Strategies for personal mobility: A study of consumer acceptance of subscription drive-it-yourself car services

Scott Le Vine

Permanent contact via: scottelevine@hotmail.com

Imperial College London
Department of Civil and Environmental Engineering

Submitted for the Diploma of the Imperial College (DIC), PhD degree of Imperial College London
This study is dedicated to Natasha, and to the future

Acknowledgements

I am indebted to many for their advice, support and inspiration that made this study possible.

My supervisors Aruna, John, and Martin went beyond the call of duty on many occasions, and this research could not have proceeded if not for their encouragement and guidance. I would also like to thank my examiners Kay and Dan, and Alison for placing her faith in me by providing me the opportunity to teach whilst undertaking this research.

This study would not have been possible if not for financial support from the RAC Foundation, and I would also like to thank Elizabeth for her insights and suggestions along the way which have led to a better project.

Warm thanks are due to Carplus, City Car Club and RAC Breakdown Services for providing access to customers of theirs willing to take part in the fieldwork, and to the study participants who generously chose to share their time. Further gratitude is extended to Benefit Technology Inc, DfT’s NTS team, Imperial’s HPC team, NatCen, and SRA for their support of this research in a variety of ways.

Discussions with a number of public sector staff, academics, and industry partners helped to shape the contours of this study, and I would like to thank each of you for your time and thought. (The usual statement applies: any remaining errors are my responsibility.)

I could not have taken on a multi-year assignment of this nature without the unflinching support of my family and friends, in particular those friends who have become like family to me during the course of this research.

The research presented here is my own, except where the work of others is referenced.
Abstract

This thesis investigates consumer acceptance of subscription drive-it-yourself car services [SDCSs], which are an evolution of car hire that began entering the commercial marketplace in the mid-1990s. The aim of this research is to develop techniques to forecast how consumer demand for SDCSs may develop.

On the basis of research reported in this thesis, it is argued that a person’s [strategic] decision to subscribe to an SDCS can be reasonably considered to have a dependency with their expectation of [tactically] using it to access particular out-of-home personal activities. It is shown that people can also be thought to view subscribing to an SDCS as part of a larger ‘portfolio’ choice of travel options. Traditional analyses of people’s travel choices are insensitive to these two issues.

Two datasets, one revealed-choice and the other stated-choice, were designed in order to provide empirical data to test the proposed ‘strategic/tactical’ and ‘portfolio’ analytical form. The revealed-choice dataset made use of web-based data-mining techniques, whilst the stated-choice survey is novel in several respects to address the challenges presented by the SDCS context.

The methodological innovations proposed in this research proved successful in forecasting consumer demand for SDCSs in the empirical application, and appear promising for wider use within the transport domain and related research fields.

Key words: subscription drive-it-yourself car services, mobility resource, mobility attribute, mobility tool, car club, carsharing, cars-on-demand, one-way carsharing, free-floating carsharing system, integrated mobility system, servicisation, servicised mobility, automated mobility, collaborative consumption, research, doctoral thesis, car ownership, public transport season ticket ownership, bicycle ownership, discrete choice model, strategic choice, multi-horizon choice, accessibility, perceived activity set, activity repertoire, online travel planning services, stated-preference, stated-choice, Car2Go, DriveNow, Zipcar
Table of Contents

Chapter 1: Introduction .................................................. page 9
Chapter 2: Background .................................................. page 15
Chapter 3: Gaming-simulation task ................................. page 48
Chapter 4: Analytical framework .................................... page 58
Chapter 5: E-NTS dataset .............................................. page 75
Chapter 6: AVATAR survey ............................................ page 103
Chapter 7: Independent analyses of E-NTS and AVATAR datasets page 129
Chapter 8: Substantive results ....................................... page 158
Chapter 9: Summary & conclusions ................................ page 183

Appendix A: Comparison of ‘distinct’ and ‘combinatorial’ model forms
Appendix B: Derivation of the ‘plateau effect’
Appendix C: Detailed specification and results from the analysis of simulated data
Appendix D: Further results from joint (E-NTS/AVATAR) estimation of parameters
Appendix E: Sample gaming-simulation survey instrument package
Appendix F: Sample AVATAR survey instrument package
List of figures

Figure 1.1: Sample publicity highlighting SDCS marketing messages
Figure 1.2: Workflow of this study
Figure 2.1: Schematic of the course of a person’s car club SDCS engagement
Figure 2.2: Time trend in car club members worldwide and in the UK
Figure 2.3: Month-on-month retention of subscribers for the Communauto car club in 2004
Figure 2.4: Schematic of alternative trajectories for carsharers
Figure 2.5: Overview of studies which have forecasted the potential market penetration of SDCSs
Figure 2.6: Daily usage patterns of the Communauto car club for the year 2004
Figure 4.1: Summary of 'Distinct' and 'Combinatorial' structures
Figure 5.1: Map of England’s Government Office Regions
Figure 5.2: Cumulative distribution plot of the number of journeys performed by NTS respondents that are within the E-NTS sample
Figure 5.3: Workflow of web scraping task
Figure 5.4: Screen capture of formatted HTML output from the Journey Planner travel planning service
Figure 5.5: Plot of observed and predicted journey speed by time-of-day for car journeys
Figure 5.6: Plot of observed and predicted journey speed by time-of-day for public transport journeys
Figure 5.7: Plot of residual values from the web scraping task, disaggregated by mode of transport
Figure 5.8: Plots of residual values from the web scraping task, separate plots for each mode of transport
Figure 6.1: Sample screen introducing the survey respondent to her [his] avatar
Figure 6.2: Sample of the main survey screen (as re-designed following field testing)
Figure 6.3: Screen capture of the request to [female] survey respondents to advise their avatar
Figure 6.4: Sample of the main survey screen prior as pilot-tested (prior to re-design)
Figure 6.5: Cumulative distribution of the duration of the interviews
Figure 7.1: Response of likelihood function to varying the mode-choice-level alternative-specific constant for car passenger travel, using dataset G (Run G2)

Figure 7.2: Response of likelihood function to varying the mode-choice-level alternative-specific constant for taxi/minicab travel, using dataset G (Run G2)

Figure 7.3: Response of likelihood function to varying the mode-choice-level alternative-specific constant for driving a personal car, using dataset G (Run G2)

Figure 7.4: Response of likelihood function to varying the portfolio-choice-level alternative-specific constant for owning a personal car, using dataset G (Run G2)

Figure 8.1: Cumulative distribution of the change in the number of car driving journeys per week between the baseline scenario and Scenario #1 by the 19 people predicted to subscribe to a car club SDCS in scenario #1

Figure 8.2: Response of likelihood function, using the ‘portfolio’ specification and the E-NTS dataset only, to successive increases in the ‘observed’ number of each journeys performed by each person

Figure 8.3: Response of values of the salience (gamma) parameters to increases in the number of people’s journeys which are taken into account
List of tables

Table 2.1: Selected early carsharing projects
Table 2.2: Examples of SDCS studies employing qualitative market research techniques
Table 4.1: Matrix of the set of means of travel enabled by various resource portfolios
Table 5.1: Comparison of online travel planning services
Table 5.2: Spatial match quality for journey origins and destinations
Table 5.3: Proportion of journeys for which no itineraries for various travel modes were reported by the web scraping task
Table 5.4: Summary of descriptive statistics for travel time of journey itineraries, by observed method of travel
Table 5.5: Average residual errors (predicted minus observed) disaggregated by source of spatial match for journey from the web scraping task
Table 5.6: Results of diagnostic mode choice model run using the E-NTS dataset
Table 6.1: Sample asymptotic variance/co-variance matrix, containing dummy values
Table 6.2: Matrix of correlations from the AVATAR survey results
Table 7.1: Comparison of data and model form characteristics for estimating the 'mode' and 'portfolio' choice models
Table 7.2: Summary of observed portfolio choices in the E-NTS and AVATAR datasets
Table 7.3: Summary of observed mode choices in the E-NTS and AVATAR datasets
Table 7.4: Correlation matrix of observations ('portfolio' holdings and 'mode' usage) from the E-NTS dataset
Table 7.5: Correlation matrix of stated choices ('portfolio' holdings and 'mode' usage) from the AVATAR dataset
Table 7.6: Listing of simulated datasets and characteristics
Table 7.7: Results from estimation using simulated data
Table 7.8: Comparison of target and obtained parameter values for run F2
Table 7.9: Results from parameter estimation of ‘portfolio’ choice using only the E-NTS dataset
Table 7.10: Results from parameter estimation of ‘portfolio’ choice using only the AVATAR survey dataset
Table 7.12: Results from parameter estimation of transport mode choice using only the E-NTS survey dataset

Table 7.13: Results from parameter estimation of transport mode choice using only the AVATAR survey dataset

Table 8.1: Results from parameter estimation of ‘portfolio’ choice using the combined E-NTS and AVATAR survey datasets

Table 8.2: Results from parameter estimation of mode choice using the combined E-NTS and AVATAR survey datasets

Table 8.3: Comparison of ‘values of time’ estimates for non-SDCS modes of transport

Table 8.4: Comparison of ‘values of time’ estimates for SDCS modes of transport

Table 8.5: Summary of results from baseline scenario & Scenarios #1 through #7

Table 8.6: Correlation matrix of simulated choices (‘portfolio’ holdings and ‘mode’ usage) from the baseline scenario (#0)

Table 8.7: Cross-tabulation of observed and predicted (baseline scenario) ‘portfolio’ holdings

Table 8.8: Correlation matrix of simulated choices (‘portfolio’ holdings and ‘mode’ usage) from Scenario #1

Table 8.9: Correlation matrix of simulated choices (‘portfolio’ holdings and ‘mode’ usage) from Scenario #5

Table C.1: Target parameter values for the simulated datasets

Table C.2: Obtained parameter values for runs with simulated dataset A

Table C.3: Obtained parameter values for runs with simulated datasets B & C

Table C.4: Obtained parameter values for runs with simulated datasets D & E

Table C.5: Obtained parameter values for runs with simulated datasets F & G

Table D.1: Results from parameter estimation of ‘portfolio’ choice using the combined E-NTS and AVATAR survey datasets

Table D.2: Results from parameter estimation of mode choice using the combined E-NTS and AVATAR survey datasets
Chapter 1: Introduction

Information technologies and the business strategies they enable have fundamentally changed industry after industry in recent decades; the transport sector is no exception. Low-cost commercial aviation has emerged, automotive assembly lines operate on the basis of modern just-in-time management principles, and public transport operators routinely make real-time service information available to the general public.

