not sit with people travelling with.	0.89	-	0.84
	A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	1.00	-	1.00
	No seats free – a few others standing.	1.24	-	1.30
	No seats free – densely packed.	2.14	-	2.32
	Value of free-flow	0.99	-	1.22
	Value of slow down	1.39	-	1.36
	Value of dwell time	0.68	-	1.57
	Value of headway	1.68	-	1.60
'Other PT'	Value of early	-2.40	-1.66	-2.98

		Commute	Employees' business	Other non-work
	Value of late	1.75	1.95	2.24
	Plenty of seats free and did not have to sit next to anyone.	0.95	1.00	1.00
	A few seats free but had to sit next to someone/could not sit with people travelling with.	0.97	1.00	1.00
	A few seats free but had to sit next to someone/could not sit with people travelling with. Some standing.	1.00	1.00	1.00
	No seats free – a few others standing.	1.13	1.17	1.10
	No seats free – densely packed.	1.70	1.78	1.87
Rail	Value of Early	-1.77	-1.55	-2.34
	Value of Late	2.86	2.76	3.21
	seated 50% load	0.73	0.75	0.72
	seated 75% load	0.79	0.76	0.72
	seated 100% load	1.00	1.00	1.00
	seated 1 pass per m2	1.09	1.13	1.14
	seated 3 pass per m2	1.31	1.36	1.39
	standing 0.5 pass per m2	1.16	1.29	1.21
	standing 1 pass per m2	1.19	1.38	1.27
	standing 2 pass per m2	1.32	1.56	1.57
	standing 3 pass per m2	1.57	1.61	1.79
	standing 4 pass per m2	1.86	2.03	2.17

Note: SP2 VTT taken as base for reliability and early/lateness. SP1 VTT taken as base for car free-flow, light and heavy traffic. SP1 VTT taken to represent 100% occupancy of seats for PT.

8.3 Findings and recommendations for appraisal values

We begin by offering some generic findings and recommendations, before identifying other recommendations which apply to specific parts of the brief.

8.3.1 Generic issues

We have found that there are significant differences between the VTT of different trip purposes, even after controlling for the characteristics of the trip and traveller.

R1: We recommend that values of travel time (VTT) savings should continue to be distinguished by business, commute and other non-work purposes.

We have found clear evidence of values of reliability and of variation in VTT with traffic conditions and crowding. In this context, it is appropriate to note that current WebTAG guidance on VTT incorporates reliability multipliers, but not multipliers for traffic conditions and crowding.

R2: We recommend that the Department should undertake work to examine the case for extending the scope of VTT guidance to include multipliers for traffic conditions and crowding.

As part of this study, we have developed an Implementation Tool, as a means of translating modelled values of travel time savings into appraisal values (in perceived prices) at whatever level of aggregation is required.

R3: We recommend that the Implementation Tool should be used by the Department to generate appraisal values for scheme appraisal.

We have estimated VTT using three different SP games (SP1: time vs. cost; SP2: time vs. cost vs. reliability; SP3: time vs. cost vs. crowding/congestion).

R4: In the immediate term, we would recommend the values from SP1 as the basis for the 'headline' VTT, since these provide the closest comparator to the 2003 game, and most readily lend themselves to implementation in appraisal. It should be clarified that we interpret VTT from SP1 as referring to 'average' travel conditions, rather than free-flow or uncrowded conditions.

If however crowding/congestion data at an appropriate level of detail can be sourced, then there is a case for basing 'headline' VTT on appropriately weighted values from SP3 – instead of SP1.

R5: We recommend that the Department should undertake further work to examine the viability of using SP3, and its relative advantages/disadvantages against SP1.

Our behavioural model, which forms a key input to the Tool, accommodates size effects within its specification. To neutralise this effect in appraisal values, it is necessary to calculate the values for a given "size". When applying the Tool to generate appraisal values, we examined the sensitivity of values to different assumptions regarding the appropriate "size" to use, and found that 10 minutes produced the most representative values. We note that 10 minutes is an assumption employed in other recent European national VTT studies, and our own analysis therefore supports this convention.

R6: We recommend that all time savings are assigned a constant unit value calculated for a change in travel time of 10 minutes.

In moving from modelled values estimated on the behavioural sample to standard average values based on the population (i.e. NTS), two issues arise, namely the method of (re)weighting and the treatment of income.

We reviewed the conceptual arguments for distance vs. trip weighting of VTT, and tested the empirical divergence between the resulting valuations, finding this divergence to be modest if valuations are segmented by distance.

R7: We recommend the retention of distance weighting. This is for the conceptual reason that the probability of a trip benefitting from the scheme being appraised is proportional to trip length within each distance band.

R8: In combination with R7, we recommend that the Department should disaggregate VTT by distance or some geography typology (e.g. urban/inter-urban) that reflects differences in distance. This will require further work to identify appropriate distance disaggregations. Such work would involve the use of distance profiles from real scheme appraisals, to explore the full implications of the approximations to the 'real' VTT of the distance-weighted VTT under different distance disaggregations.