This thesis investigates the consumer market for subscription drive-it-yourself car services [SDCs], which are an evolution of car hire that began entering the commercial marketplace in earnest in the mid-1990s. The aim of this research is to develop techniques for forecasting consumer acceptance of SDCs. This research considers both the current operating model and prospective future developments in service offers; data generated for this task are sourced from within Greater London.

In the traditional car hire business model the customer travels to a place of business to pick up the hire car. SDCs\(^1\) operate distributed fleets rather than storefronts, with cars located at on- or off-street parking spaces throughout a neighbourhood or a wider geographic area. Operators use dedicated telematics systems to manage fleet operations, customer subscriptions, bookings, billing, etc. The service concept has its roots in the early post-WWII years, though modern information technologies proved necessary for SDCs to become commercially viable.

SDCs market the car access that their services offer as an alternative to car ownership. (See Figure 1.1) Management theorists classify this in a broader economic phenomenon of servicization wherein consumers choose to access a stream of services without the responsibilities (and explicitly forgoing the privileges) of owning and maintaining a physical asset. Examples include cloud computing, fast food, movie-rental-by-post services, and outsourcing of business services. This trend is not universal, however; the proportion of UK households with home computers, for instance, continued to rise through the 2000s despite it being possible to access serviced computing resources at public facilities.
It was found, as reported in subsequent sections, to be fruitful to view the structure of people’s options with respect to SDCSs as having both ‘strategic-tactical’ and ‘portfolio’ dimensions. Thus it is argued that a person’s [strategic] decision to subscribe to an SDCS can be reasonably considered to have a dependency with their expectation of [tactically] using it to access particular out-of-home personal activities. We show that people can also be thought to view subscribing to an SDCS as part of a larger ‘portfolio’ choice of travel options. The cost profile makes it relatively impractical to use an SDCS vehicle for all of one’s travel needs (i.e. as a pure substitute for car ownership), but, depending on the opportunities and constraints in a person’s life, it may be practical to use an SDCS for a portion of one’s travel needs, as part of a ‘portfolio’ together with non-car means of travel to serve the residual travel needs. Such techniques, as they are shown to be practical in later sections of this thesis, are thought to have implications for providing more robust estimates of the knock-on effects of SDCS use on issues such as parking demands in urban neighbourhoods, patterns of road congestion, and the net environmental impacts.

The line of research reported in this thesis has a number of applications, both concrete and prospective. Understanding the trajectory and limits of consumer demand for SDCSs is of direct commercial interest to the businesses which operate them, as well as those in adjacent markets (e.g. traditional car hire firms, the broader automotive sector, public transport operators, etc.) Researchers in the transport domain may find wider applications for the techniques developed to address the challenges raised by this particular new form of personal mobility. Practitioners and policymakers, in both industrialised and developing societies, may likewise find the techniques applicable in evaluating policy options. It is thought that the concepts discussed here – principally the ‘strategic-tactical’ and ‘portfolio’ nature of many of the choices that people face – may also perhaps yield insights into patterns of consumer demand in other areas where policies affect the
spatio-temporal patterns of people’s lives in structural ways, such as city/regional planning, energy systems, and environmental policy.

Figure 1.2 illustrates the work flow of the research reported here.

**Figure 1.2: Workflow of this study**

The remainder of this thesis is organised as follows:

Chapter two presents (in section 2.1) an overview of the present SDCS marketplace and how the services have developed to date. It also contains a literature review relevant to the techniques employed in this research and the central concepts on which the analytical framework rests:
§ 2.2 Consumer acceptance of new products/services: This section describes efforts by researchers to understand and predict consumer acceptance of innovative products and services.

§ 2.3 Discrete choice analysis:

§ 2.4 Stated-response survey design: Two stated-response surveys were designed and taken forward as part of this research. The first was the initial task of this research – a set of qualitative research methods called gaming-simulation. Later in the research, a stated-choice survey was developed, based on the findings from the gaming-simulation research.

§ 2.5 Car ownership: It is proposed to analyse people’s participation in SDCSs as part of a broader choice in which another option is to own a car. This section discusses the development of techniques to analyse people’s car ownership.

§ 2.6 Accessibility measures: The analytical framework introduces a new concept of personal accessibility termed the perceived activity set; existing concepts of personal accessibility are discussed in this section.

Chapter three presents the gaming-simulation research task, with presentation of the methodology followed by a discussion of the findings. Respondents’ [stated] use of SDCSs was found to lead to structural changes in activity participation patterns and use of various means of travel.

The fourth Chapter is a presentation of the analytical framework developed for the main quantitative aspects of this research, largely on the basis of the findings from the gaming-simulation research task. Newly-introduced concepts are defined and the mathematical notation used in subsequent chapters is introduced. The term StrAP is proposed as shorthand for the analytical framework, as the principal analytical innovations are the techniques to incorporate the STRAtegic/tactical and Portfolio [portfolio] dimensions into choice analysis methods.

Chapters five and six present, respectively, the design of the two instruments for generating the datasets used in the main quantitative analysis: the spatially-and-travel-option-enriched National Travel Survey [NTS] dataset purpose-designed stated-choice survey and the. The term E-NTS is proposed to refer to the [spatially-and-travel-option-] Enriched National Travel Survey dataset. NTS data is released publicly at a very coarse level of geographic detail (the Government Office Region level – i.e. London, Rest of South East England, etc.) The E-NTS dataset was prepared by processing (after suitable randomisation to maintain respondent privacy) spatially-enriched E-NTS data through online journey planning services to identify the characteristics of unused-but-available travel options for performing journeys captured in this dataset. The term ‘AVATAR survey’ is employed throughout
this thesis as shorthand for the Advanced Vehicular/Activity/Travel And Resource [AVATAR] survey, in which respondents were asked to make choices within a complex choice context in keeping with the concepts outlined in Chapter four. The functional links designed to facilitate joint analysis using the AVATAR and E-NTS sources of data are also discussed.

Results from analysing the E-NTS dataset and AVATAR survey dataset independently are presented in Chapter seven.

The substantive results from this research are presented in Chapter eight, which include a set of scenarios in which representative market forecasts of SDCSs for Greater London are presented and discussed.

The ninth and final Chapter summarises the work performed, the principal findings, and the questions raised through the course of this research. Avenues for extending the lines of enquiry reported here are discussed.

The principal contributions of this research are:

1) An application of [well-established] gaming-simulation techniques to study the [novel] take-up and use of SDCS services
2) The successful use of newly-available distributed computing resources [online journey planning services] to address spatial and temporal imprecision in traditional PC-based models of transport and land use systems
3) The AVATAR survey design, which, as reported in chapter six, proposes an abstraction between the survey respondent and the agent acting in the survey’s choice situations, in order to address the requirement of generating data of sufficient quality and quantity for use in the StraP framework.
4) The StraP specification of choice model, which challenges the classical assumption in travel demand analyses that people make sequential [as opposed to strategic] decisions regarding ‘structural’ choices (such as car ownership and residential location) and their day-to-day travel habits
5) An application of the StraP system to develop indicative forecasts for the take-up and usage of SDCSs.
The following scholarly articles were prepared as part of this research:


Notes

1 The service model most widespread today is known as carsharing in North American vernacular, which is functionally equivalent with the British term car clubs. The term station cars can be found in the literature, and cars-on-demand is becoming the term of choice for describing the emerging service model in which one-way journeys are accommodated. The term subscription drive-it-yourself car services is employed in this thesis as it is descriptive of the nature of such services and encompasses a wide range.
Chapter 2: Background

This Chapter discusses the state of knowledge as regards the substantive focus of this research, subscription drive-it-yourself car services [SDCSs], and subsequently methodological issues that were identified as relevant during the course of the study.

Section 2.1, the bulk of the material in this Chapter, presents background relating to SDCSs. Section 2.2 describes the general literature regarding consumer acceptance of new products and services. Section 2.3 then discusses the basic quantitative analytical technique employed on this research, known as discrete choice analysis. Section 2.4 describes the literature on stated-response survey methods, and Section 2.5 discusses the body of literature regarding analyses of people’s ownership of cars. As this research proposes a novel conception of people’s spatio-temporal access to life opportunities, Section 2.6 describes the various existing notions of personal accessibility within the field of transport geography. Section 2.7 concludes this Chapter.

The original research then follows in Chapters three through eight.

2.1 Subscription drive-it-yourself car services

Since the first mass production of the motorcar in the early twentieth century, the extent of motorisation could serve as a reasonable indicator of where societies are located (at any given point in time) on a development continuum. Within this period, ownership of a private motor vehicle has tended to be closely correlated with factors such as income, household size, and residential location at the household level. Classical studies have made use of these robust and stable empirical observations to produce aggregate forecasts of the demand for private vehicle ownership. (e.g. Chow 1957; Mogridge 1983) These forecasts have served as inputs (in some cases subject to explicit feedback effects) to the development of transport policies at various geographic scales.(de Jong et al. 2002)

A recent development in the way automobiles are used in developed societies is the growth of organisations known in the UK as ‘car clubs’, which allow subscribers to use cars that are managed by the organisation. Car club SDCSs may be not-for-profit or commercial operations, with membership size ranging from a handful to several hundred thousand. Regardless of their organisational characteristics, they all operate on the principle of separating car ownership and use. In this respect, the car club SDCS mode of transport bears some similarities to both taxi services and traditional car rental services. The differences with these other modes are related to attributes such as the relative ease of engaging a car club vehicle and pricing structures.
Car club SDCS subscribers typically pay some combination of a fixed monthly/annual subscription fee, a time-varying usage fee, and a distance-varying usage fee. Figure 2.1 shows a schematic of these steps that a car club SDCS subscriber experiences in using a carsharing vehicle, from initially joining the service through invoicing. The subscriber reserves a vehicle, accesses it, performs their travel/activity episode, returns the vehicle at a specified time, and is typically billed for usage on a monthly basis. In modern operations, many of these steps are performed automatically through dedicated IT systems, which require significant though not prohibitive investment.

![Usage Cycle Diagram]

**Figure 2.1: Schematic of the course of a person’s car club SDCS engagement**

### 2.1.1. History of SDCSs

The first experiments with shared-car services took place in Europe from the late 1940’s, and achieved relatively little commercial success. Readers are referred to (Shaheen et al. 1998; Parvianen 1983; Robert 2005c) which provide detailed listings of early carsharing programs. Several noteworthy ones in are listed in Table 2.1.
<table>
<thead>
<tr>
<th>Name of service</th>
<th>Location</th>
<th>Year of founding</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEFAGE</td>
<td>Zurich, Switzerland</td>
<td>1948</td>
<td></td>
</tr>
<tr>
<td>Procotip</td>
<td>Montpelier, France</td>
<td>1971</td>
<td></td>
</tr>
<tr>
<td>Witkar</td>
<td>Amsterdam, Netherlands</td>
<td>1973</td>
<td>(Bendixson and Richards 1976)</td>
</tr>
<tr>
<td>Green Car</td>
<td>London and Milton Keynes, UK</td>
<td>1975</td>
<td></td>
</tr>
<tr>
<td>Taxistor</td>
<td>Brussels, Belgium</td>
<td>1978</td>
<td></td>
</tr>
<tr>
<td>Sambil</td>
<td>Västerås, Sweden</td>
<td>1980</td>
<td></td>
</tr>
<tr>
<td>Hyrbil</td>
<td>Örebro, Sweden</td>
<td>1983</td>
<td></td>
</tr>
<tr>
<td>Short Term Auto Rental (STAR)</td>
<td>San Francisco, California, US</td>
<td>1983</td>
<td>(Walb and Loudon 1986)</td>
</tr>
<tr>
<td>Mobility Enterprise</td>
<td>Purdue University, Indiana, US</td>
<td>1983</td>
<td>(Doherty et al. 1987)</td>
</tr>
</tbody>
</table>

Table 2.1: Selected early carsharing projects

Two of the early experiments in the US, the Mobility Enterprise and STAR programmes, are illustrative. The Mobility Enterprise project – in which participants were given small ‘minimum attribute vehicles’ to use for most of their travel, while having as-needed access to shared larger vehicles – was discontinued after the initial research grant cycle, when it would have otherwise required another financial sponsor. (Doherty et al. 1987) The STAR programme was a demonstration venture that did not continue, apparently due to issues with its clientele and pricing structure. Its inability to collect revenues which it was due, together with issues related to its pricing structure and vehicle maintenance, contributed to the programme’s cancellation roughly halfway into its intended three-year duration. (Shaheen et al. 1998; Millard-Ball et al. 2005; Walb and Loudon 1986)

The first non-demonstration SDCS in North America was Autocom in Quebec City in 1994, which is now part of the Communauto organisation. (Robert 2005a) By the late 1990’s, car club SDCSs began to experience scalable success in Europe, North America, and Asia, and their numbers and breadth have expanded rapidly in recent years. A variety of factors – growing environmental awareness, commercialisation of the relevant information technologies, and visionary entrepreneurs – appear to have catalysed the market viability of the SDCS concept.

Two of the largest shared-car services are Mobility CarSharing Switzerland and Zipcar with c.100,000 and c.400,000 subscribers respectively. (Mobility carsharing 2011, Griffith 2010) Zipcar has very recently been joined by Connect by Hertz as the only operators with inter-continental operations.
(both operate in North America and Europe). Zipcar reached a dominant position in the North American market through its 2007 acquisition of Flexcar, which was at the time the second-largest U.S. operator (behind Zipcar), and has engaged in a similar strategy of acquisition in the UK and continental Europe. Although car club SDCSs are expanding rapidly at present, the market is still to a certain degree geographically segmented; many SDCSs operators are without direct competition in their markets.

Novel operating models are currently in planning or piloting stages in various locations. It appears likely that shared-car services in coming years will have quite different operating characteristics than the type of services which are in today’s commercial marketplace.

One recent innovation is the ‘car2go’ SDCS that has been introduced by Daimler, the parent company of the Mercedes-Benz brands. It is presently operating in roughly five moderately-sized cities in continental Europe and North America. (Firnkorn and Muller 2011) The service is unique in that it is designed for customers to access the cars primarily without a reservation, pay-by-the-minute whilst driving to their destination, and then end the accrual of fees and their responsibility for the vehicle by parking it in any public parking space at the end of their one-way journey. Another novel system is being tested by Peugeot, in which an operating model broadly similar to the traditional SDCS model is extended by providing customers a wide variety of [Peugeot] vehicle types, including sedans, electric bicycles, scooters, and vans. The expectation is that subscription to this service, called ‘Mu’, will be bundled with purchase of en electric vehicle from the firm in order to provide EV buyers with access to internal-combustion vehicles on instances where the limited range of an EV would be problematic. The Mobility SDCS in Switzerland, which is tightly-integrated with public transport operators, offers customers the option of casual use of a variety of cars without a subscription, a service which is similar to traditional car hire in many respects (the main distinction being the fine temporal granularity of the reservation.) Start-up firms are also beginning to emerge which operate peer-to-peer SDCS business models, in which car owners rent to pre-screened drivers through a trusted network and with the protection of a bespoke insurance product to manage liability. The City of Paris has been planning for several years an SDCS network of electric vehicles, with a one-way service model similar to that of ‘car2go’, though the requirement for charging means that the system will likely have discrete stations where the vehicles are accessed and must be dropped off. (Royer 2008)

2.1.2. Defining SDCSs
Defining the term SDCS is not straightforward, in particular with respect to drawing a line between SDCSs and the service provided by traditional car hire outfits. In some cases, in fact, car club SDCS
usage has been deemed to be subject to pre-existing car hire taxes, which some researchers argue is inappropriate. (Nassauer 2008, Bieszczat and Schwieterman 2011)

A 2005 overview of carsharing identified the U.S. State of Washington as having a concise and effective definition:

A membership program intended to offer an alternative to car ownership under which persons or entities that become members are permitted to use vehicles from a fleet on an hourly basis. (Millard-Ball et al. 2005)

The same report noted an alternative definition from the City of Toronto:

... the practice where a number of people share the use of one or more cars that are owned by a profit or non-profit carsharing organization [CSO]. To use a vehicle a person must meet the membership requirements of the CSO, including the payment of a membership fee that may or may not be refundable. Cars are reserved in advance and fees for use are normally based on time and miles driven. CSO’s are typically residentially based with cars parked for convenient access within the area of the membership served by the organization.

By design, these definitions draw a contrast between SDCSs and traditional car rental services. It appears that one of the reasons for this distinction may be a desire for a market identify distinct from car hire.

From the vantage point of the academy, and recognising the pace at which SDCSs continue to evolve, a broader definition is proposed. An SDCS is thus defined here as:

An arrangement which provides on-demand access to a car to drive in exchange for monetary or other consideration.

Such a definition excludes only private car ownership and informal shared-car arrangements for which no exchange of value is made.

### 2.1.3. Dimensions of SDCSs

SDCSs operate a variety of different business models. Some, especially smaller ones, are cooperative or non-profit organisations. Larger operators tend to be for-profit concerns, although present market dynamics see some shared-car services undertaking market-share-building in lieu of pursuing short-term profitability. (DiStefano 2008)
The service offers vary considerably, both between different operators and amongst different services offered by a given operator. The most significant differences amongst SDCSs fall into several categories:

- **Membership Selectivity and Definition**: Different operators have different policies regarding their selectivity. The age of eligibility varies, for instance, as does the willingness to accept prospective subscribers with limited or imperfect driving histories. Further, membership may be offered at either the individual or household level by various operators.

- **Pricing Structures**: Similar to mobile phone plans, shared-car service memberships are available on different terms based on a person's expectations of usage. Some plans have high variable costs, with low or no fixed costs, while others are structured differently. Certain operators have rate plans targeted at markets such as independent consultants, for instance, who may need a car for several days or weeks at a time to serve particular clients. Options for insurance may also vary, with higher fees for policies providing a subscriber with better coverage.

- **Ease of Transaction**: Operators with more advanced IT systems can generally facilitate their subscribers' usage better than those with smaller or more primitive systems. This category includes the car-scheduling system, car-access system, mileage-reporting system, and also the proximity of an available vehicle to the user's location. Additionally, some operators have in-vehicle credit cards with which to pay for fuel, while in other cases subscribers pay for fuel, subject to reimbursement afterwards. The use of a shared car may be more likely to require a stop at a petrol station than when driving one's own vehicle; operators may require subscribers to return the vehicle with a stated level of petrol in the tank.

- **Customer Service**: As with any retail business, shared-car service operators interact with their clients in any number of ways. The light staffing levels made possible by automated operations mean that customers may come into contact with staff only when there is a problem of some sort (unclean or damaged vehicle, vehicle missing at start of reservation, mechanical problem, etc.). Anecdotal evidence (e.g. discussion on consumer review websites) indicates that standards of customer service vary widely, which may be associated with operators' rapid organisational growth.

- **Vehicle Availability**: The fleet size, network density, and operational decisions of a car club will determine the subscriber's experience in terms of vehicle availability. Research has unsurprisingly shown that customers place a high value on vehicles being available close to
their home with reliable availability. (Millard-Ball et al. 2005) Further, some operators are present in multiple cities and are thus able to offer subscribers the use of vehicles in different urban areas.

- **Vehicle Fleet Composition:** Some shared-car service operators have relatively homogeneous vehicle fleets, which facilitates maintenance and otherwise simplifies fleet management. Others, however, have more varied fleets which can include cars with child seats, small sedans, sports cars, sport utility vehicles, minivans, etc. to better match a subscriber's requirements for any given reservation. Operators also vary in terms of the condition and cleanliness in which vehicles are maintained.

- **Reservation Attributes:** Nearly all operators require subscribers to: 1) reserve a vehicle in advance, 2) return a vehicle at the same location from which it was obtained, and 3) identify the usage duration when making the reservation. The subscriber is billed for the entire duration of the vehicle reservation, including their time at a destination. All three of these reservation attributes are subject to ongoing evolution. Other relevant attributes include pet-friendliness, the provision of several minutes of non-billed ‘buffer time’ before and after a subscriber’s reservation, and the schedule of penalties for infractions such as returning a vehicle late or smoking in vehicle.

- **Partnerships with Merchants:** Some operators have relationships with merchants through which they offer discounts to their subscribers. For instance, subscribers may be eligible for products and services such as reduced-rate transit tickets/passes, taxi fares, or ride-matching services. In another type of partnership, the car2go service has negotiated dedicated parking arrangements at a number of commercial establishments.

It is noted that the customer’s experience of using a shared car is qualitatively different from using one’s own personal car, despite the on-road similarities of self-driven travel in a passenger car. The listing below of the major differences is reproduced from (Katzev et al. 2001) quoting Bernard writing in 1998. (Bernard 1998) The notes in italics are this author’s (Le Vine) comments pertaining to how the qualitative user experience has since evolved:

- A user has to plan their trips in advance. So in most cases spontaneity is lost.

  *This attribute of the carsharing experience remains true today, but less so than in the late 1990s. The ability to make reservations with one’s mobile phone in near-real-time mitigates this drawback of SDCSs, provided a vehicle is available.*
The user has to remember and take the time to make a reservation.