We have considered a range of income weighting options[107], built them into our Implementation Tool, and provided the comparative results. Such weighting removes the direct effect of income differences on the standard value of time, while retaining the differences due to trip length, geography etc. Our recommendations concerning income weighting differ according to the scheme level, and the level of disaggregation in VTT, as follows.

R9: For business VTT, we recommend income option (1), basing the average value on observed variations in income by person and trip. For non-work VTT, the appropriate treatment of income depends on the level of segmentation in the values. Where VTT is disaggregated by purpose (i.e. commute and other non-work) income option (1) should be used. With further segmentation (e.g. by mode, distance and/or geography) we recommend income option (2), based upon the average income of motorised travellers. For larger schemes, strategies and "user pays" projects, we recommend use of income option (1), irrespective of the level of disaggregation.

While we think there is some argument in favour of the 1980s recommendation to adjust car VTT for group size, we conclude that we have not produced sufficient evidence to justify departure from the current convention.

R10: We recommend that driver values should be treated as representative of all occupants in the car, and should be applied separately to each occupant. Separate vehicle occupancy values should be used for the different purposes.

In our preferred behavioural models, on which the reported VTTs are based, we have used household income as the income variable for commuting and other non-work and personal income for business. This position is informed by empirical tests of alternative income variables within the behavioural model.

R11: We recommend that appraisal values for non-work should be based on household income, whilst business values should be based on personal income.

8.3.2 Business travel

Recalling that professional drivers were outside the scope of the present study, it is appropriate to comment on how we anticipate this segment being treated for appraisal purposes.

R12: We recommend the continued use of the Cost Saving Approach (CSA) for professional drivers. It follows from this that for these categories, separate

[107] Option (1) = Averaging over income, but not segmenting by income. Option (2) = Calculating values at 'average' income.

appraisal values for goods vehicle drivers, light van drivers, bus and coach drivers etc. based on their gross wage plus on costs will continue to be required.

For briefcase travellers, we found a fair degree of correspondence between the values from the employers SP survey and the much larger employees SP survey. This gives confidence that the employee values can stand proxy for the benefits of changes in business travel time. More generally, we found that the employee values, when controlled to NTS incomes and trip lengths amongst other things for business travellers, are on average around 60% of the values in the CSA; this rises to 72% if this is restricted to travellers who are reimbursed for their trip. Longer distance trips yield VTT close to the CSA while short distance trips VTT are well below.

R13: We recommend that the Department reviews the respective merits of continuing with the CSA or moving to WTP values based on the employee survey. Within the option of moving to WTP values, the Department should, with reference to R2, explore the practicalities of incorporating reliability ratios and congestion/crowding multipliers.

We found that VTT for car and rail from the employees' business survey are quite different. We are also mindful that the majority of the benefits from business travel time savings percolate through the economic system, and do not necessarily end with the travellers themselves. Therefore:

R14: The Department should continue to use mode specific values for trips in the course of business.

The business values that we estimated cover a wide range, and are sensitive to certain assumptions concerning the classification of business trips/travellers.

R15: Attention should be paid in the NTS and other data sources to ensure that trips in the course of business are carefully categorised. Specifically we recommend that business trips are restricted to those trips where the employer is paying or could legitimately pay for the trip costs (though not necessarily the time), whether directly by providing a vehicle or tickets or through reimbursement to the employee.

If the Department decides to move to WTP values for business travellers, then it will require suitably segmented values by variables of interest such as income and trip length.

R16: We recommend that the Department undertakes work to explore alternative banding options for business VTT by income and trip length.

Whilst our scope did not include specific research on business VTT over time, we reviewed the Department's current approach in this regard.

R17: With regards to changes in business VTT over time, we believe that the most appropriate recommendation at this point is to retain the status quo of applying a unit elasticity to forecast changes in GDP/capita. Periodic adjustment will be required to account for the difference between forecast and out-turn.

8.3.3 Non-work travel

For non-work trip purposes, the evidence is that the average VTT for 'other' non-work is significantly lower than the average commuting value, all else equal.

R18: We recommend that the Department should maintain its distinction between commuting and 'other' non-work trip purposes.

We found that the differences in values across modes cannot be explained solely by comfort differentials.

R19: We recommend a weighted average of non-work VTT by modal share.

Assuming that the Department wishes to maintain its current policy of having segmented values available for use in some types of analysis and appraisal, the question arises as to the appropriate definition of the segments. Clearly these should include income and trip length and possibly geography.

R20: We recommend that the Department undertakes analysis with suitable scheme data, together with the values and elasticities from the behavioural model and Implementation Tool, to explore the most appropriate form of segmentation.

Whilst our scope did not include specific research on non-work VTT over time, we reviewed the arguments for the treatment of income growth over time.

R21: We recommend an interim position of retaining the existing method of applying a unit elasticity to GDP/capita.

R22: In conjunction with R21, we recommend that further review is carried out. This would cover the results of the present study, and how they cohere with meta-analysis data. This review should also take into account further evidence on temporal stability from the recent studies in Denmark, Norway, Sweden and the Netherlands.