_The requirement to make a reservation remains the industry norm. Making a reservation has become easier, however, as the options in terms of reservation have increased and become more user-friendly._

The car is probably parked further from the user’s residence than their personal car would be.

_This service feature remains today, though the increasing density of car club vehicles is mitigating the difference somewhat in some places._

The user has to leave it clean, every time, even if they are in a hurry.

_This attribute remains today. In many cases, shared cars are not reviewed by staff between reservations, hence cleanliness is policed by the next user of the vehicle. This is a major difference in service levels between SDCSs and traditional car hire._

The user has to deal with some form of paperwork, pin numbers, lock boxes, etc. every trip.

_This attribute is less true today. Contemporary information technology systems have automated many of these steps which were previously performed manually by the customer. Generally speaking, the larger operators have more-advanced IT systems as they are able to spread the fixed costs more widely._

The user has to worry about getting the car back on time – another loss of spontaneity.

_This aspect of carsharing is still salient, though operators typically allow subscribers to request time extensions during a reservation. The operator then seeks to allow the extended usage, which depends on the ability to meet reservations previously made by other clients. In any case, there is an additional element of inconvenience to the customer in having to actively request the time extension._

### 2.1.4. Review of Academic Literature on Shared-Car Services

The body of literature on SDCSs can be characterised as broader than it is deep. As with any new product or service, there are many aspects which have been subject to academic inquiry as researchers probe the service features and its implications for society at large. This section explores the breadth of the literature at a relatively high level.

It is proposed that the academic literature on SDCSs can be organised into the following categories:
• Economic theory in support of carsharing
• Findings of demonstration projects
• Descriptions of the scope of existing regional / national / world shared-car markets
• Operational analyses of individual systems
• Qualitative research exploring shared-car services
• Descriptive results from surveys of carsharers
• Quantitative analyses of subscribers’ travel behaviour
• Environmental impacts of carsharing
• Forecasts of the potential market penetration of shared-car services

The early literature on carsharing identifies a set of challenges associated with the private automobile and provides the theoretical underpinnings for the SDCS form of travel. Researchers considering urban transport issues investigated the cost structure of the private motorcar, which is characterised by relatively large fixed costs and small operating costs. SEFAGE\(^1\), the earliest recorded commercial car club SDCS, started in Zurich in 1948 with an objective to increase car access to non-car-owners by providing shared car access to about a dozen members of a housing co-operative who apparently could not otherwise afford personal cars. (Harms and Truffer 1998)

SDCSs were also seen as a mechanism to manage the spatial challenges of rapid motorization, such as acute infrastructure needs in historic city centres, particularly in post-war Europe. (d’Welles 1951; Fishman and Wabe 1967) At the same time, the social costs that individual car trips imposed (through the inefficiency of traffic congestion) were recognised. (Smeed 1964) It was further noted that private vehicle ownership meant that many people appeared to acquire vehicles based on travel patterns they might do occasionally (such as family outings), even though a smaller vehicle would be sufficient for many of their routine travels. (University of Pennsylvania 1968) Fishman and Wabe (1967) present an early justification for carsharing on the grounds of economic efficiency – the authors argue for the efficiency of average-cost pricing for car use to match the pay-as-you-go pricing model public transport prevailing at that time.

The first SDCS projects were experimental and thus tended to be closely-monitored by researchers.\(^2\) A common feature across them is that the operating concepts were quite fluid as different service models were tested. As the SDCS form of transport was subject to the rigours of real-world field-testing and the commercial marketplace it began to evolve into the operating model that is typical of contemporary operations (see Section 2.1.7). One of the lessons learned in the early days of carsharing was put succinctly by Robert, founder of Communauto: *Operations must be kept simple* – he warns of the risks of attempting to implement multiple innovations simultaneously. (Robert 2000)
As subscription to SDCSs has grown rapidly in recent years, research performed at the University of California, Berkeley has tracked the growth and evolution across different nations and continents. (Shaheen et al. 1998; Shaheen and Cohen 2006, 2007) In addition to snapshots of the SDCS market (see Figure 2.2), these articles provide a meta-analysis of reported knock-on effects of SDCSs and brief descriptions of how car club SDCSs operate in different places.

Figure 2.2: Time trend in car club subscribers worldwide, left, (reproduced from Shaheen and Cohen 2007) and in the UK (courtesy of Carplus via Roberts 2008)

The high rate of membership growth has overlaid setbacks in some instances, such as termination of various uneconomic schemes and significant subscriber turnover rates. One study, for instance, has found member turnover rates on the order of 50% per annum. (See Figure 2.3) (Morency, Trepanier, Agard, Martin and Quashie 2007)

Figure 2.3: Month-on-month retention of subscribers for the Communauto car club in 2004. The different-coloured bands correspond to the cohort of subscribers whose subscriptions commenced in each month. ‘January’ corresponds to subscribers joining in January 2004 or previously (reproduced from Morency, Trepanier, Agard, Martin and Quashie 2007)
Qualitative procedures have been employed to investigate how SDCS subscribers and the public at large relate to the service. Qualitative research techniques are especially useful in situations where behavioural patterns and underlying motivations are not well-understood – precisely the case with an innovative service offer such as SDCS. Many analyses have explored the process that SDCS subscribers go through in deciding whether to change their travel behaviour. Table 2.2 below lists several inquiries that fall within this category.

<table>
<thead>
<tr>
<th>Study</th>
<th>Qualitative Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes towards Car Clubs (Harvey et al. 2007)</td>
<td>Focus groups, attitudinal surveys</td>
</tr>
<tr>
<td>A Step Towards Sustainable Transportation Behaviour (Dalla Rosa 2007)</td>
<td>Depth interviews with car club SDCS subscribers</td>
</tr>
<tr>
<td>Changing Consumer Behaviour Through Eco-Efficient Services (Meijkamp 2000)</td>
<td>Focus groups, attitudinal surveys</td>
</tr>
<tr>
<td>Dynamics in Behavioral Adaptation to a Transportation Innovation (Shaheen 1999)</td>
<td>Focus groups; attitudinal surveys</td>
</tr>
<tr>
<td>Paris Carsharing Study: Analysis of Behaviour and Associated Meanings (Title translated from French original) (Jemelin and Louvet 2007)</td>
<td>Focus groups</td>
</tr>
<tr>
<td>The Long Way from Interest to Participation: When Does the Car Owner Change to Car Sharing? (Harms and Truffer 2000)</td>
<td>Focus groups and in-depth interviews</td>
</tr>
<tr>
<td>A Psychological Analysis of Acceptance of Pro-environmental Use of Automobile: Cases for Carsharing and Eco-car (Ohta et al. 2009)</td>
<td>Attitudinal surveys</td>
</tr>
</tbody>
</table>

Table 2.2: Examples of SDCS studies employing qualitative market research techniques

As SDCSs are a relatively new phenomenon, operators engage in fairly regular and intensive customer surveying. When combined with the relatively small population of SDCS subscribers, the result has been a distinct risk of survey fatigue that experts believe may be affecting response rates. (Carplus 2008) In some places, such as the UK, customer surveying is a requirement for an SDCSs accreditation, which itself is frequently a precondition for on-street parking concessions and other public sector support. Such member surveys have served as data sources for much of the literature. In some cases the surveys simply ask customers about their travel behaviour (see e.g. Carplus 2007 and Vance et al. 2005); in others carsharers are asked to record travel diaries (e.g. Cervero et al. 2006). Typical issues explored on nearly all subscriber surveys include:
• Subscribers’ demographics
• Subscribers’ self-reported driving mileage and usage of various methods of transport both before and after subscribing
• Whether a subscriber has sold a car, or decided against purchasing one, since subscribing

In a very small number of instances, data from surveys of subscribers has been used to perform detailed analysis of subscribers’ travel patterns. The work done by Robert Cervero and colleagues (Cervero 2002; Cervero and Tsai 2003; Cervero et al. 2006) has perhaps the most appealing experimental design – including a control group, the previously-mentioned travel diaries, and a panel design beginning prior to the introduction of the SDCS. The researchers prepared basic statistical analyses of mode choice, automobile ownership, and daily vehicle kilometrage. Other researchers (e.g. Morency, Trepanier and Martin 2007) have made use of operational data to investigate subscribers’ travel behaviour, but such efforts have been limited in that the operational datasets – essentially the information that is recorded with each subscriber and each usage episode – only record carsharing behaviour, not subscribers’ overall patterns of mobility. Morency and colleagues have also used operational datasets to explore the supply-side attributes of SDCSs. (e.g. Martin 2007)

The purpose of analysing SDCSs is frequently to perform accounting of the environmental effects. Operators’ surveys frequently find subscribers reporting that on balance they drive less. Combined with the relative efficiency of SDCS fleets, it is quite plausible that significant overall emissions reductions are associated with carsharing. Researchers have in fact developed estimates of such environmental benefits associated with carsharing (e.g. Lightfoot 1997; Muheim 1998; Sloman 2003; Ledbury 2004, 2007, Martin and Shaheen 2010)

The research performed by Cervero and colleagues, using the control-group research design described above, highlights the challenges of identifying the environmental impacts of SDCSs. (Cervero 2002; Cervero and Tsai 2003; Cervero et al. 2006) It will matter, for instance, whether a person was a car owner beforehand, and, if and when they end their SDCS subscription, whether they then decide to own a car. The particulars of the initial pre-SDCS state, the period of subscription, and the end state will determine whether a person’s overall automobile travel increases or decreases. It should be noted that analysis of effects is in theory based on comparison between alternative futures rather than a current state and a future state. For instance, if a non-car-owner joins an SDCS and then subsequently leaves to become a car owner, the role of the SDCS in facilitating such a pathway could be viewed as catalysing car purchase and thus facilitating higher automobility. If the person would otherwise have purchased a car at an earlier point, however, and
the availability of an SDCS played a part in delaying the purchase, such an interpretation of the net environmental effects associated with SDCSs could be incorrect. Martin and Shaheen 2010 address the first of these issues (relating to ‘alternative futures’) through the design of several questions in their survey instrument.

Finally, the literature includes forecasts of the potential market for shared-car services, based on a variety of procedures. Figure 2.4 shows the range of forecasts that have been prepared. Note that the studies cannot be directly compared to each other due to different techniques, data sources, populations of interest, and research goals; the results are however collated in this graphic in the interests of clarity.

![Figure 2.4: Overview of studies which have forecasted the potential market penetration of SDCSs](image)

**Figure 2.4: Overview of studies which have forecasted the potential market penetration of SDCSs**
A frequent technique is to use a break-even point for aggregate driving mileage – with the assumption that individuals (or vehicles, depending on the unit of analysis) that drive less than a given threshold would be better off using an SDCS rather than owning a private car. (e.g. Schuster et al. 2005) Another set of procedures employed regularly is to divide the population into discrete market segments (demographic, geographic, behavioural, etc.) and project market penetration rates for each segment. (e.g. Dallaire et al. 2006, Clark 2010)

One study, as reported in (Fukuda et al. 2005) employed a stated choice exercise to investigate the carsharing market. The focus of this study was identifying salient service attributes and people’s attitudes towards carsharing that would support a prospective service in Bangkok. The instrument asked respondents to choose amongst various methods of travel, as in a typical ‘mode choice’ analysis. The results appear to have been used in a diagnostic sense, to identify the key drivers of demand for a prospective SDCS, rather than to forecast take-up.

A similar study was reported by Yang (2010), who explored the take-up of a number of innovative forms of travel, including a ‘one-way’ SDCS (akin to the ‘car2go’ service model, though with vehicles available for customers to access at nodal stations rather than free-floating within an urban area.) This study addressed the issue of the use of various modes of transport in combination with the time-of-day of travel. A distinction with the present research is that the survey instrument for this study was designed with an explicit assumption that the addition of the new modes of transport would not lead to structural changes in people’s travel capabilities, such as ownership of a car.

Shaheen prepared perhaps the most rigorous analysis of SDCS take-up in her doctoral dissertation. (Shaheen 1999) The research programme tracked the implementation of a ‘station car’ type of SDCS project. The market research consisted of focus groups, an introductory video, a paper brochure, and a trial clinic. Would-be SDCS participants were surveyed in a longitudinal manner at several points after they received the informational media, and completed a three-day travel diary following the SDCS trial clinic. The findings are focused on the relationships between personal values, attitudes, and propensity to subscribe to an SDCS.

2.1.5. Prior findings related to SDCSs
The theories underlying the SDCS concept fall broadly within the same school of thought as that of congestion pricing: *variabilisation*. In the case of SDCSs, this implies conversion of the ‘high fixed cost / low variable cost’ model of private automobile ownership into a ‘low fixed cost / high variable cost’ model more closely resembling those of a number of competing urban transport modes.

The theorised effects of SDCSs are:
• ‘Levelling' of the degree of mobility available to (and consumed by) individuals: Those not previously owning a car are expected to see their mobility increase, and vice versa for those formerly keeping a private car (who may choose to forego direct car ownership.) This effect, as measured in a variety of metrics, has been borne out by field experience. However, market research shows that car club subscribers tend to be well-educated and earn middle/upper incomes. The evidence suggests shared-car services have shown less effectiveness in providing mobility to disadvantaged individuals. (Shaheen and Cohen 2006)

• Overall greater efficiency of vehicle use: While the typical private automobile remains stationary for roughly 95% of a given day, SDCS cars tend to need to be in revenue operation on the order of 10 hours/day. (City Carshare 2005) (Although it is noted that a component of revenue operation is that time which the user is paying for but when the vehicle is parked.) On the basis of a large body of subscriber surveys indicating that in the absence of SDCS many subscribers would otherwise own cars, SDCS cars are said to 'replace' multiple private automobiles and thereby reduce overall fleet size and associated parking demand. Of additional note, SDCS cars tend to be relatively new, fuel-efficient, maintained in good mechanical condition, and carry higher-than-average passenger occupancies. (Cervero et al. 2006)

• Generally more ‘prudent’ automobile travel by SDCS subscribers than drivers at large: Use of a shared car requires promptness, cleanliness, diligence, forethought and schedule-adherence on the part of the traveller – in addition to involving higher variable costs than private car travel. Some survey results from North America show that subscribers use a car club vehicle on average for 2-3 trips per month, or approximately 5% of their overall travel. (Millard-Ball et al. 2005) Recent research in the UK has found considerably higher percentages, in the range of 20-25%. (Carplus 2008) Carsharers have also been found to exhibit a higher proportion of transit and non-motorized trip-making than comparable non-subscribers. (Cervero et al. 2006)

• Use of shared-car services as supplementary to private cars, or as might be termed ‘mobility insurance.’ Researchers have found that a substantial proportion of SDCS subscribers continue owning a private vehicle [in their household] (Cervero et al. 2006, Carplus 2008, Martin and Shaheen 2010) For these subscribers, the SDCS may allow them to forego the purchase of a second household vehicle (subject to internal household dynamics.) Interestingly, it was found that when an SDCS in Portland, Oregon introduced a fixed-cost
subscription fee for the first time, approximately 30% of the subscribers departed from the
service though the effect on usage levels was relatively minor. (Millard-Ball et al. 2005)

- \textit{Greater propensity to use shared cars (relative to private cars) only for specific mobility
  needs, such as:}
  - Journeys with requirements for goods-carrying capacity
  - Round-trip journeys where the duration of the activity does not dominate the
duration of the travel to the activity (e.g. commuting to/from an eight-hour workday
is less conducive to SDCS use than a two-hour shopping tour with multiple stops.)
  - In terms of traditional journey purpose categories, journeys to access shopping and
leisure-related activities are typically cited in the literature (e.g. Transport for
London 2008)

2.1.6. Profile of SDCS subscribers

The archetypal SDCS subscriber has a quite distinct set of socio-demographic characteristics,
although there is some evidence that the customer profile is moving from 'early adopters' to a more
mainstream market. (Robert 2005c) identifies the following characteristics of SDCS customers and
the places in which they live:

- Professionals typically between 30 and 50 years old
- Relatively high educational attainment
- Moderate and slightly higher incomes
- Residents of neighbourhoods with high residential densities
- Proximate to commercial services
- Proximate to good public transit services

These are similar to findings by other researchers, though the particulars vary somewhat in different
contexts (e.g. Millard-Ball et al. 2005; Shaheen et al. 1998; Harvey et al. 2007). For instance,
subscribers in London were found to be disproportionately male, while the opposite was found in
Quebec, Canada. (Harvey et al. 2007; Dallaire et al. 2006) Subscribers, as noted above, are in general
observed to use SDCS cars relatively infrequently: (Cervero et al. 2006) report that SDCSs promote
‘judiciousness’ in travel behaviour. The mode seems to fill a niche wherein subscribers perform
relatively high-value, infrequent trips with an SDCS car, and different types of journeys by other
forms of transport. This pattern of demand provides mutual benefits for SDCS and public transport
operators to offer their services cooperatively, as is now done in several instances. (Shaheen and
Cohen 2006; Dallaire et al. 2006; Hilkevitch 2008) Further, it is noted that several of the
characteristics identified above (education and income levels in particular) are broadly consistent with the classic early adopters of innovations in general. (Meade and Islam 2006)

2.1.7. Contemporary SDCSs
At the time of writing there are in the range of 1.25M SDCS subscribers in 1,100 cities, 26 countries and five continents, with a combined SDCS fleet of 32K vehicles. (Sheen and Cohen 2011) There are c.170K/160K subscribers in the UK / Greater London, respectively. (Morgan 2011). Modern SDCS operators exist as both commercial and non-profit ventures in different contexts. Subscribers typically pay any monthly subscription fee and variable use charges via pre-arranged debit. Some organizations have per-day rates to support the excursion market, while others bill only at standard hourly rates to encourage rapid vehicle turnover.

Reservation management for many operators has evolved from pencil-and-paper in the early projects to RFID wireless vehicle access, internet and SMS (text messaging) reservations, and GPS-based fleet management. Subscribers typically are able to make reservations via the internet or their mobile phone, access vehicles with a proximity card, and be debited monthly based on usage determined by a combination of onboard-vehicle and back-office technology. Transaction costs are kept low as a reservation can be made in many cases without interaction with staff.

The greater value of using IT systems however is perhaps the facilitation of the consumer experience. The freedom and time/cost efficiency of self-service by subscribers is one of the key reasons SDCSs have become commercially viable, as the technology has become available at reasonable cost only in recent years. Likewise, SDCS networks require a high level of trust that one will have access to a shared car when needed. The use of advanced IT systems helps manage such user confidence, as subscribers receive immediate feedback when making a reservation on whether it is accepted or cannot be accommodated as requested. If one subscriber is anticipating returning a vehicle late, to the inconvenience of the renter with a subsequent reservation, the operator is able to contact the latter to attempt to minimise the inconvenience and typically the offending party would pay a moderate charge (according to a set schedule of penalty charges) to their counterpart as compensation.

A central tension of operating an SDCS is managing the competing pressures of meeting customer demand, operating high vehicle utilisation rates, and accommodating a growing base of subscribers. Vehicle availability can be extremely limited at times of high demand, such as holiday weekends. The empirical evidence shows that carsharing usage exhibits strong temporal peaks. Figure 2.5 shows the distribution of daily carsharing usage for the Canadian Communauto car club. Daily, weekly, and seasonal variance can be seen. Note that the graphic shows usage at a daily level of
temporal resolution; there is also, though it is not shown considerable heterogeneity in demand patterns through the course of a day. Of further interest, the graphic shows only completed reservations; the one to four percent of reservations which the system was unable to accommodate (which would presumably accentuate the peaks in the graphic) are not shown. (Communauto 2008)

Additionally, the graphic does not include the unknowable number of ‘discouraged reservations’ which were never attempted by users able to see vehicle [un]availability prior to formally requesting a reservation.

![Average number of transactions per member (2004)](image)

NB: The separate lines represent different groups of car sharers that the authors identified through cluster analysis of usage patterns. The upper curve represents a group the authors term ‘frequent users’, and the lower line represents ‘occasional users’.

Figure 2.5: Daily usage patterns of the Communauto car club for the year 2004 (reproduced from (Morency, Trepanier, Agard, Martin and Quashie 2007))

As noted above, following earlier experimental trials (Bendixson and Richards 1976, Honda Motor Company 2008) commercial services offering ‘one-way’ journeys as their standard service are currently expanding in both continental Europe and the US, with indications that cities within the UK may well find themselves host to such systems in the coming years. The parking nodes themselves (alternatively called pods, bays, or stations) tend to be relatively small-scale and dispersed throughout neighbourhoods, with the relative ease of accessing them being a key advantage over traditional hire car outlets.

(Robert 2005c) identifies a goal of a 5-10 minute maximum walking time for customers to reach a car club vehicle, and an acute difficulty obtaining parking spaces that would allow operators to meet this criterion. A shortage of dedicated parking for SDCS fleets is quite common in the more successful geographic markets; many SDCSs solicit private parking spaces on their websites. Some localities,
including the London boroughs, provide on-street parking spaces as concessions to SDCSs through competitive procurement processes. (Transport for London 2008)

In recent years, institutions have widely begun to make arrangements with SDCSs as a way to augment their institutional fleet or eliminate their fleet management responsibilities entirely. At present, these tend to be universities (e.g. University of California, Berkeley) or government agencies (City of Philadelphia) rather than commercial enterprises. (Millard-Ball et al. 2005) Many SDCS operators, however, experience relatively low daytime usage on weekdays. A potential solution which appears promising is to diversify through expanded take-up by business and institutional clients. (Shaheen and Cohen 2007, TFL 2008) Tackling the business-to-business market in a serious way can be expected to require the re-thinking of organisational structures and service offers.

The peer-to-peer form of SCDS is at present beginning to emerge simultaneously in North America and the UK (though it has yet to, at the time of writing, appear elsewhere). In this model, the SDCS serves merely as the market-maker that links private car owners with car drivers wishing to rent to them, as well as fulfilling a role of protecting either in case of a transaction not going smoothly. The key enabler of such services is the insurance product; ‘Whipcar’ in the UK for instance has secured a novel policy which protects car owners nearly-completely from events arising from renters’ usage. The growth trajectory of this service in the UK appears to be rather steep; they claim to have reached a fleet of 1,000 shared-cars within six months, a standard which the UK’s car club SDCS reached after a period of six years. (Huang 2010).

The UK has an active marketplace of such ‘traditional’ car club SDCSs, with no fewer than four major operators and fourteen smaller organisations at the time of writing. (Carplus 2011) The ‘Carplus’ charity administers an accreditation programme for car club SDCSs, which some local authorities use to screen operators for concessions. In order to be accredited, an operator must meet particular operating criteria, an element of which is an annual survey of subscribers.

In North America and the UK, car club SDCSs for the most part operate as independent businesses. This model is not universal, however. In Germany, multiple SDCSs operate under the ‘cambio’ brand name, and back-office functions with economies of scale are performed jointly. Operators associated with the network hold an annual meeting to make decisions which affect cambio’s operations or the brand as a whole, and subscribers are privy to network-wide operational data useful in benchmarking their operations. (Brook 2007) Another innovative business model which has also appeared in Germany is the integration of SDCSs with Deutsche Bahn’s intercity rail services; the *Flinkser service* is owned by DB. (Flinkster 2011)
2.2 Forecasting take-up of new products/services

It is argued that SDCSs fall within a category that the marketing literature terms *discontinuous innovations*:

Discontinuous innovations are new products or services that shift market structures, represent new technologies, require consumer learning, and induce behaviour changes...The potential customer, for example, has no direct experience of the new product, so the market researcher has no history to draw on and cannot use questions about usage and attitudes to determine demand.(Mackay and Metcalfe 2002)

Turrentine and colleagues describe the challenge of forecasting take-up of electric vehicles, a type of discontinuous innovation bearing some similarities to SDCSs:

The problem for those wishing to predict the market [for electric vehicles] is that consumers themselves cannot usually tell us with any certainty what they need themselves, seeing [as] they have even less experience than we do with electrics to assess needs and wants. (Turrentine et al. 1992)

The principal challenges to forecasting carsharing take-up are adapted from (Urban et al. 1994), which describes the challenges to forecasting demand for any type of 'really-new product':

- **Diffusion of Information**: A 2006 survey in London found that only one-third of non-subscribers were aware of car club SDCSs, and a similar 2007 survey in Paris found that only eight percent of the population knew of SDCSs and had a basic understanding of the service features. (Harvey et al. 2007; Jemelin and Louvet 2007) While awareness of SDCSs is growing, we remain far from the point where all potential subscribers are familiar with the concept.

- **Evolution of Technology**: Most operators' systems are moving towards automation supported by mobile IT technology, but some successful car club SDCSs continue to use manual operating methods. This, as well as many other aspects of SDCS operating models can be expected to change in coming years. We can reasonably expect that competing mobility products and services will be concurrently evolving as well.

- **Discovery of New Uses**: As consumers incorporate a new product into their routine they naturally employ the innovation in ways that suit their needs, which may not align precisely with those envisioned by the product developer. When the SMS standard for mobile communication was being designed, for instance, it was not foreseen that consumers would use SMS for such applications as vehicle fleet tracking or ringtone sharing. (Le Bodic 2005)
Reduction in Price from High Initial Levels: Many investments required for SDCSs (e.g. IT systems for facilitating operations) have inherent economies of scale, which generally implies lower average fixed costs as an operator’s subscriber base and usage levels grow. In addition, the price of the specialised fleet management technology can be expected to trend downwards (and/or improve qualitatively) over time. While other cost elements do not have similar structures (e.g. vehicle fuel costs), one can plausibly expect that the price of carsharing may fall in coming years in relative terms.

Growth of an Infrastructure: A recurring complaint from subscribers relates to system reliability. At peak times such as summer weekends vehicles may be fully booked many months in advance, and there is evidence of clients making strategic use of cancellation policies when demand is heavy. (Theriault 2008) However, improved infrastructure and increasing fleet sizes in future are likely to have the effect of increasing the reliability that subscribers experience. A further aspect of the growing SDCS infrastructure is that several automobile manufacturers either currently have assembly lines producing models dedicated for the SDCS market or have stated publicly they will do so. (Belson 2009)

Entry of Competition: From its beginnings as a commercial service, the SDCS market has attracted financing from some traditional car hire firms. In the past few years, however, as SDCSs have begun to appear as a mid-term threat to disrupt the prevailing industry structure and standard practices, car hire firms have begun to modify their service delivery to more closely match the convenience of other SDCSs. In addition, other players within the broader automotive and public transport industries have begun to experiment, at substantial expense, with SDCS. The immediate trend is one of consolidation (Davis 2010), though entry by new service providers may occur with time.

Though it is a challenging task, the literature offers a number of suggestions for forecasting the demand for a ‘really-new’ product like SDCSs. Generally-speaking, a sample of target consumers is provided with a good understanding of the product and its context, and stated-response techniques are employed to elicit various types of information from participants on whether they are inclined towards the product.6 Research in the 1990’s into take-up of electric vehicles (Turrentine et al. 1992; Urban et al. 1994) and SDCSs (Shaheen 1999) provides background on techniques employed in similar contexts.

The study reported by Turrentine et al. centred on a simulation technique titled PIREG (Purchase Intention and Range Simulation Game), which itself is based on the earlier gaming simulation CUPIG (Car Use Patterns Interview Game). (Turrentine et al. 1992; see Lee-Gosselin 1990 regarding CUPIG.7)
Participating households were asked to record their members’ (and their cars’) itineraries in a week-long activity-travel diary. After completing the diaries, the drivers in the respondent household were asked (with all drivers present at the interview, to facilitate joint activity scheduling, vehicle swapping, etc.) to adjust their activity-travel patterns to be consistent with the operating limitations of replacing one of their cars with an EV. Only after gaming out the strategies for dealing with the EV’s charging requirement and reduced range, in the context of their activities of the past week, were they queried for their likelihood of purchasing an EV. Interestingly, however, researchers have found (e.g. Raux 1997) that people in some instances make choices that affect the structure of their mobility needs (in this instance moving home) without taking in advance a broad view of the implications on their lifestyle.

2.3 Discrete choice analysis

The foundations of economic theory rest, in large measure, on notions of people acting in a self-interested manner. People’s behaviour related to producing goods to be traded in the marketplace, for instance, is classically analysed as if they were maximising income or profits within the set of constraints and resource limitations which describe their circumstances, though more modern extensions of production theory account for subtle variations.

Quantifying people’s behaviour as consumers requires rather more abstract concepts as in this role, in contrast to that of a producer, a person does not accumulate a currency which is tangible, traded in a marketplace, and has an agreed value. Indeed, consumption in economic terms implies the conversion of something of value in a [public] marketplace into the consumer’s [private] welfare.

Classical economic theories of people’s behaviour as consumers introduce the term utility, which in its standard form is defined as an arbitrarily-scaled and unit-less scalar measure onto which a consumer’s preferences are mapped. (McFadden 2000) In a situation where a consumer faces several ‘discrete’ options from which to choose, the utility metric is defined such that the most-preferred option has the largest level of utility, the second-preferred option has the next-largest level of utility, and so on. Utility is an ordinal, rather than cardinal, scale in this interpretation\(^6\). In other words the ordering of options on the scale of utility matters, whilst the magnitude of the differences in the level of utility between options do not. By definition a consumer prefers, in a deterministic manner, options which would provide him with larger levels of utility to those which would yield smaller levels of utility.

It is typically assumed that the utility which competing options would provide a consumer is in some sense known to them, but that these are latent quantities – not directly knowable – from the ex post
outside perspective of a researcher analysing the consumer’s behaviour. By appealing to the
definition of utility, however, a researcher who is able to observe the expression of a consumer’s
free will in a given situation, together with observations of other available options which the
consumer could have chosen instead but did not, can use this information to infer that the option
which the consumer chose must have provided him or her with a higher level of utility than all other
options.

Having structured her view of the situation in this manner, with attendant assumptions about the
qualitative nature of the circumstances in which the consumer is acting (often referred to as a choice
context) and the consumer’s knowledge of the available options which they did not choose, the
researcher is in a position to explore statistical relationships between the features of each of the
consumer’s options, the characteristics of the consumer himself, and the relative utility which each
of the consumer’s options would yield him. This structure is termed a discrete choice model [DCM].

The researcher’s imperfect knowledge about how the consumer perceives the various options
available to him is taken into account by way of Random Utility Maximisation [RUM] theory, which
posits that the utility that an option would provide to a consumer can be decomposed (usually, but
not necessarily, in an additive manner) into a systematic and an error component:

\[ U_{ni} = V_{ni} + \epsilon_{ni} \]  \hspace{1cm} (2.1)

In equation 2.1 \( U_{ni} \) is the utility which option \( i \) would provide to consumer \( n \), where \( V_{ni} \) is the
systematic portion of utility and \( \epsilon_{ni} \) is the error term. The systematic component comprises the
statistical inferences which the researcher is able to draw from the available ex post information
which describes the choice context.

The systematic component is typically decomposed into an additive functional form with a vector of
one or more unknown parameters [normally denoted as \( \beta \)] multiplied with an equally-sized vector
of ‘known’ information \( [X] \) describing the consumer and the options available to him:

\[ V_{ni} = \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_k X_k \]  \hspace{1cm} (2.2)

As the researcher can ‘observe’ the systematic component of utility that any option \( i \) provides
customer \( n \), but not the \( \epsilon_{ni} \) error term, she cannot deterministically predict option \( i \)'s utility \( U_{ni} \)
with certainty. She can, however, predict the option that the consumer will choose in a probabilistic
manner, which will depend on her assumptions about the nature of the \( \epsilon_{ni} \) terms.
The researcher has wide latitude in making distributional assumptions regarding the [stochastic] error term in Equation 2.1. (Ben-Akiva and Lerman 1985, Train 2003) The most common assumptions lead to the logit DCM form:

Assumption 1) All of the \( \varepsilon_{ni} \) error terms are *identically* distributed, meaning that their variances are the same

Assumption 2) All of the \( \varepsilon_{ni} \) error terms are *independently* distributed, meaning that they are uncorrelated

Assumption 3) All of the distributions of the \( \varepsilon_{ni} \) error terms are of the Gumbel form

In addition to these assumptions, the researcher must also specify the scale of utility by selecting a single scalar value for the variance of the \( \varepsilon_{ni} \) error terms. This selection is arbitrary as the scaling of utility is unit-less, though in the interest of simplicity it is typically set to a convenient value such as one.

These assumptions lead to the logit DCM form of the probability of consumer \( n \) choosing option \( i \) from set \( C_n \) which contains all \( j \) options available to consumer \( n \), including option \( i \):

\[
P_{ni} = \frac{e^{\nu_{ni}}}{\sum_{j \in C_n} e^{\nu_{nj}}}
\]  

(2.3)

Discrete choice models of the logit and related forms (termed *Generalised Extreme Value* [GEV] models) have been widely applied, particularly within the transport domain, to analyse choice contexts such as choice of travel mode, route choice, activity location choice, residential choice, etc. Their attractiveness appears to arise from their ability to accommodate situations where consumers face discrete and qualitatively-distinct options, computational ease (relative to alternative DCM forms such as ‘probit’), and the flexibility they offer by way of relaxing the restrictive distributional assumptions regarding error terms.

### 2.3.1. Portfolio choice

As discussed above in Section 2.1.6, the literature suggests that SDCSs tend not to dominate a person’s travel as owning a car may do. It is suggested that that they may instead fill a ‘niche’ of a person’s travel, and thus be a viable option in cases where the remainder of the person’s travel can be completed to their satisfaction by other non-personal-car methods of transport. Later portions of this thesis describing the original research undertaken here provide concurrence for this view, leading to a proposed analytical form in which some of a person’s mobility decisions are viewed as
the composition of a ‘portfolio’, rather than the more-typical form in which independent choices of unitary elements are made.

Such ‘portfolio’ choice analysis is well-established outside the transport domain (Elton et al. 2009, Dube 2004). The focus on the determinants of portfolio choice within the transport field of study, however, is in its infancy, reflecting the broader ongoing development of subtler and more flexible forms of quantitative choice analysis in the domain. (see Hess 2005 for a discussion)

A form of preference inference termed the Multiply Discrete-Continuous Extreme Value [MDCEV] model, discussed in detail in Section 4.4.1, has recently been used to analyse situations in which a choice maker composes a portfolio from a set of elemental options, where each discrete elemental option is also characterised by a continuous quantity. (Bhat 2008) For instance, one recent application was a study attempting to explain the [discrete] number and type of automobiles/light trucks a household owns and the [continuous] distance that each is driven. (Bhat and Sen 2006) In essence, the choice-maker is specified to act in a situation in which they may choose zero or any positive amount of a number of continuous classes of quantities. All discrete elements for which they are simulated to choose a non-zero amount from the continuous dimension are simulated to be a part of the choice-maker’s ‘portfolio’.

Another recent application of the MDCEV analytical structure analyses how households might adjust their portfolio of spending on transport and other classes of expenditure in response to structural increases in fuel price. (Ferdous et al. 2010)

A further instance of exploration of ‘portfolio’ choice contexts within the transport domain is recent work by Wiley and Timmermans, who argue that ‘portfolio’ choice specifications may be more widely-applicable within the transport domain than previously realised. The authors raise a number of issues associated with techniques to explicitly accommodate portfolio choice within discrete choice analysis methods. (Wiley and Timmermans 2009)

2.3.2. Strategic-tactical analysis

Later portions of this thesis report the finding that people’s engagement with SDCSs can fruitfully be considered to involve both strategic and tactical dimensions, in which the strategic dimension of choice is some function of a person’s expectations across future tactical choice occasions.

One of the definitions of strategy found in the Miriam-Webster dictionary is ‘the art of devising or employing plans or stratagems toward a goal’. It is in this sense that the term is employed in this research, in which a strategy is conceptualised as a broad plan, whilst ‘tactics’ are actions for which a
person’s range of options is constrained by their earlier selection of strategy. A strategic choice, in this definition, is one in which outcomes of future linked tactical choices are taken into account.

The term strategy is closely associated with military and business applications. (e.g. Porter 1987, Radford 1977). There is a tradition of transport researchers specifying the choices that people face to be ‘multi-dimensional’ and thus analysed jointly, such as a person’s choice of where to go and how to get there, (‘mode/destination choice’: Ben-Akiva and Lerman 1985) or of how to go and what time to leave (‘mode/departure time choice’: Bhat 1998). A relatively small number of studies within the transport domain, however, have taken an approach that takes account of possible strategic/tactical behaviour. An example is Ben-Akiva and Bowman (1998), where the authors attempted to specify people’s residential location choices as a function of people’s daily activity patterns; though the authors report that a traditional work-trip-based provided a better fit to the observed patterns in the estimation dataset.

A study by Train (also discussed below in Section2.5) included a link between simultaneous analyses of car ownership and mode choice for commute journey – one of the explanatory variables for the attractiveness of owning a car was specified to be its value to perform one’s commute-to-work relative to other options such as public transport. (Train 1980) A more recently-reported study by Choudhury investigated drivers’ lane-changing behaviour, and found evidence to suggest that drivers engage strategies to reach their ‘target’ lanes. (Choudhury 2007)

A number of studies have investigated ‘choice set generation’, in which people are specified to face a two-stage choice process where the second stage is the classical ‘discrete choice’ described in Section 2.3. (c.f. Swait and Ben-Akiva 1986, Thill 1992, Ben-Akiva and Boccara 1995, Cascetta et al. 2002, Meyer 1979) The novelty is that the set of options from which the choicemaker is specified to select is ‘generated’ in the first step. In cases (e.g. Ben-Akiva and Boccara 1995) where the composition of the choice set depends on the attractiveness of the options in the ‘universal’ choice set, there can be said to be a ‘strategic’ aspect to choice set generation techniques.

2.4 Stated-response methods

As is described later in this document, the trajectory in which this study proceeded included a type of semi-structured and mostly qualitative technique referred to as ‘gaming-simulation’ [GS], which was subsequently followed by a more-structured stated-choice [SC] survey.

In the overview prepared in Lee-Gosselin 1995, both SC and GS methods fall within the broad category referred to as ‘stated-response’ [SR] techniques. Table 2.3 lists several other types of SR
techniques; stated-tolerance methods, for instance, have been the focus of much attention by environmental economists using a particular form known as contingent valuation.

<table>
<thead>
<tr>
<th>BEHAVIOURAL OUTCOMES</th>
<th>CONSTRAINTS (expressed as attributes: personal/household/social/spatial/supply, etc)</th>
</tr>
</thead>
</table>
| STATED PREFERENCE    | Mostly given
| (focus = tradeoffs, utility) |
| "Given the levels of attributes in these alternatives, which would you prefer? [A]...? [B]...? etc..." |
| STATED TOLERANCE     | Mostly elicited
| (focus = limits of acceptability, and thresholds for change) |
| "Under what circumstances could you imagine yourself doing: [C]...? [D]...? etc..." |
| STATED ADAPTATION    | Mostly elicited
| (focus = reactive and trial behaviour, problem-solving, rules) |
| "What would you do differently if you were faced with the following specific constraints: [E]...? [F]...? etc...?" |
| STATED PROSPECT      | Mostly elicited
| (focus = learning processes; information seeking; the imaging, formation and testing of choice-sets; meta-decisions) |
| "Under what circumstances would you be likely to change your travel behaviour and how would you go about it: [G]...? [H]...? etc...?" |

Table 2.3: Taxonomy of ‘stated-response’ survey methods (reproduced from Lee-Gosselin 1995)

Most GS techniques fall into the Stated Adaptation quadrant of Table 2.3. They generally ask respondents to indicate, in an open-ended manner, how they would adapt some part of their real-world behaviour in response to a stimulus. In-game tools are used to ensure that the stated adaptation is consistent with the constraints the respondent would have faced in a real world situation, usually one that is based on a very recently observed period of revealed behaviour. Such techniques are thus particularly well-suited to situations where one is investigating behaviour that is not fully understood or well-practised, or may involve a range of responses not straightforward to accommodate with a small list of a-priori-developed options.

SC methods, by contrast, are characterised by the structured nature of the stated responses that are requested from him/her. Rather than being presented with constraints and asked to adapt, in an SC exercise the respondent is asked to indicate a choice from a limited set of options.

The proposed design for the survey reported here (which we term the Advanced Vehicular/Activity/Travel And Resource [AVATAR] survey) involves a choice context of rather high complexity; as Richardson (2001) notes, however, the literature on the reliability of respondent’s stated choices tends to focus on the breadth of the information load (e.g. the number of alternatives presented, the number of attributes of those alternatives, etc.) rather than the depth (complexity) of the choice task. The author further notes that one of the great challenges of complex survey design pertains to the analysis of the data afterwards; as the intensity of the cognitive task increases it may be less plausible to assume that people have engaged in choicemaking consistent with utility-maximisation (the predominant theoretical framework underlying SC methods).
Bates et al. (2001) report the design of a particular type of stated-response survey in which respondents are presented with a ‘choice set’ where the outcomes are probabilistic; in typical applications the outcomes of the choice are presented as deterministic. In this case, the authors proposed a visual representation for the probability distribution of actual arrival times.

The proposed AVATAR survey design includes both ‘strategic’ and ‘tactical’ dimensions of choicemaking. Though this is a novel design element, researchers have very recently begun experimenting with designs in which respondents are asked to consider both day-to-day and broader mobility choices. (Weis et al. 2010) In this particular study, the researchers asked each respondent to consider ‘tactical’ (mode choice) and ‘strategic’ (car type and public transport ticket type choice) choices separately, rather than jointly in a single choice context. Axhausen and Scott (2006) reported an SC design in which the choice context has a ‘strategic’ dimension, with an implicitly-linked ‘tactical’ one. Erath and Axhausen (2010) report a stated-choice survey where the ‘tactical’ level of choicemaking is aggregated into a continuous choice of car-driving mileage and public transport mileage.

Much research advancing the state-of-the-art regarding SR design principles has originated in Australia in recent years. Rose and colleagues have, for instance, reported on a set of techniques that fall within the rubric of ‘efficient’ design methods. These techniques comprise a variety of strategies aiming to maximise the amount of statistically-useful information obtained from any given sample size (or alternatively to minimise the required sample size which can be expected to provide parameter estimates meeting a target level of statistical reliability.) (Rose et al. 2008, Bliemer and Rose 2011, Scarpa and Rose 2008, Bliemer et al. 2008) A certain degree of tension exists between statistically-optimal survey design and the plausibility of the choice context that is presented to the respondent, as ‘efficient design’ theory does not take into account how the respondent relates to the choice context that they are presented; Hensher refers to the latter as ‘cognitive efficiency’. (Hensher 2006) One such strategy for managing the plausibility of stated-response choice situations, to which much attention has been directed in recent years, is to ‘pivot’ the choice context around some aspect of the survey respondent’s real-world experience. (Kreitz et al. 2000, Bradley and Daly 2000, Hensher 2006, Train and Wilson 2008.) A related strain of research has explored methods of identifying respondents engaging in non-utility-maximising behaviour such as lexicographic behaviour, or respondents who interpret the instrument fundamentally differently than intended by the researcher. (e.g. Hess et al. 2010)

Computer-aided instruments, in particular web-based ones, have unlocked previously-infeasible possibilities in stated-response design. In addition to providing a new medium for connecting
researchers with respondents, the capabilities to adapt the instrument in real-time during an interview (Fowkes 2007, Richardson 2002) offer new freedom to researchers, as does allowing the respondent to seek information from databases in-game. The Mobiplan instrument, for instance, allowed respondents to learn about how their activity-travel options would vary in different places, by drawing information from databases that is tailored to their individual activity-travel patterns, and which can be modified either systematically (in a stated-response mode of operation) or by the user (in an information-seeking mode). (Kreitz et al. 2000) Computer-aided designs also enable recording of the respondent’s pattern of engagement with the exercise, such as the tempo at which they move through it, or their style and amount of interaction with the instrument leading up to their stated responses.

2.5 Analyses of car ownership

It is proposed later in this document that whether or not a person might subscribe to an SDCS be analysed as part of a broader sort of choice where car ownership is an option, as is having only casual access to automobility (such as taxi services). This section thus reviews the body of literature that has developed regarding people’s ownership of cars.

Early academic enquiries (e.g. Chow 1957, Mogridge 1967) investigated car ownership at the level of aggregate demand, in other words aiming to understand how many cars in sum would be owned by a given group of people. Tanner, for instance, took such an approach in a series of studies during the late 1970s, where the focus was principally on identifying the growth trajectory that private car ownership might take over time. (Tanner 1978, 1979, 1980) As with other sort of analyses relating to travel demand, the state-of-the-art migrated to disaggregate techniques (i.e. where individual people or households are the unit of analysis) over time. Hensher and colleagues reported an important study in the early 1990s which made use of a discrete-continuous analytical system (car ownership and mileage) to analyse panel data. (Hensher et al. 1992)

de Jong and colleagues prepared a broad review of analytical practices related to car ownership in a study published in the early 2000s, which by then had evolved, in response to the changing policy agenda and increased computing power, to take in issues such as fuel type, emissions level, and timing of vehicle transactions. (de Jong et al.2002) Roorda and others reported agent-based (i.e. disaggregate) analyses in which car ownership is one in a broader sequence of choices that people face related to their mobility. (Roorda et al. 2009, Ciari 2010, Eliasson and Mattson 2000). Huang showed, using a quasi-panel form of data, that cohort effects could be identified with regards to car ownership – in other words that people’s ownership of a car at any juncture seemed to relate to the
path of their life course, in addition to the point along it in which they find themselves at any given moment. (Huang 2007)

The techniques which have predominantly been employed to analyse car ownership have treated the underlying demand for cars to be a function of apparently-exogenous observations such as a person’s income, the type of neighbourhood in which they live, whether they are employed, the distribution of employment opportunities across a region, etc. In a seminal paper on consumer theory in 1966, by contrast, Lancaster argued that a person’s demand for a consumer good can be decomposed into a function of ‘the properties or characteristics of the good’, in particular the ways in which the good combines multiple characteristics in a single package. (Lancaster 1966) Whilst it is relatively uncontroversial to posit that a number of relevant properties associated with automobiles relate to the provision of mobility – of oneself, one’s family, cargo, etc. – researchers have also pointed out the relevance of less-evident characteristics. Turrentine and colleagues for instance, present evidence, amongst early-adopters of hybrid petrol/electric vehicles, of car buyers’ motivations also including issues of ‘personal identity’ and the public conveyance of messages about oneself to others. (Kurani et al. 2007)

Early attempts at representing Lancasterian notions in analyses of car ownership include Ben-Akiva and Atherton 1977 and Train 1980, though in these cases limitations of data and computing resources constrained the researchers to specifying the demand for car ownership as a function of a single commuting journey. In very recent years (e.g. Salon 2006, Pinjari et al. 2008, Dissayanake and Morikawa 2010) researchers have begun to experiment further with techniques that treat a person’s car ownership as a function of the value a car would provide to perform a specific need for mobility, typically a commuting journey, though these recent studies also use a single archetypal travelling need to encapsulate the mobility value of owning a car.

In the past decade, researchers have also begun to explore the inter-relationships between owning a car and owning other durable travel products/services. (Simma and Axhausen 2001, Scott and Axhausen 2006, McElroy 2009, Weis et al. 2010, Vovsha and Peterson 2009) Scott and Axhausen, for instance, specified people’s holdings of both cars and public transport season tickets to be linked choices with the possibility for correlated error terms. (Scott and Axhausen 2006)

2.6 Measures of accessibility
This research proposes a new measure of personal accessibility – accessibility being the ability to access life opportunities. In a traditional sense accessibility refers to corporeal accessibility – the ability to transport oneself to/from locations which, at particular times, offer the possibility of
participating in life activities (e.g. shopping, employment, caring, education, socialising, etc.)

Technological developments – the post, telephone, internet, card payment, and mobile telephony / computing amongst them – have somewhat mitigated the value of physical presence for some types of human activities. On the basis of evidence of today’s young people reporting a mobile phone to be a higher-priority ‘must-have’ than an automobile, it may well be that this process has accelerated in recent years. (Motavalli 2010) Regardless of whether this short-term trend proves to be sustained over time, it does appear that despite the role played by other-than-corporeal mobilities – Urry defines and discusses object, virtual, and imaginative forms of travel – the ability to deliver oneself to particular places at particular times will remain desirable or imperative for many types of personal activities. (Urry 2002)

Hagerstrand introduced the fundamental concept of a time-space prism, which in principle sets the boundaries, and the time-space volume contained within, which can be reached by a person able to travel at a given fixed rate of speed and having a given amount of time before returning to their starting point. (Hagerstrand 1970) A closely-linked term is the action space in which a person operates, defined by Dijkstra as ‘the area in which a person undertakes activities, or could if they wished.’ (Dijkstra 2004) Schonfelder and Axhausen (2004) review related concepts such as awareness space, perceptual space, and mental maps, which introduce people’s individual life experiences and cognitive capacities. Recent developments include efforts to take into account multiple people’s time-space constraints in the case of activities undertaken jointly (Neutens et al. 2008)

The extension of these ideas to account for heterogeneity of travel speeds (in a world where much transport takes place on differentiated networks) leads to concepts of accessibility such as ‘the number of employment sites accessible from a given home location within a 30-minute commute’. Practical applications frequently involve analysts investigating the aggregation or patterns of such metrics across a region or population of interest. (e.g. GLA 2004, Soberman 2001, MTC 2005)

The activity repertoire concept bears similarities to the view of accessibility discussed later in this thesis. Axhausen (2002) defines it as ‘the set of activities (specific action streams at particular – types of – locations and times) which a person knows of, or has performed in the past.’ This differs from Hagerstrandian notions of accessibility in that time-space is viewed discontinuously, rather than as a homogeneous field in which opportunities are distributed.
2.7 Summary

This Chapter reviewed the literature regarding several aspects of this research, starting with the substantive context (subscription drive-it-yourself car services) before turning to a number of methodological issues.

The remainder of the thesis presents the original research undertaken on this study. Chapter three describes an application of gaming-simulation techniques aimed at understanding how consumers would engage with subscription drive-it-yourself car services, which informs the analytical framework presented in Chapter four.

Notes

1 SEFAGE is an acronym for the German name ‘Selbstfahrergenossenschaft’ which Robert translates roughly as ‘club of drivers’ (Robert 2005b)

2 See Table 2.1 for a listing of research products of these demonstration projects, and Section 2.1.1 for a brief description of their operating characteristics.

3 The control group consisted of people who expressed interest in subscribing to the SDCS but were unable to do so as the initial area of geographic coverage did not serve their residential neighbourhood.

4 The question regarding ‘alternative futures’ was worded:

   In total, how many miles (annually), do you think you would put on these vehicles [vehicles respondent previously indicated their household would acquire if SDCSs suddenly disappeared from their region], if you had to use them instead of carsharing?

5 See (Barth and Shaheen 2002) for detailed descriptions of alternative SDCS operating models. Station car systems are a form of SDCS that are oriented around public transport stations.

6 This is assuming that the time and/or cost required for real-world trials is infeasible.

7 The technique was introduced in the late 1970s by researchers at Oxford University, in a form known as the Household Activity Travel Simulator (HATS). See Jones 1979.

8 The branch of the economics discipline known as ‘subjective well-being’ or ‘happiness’ economics challenges this interpretation of utility as a non-cardinal measure. (Easterlin 1974, Abou Zeid 2009)
9 The term *stated-choice* is used here, though *stated-preference* is also encountered in the literature (and is found in Table 2.3).

10 Readers generally interested in SC methods are referred to Louviere et al. 2000; this review is primarily focused on the principles and design of complex SC instruments.
Chapter 3: Gaming-simulation task

The previous chapter presented a two-part overview of the literature; on the one hand of the state of knowledge relating to subscription drive-it-yourself car services (SDCSs), and on the other of the specific methods which relate to the research reported here.

This Chapter begins the presentation of the original research performed on this study, which proceeds through the remainder of this thesis. It reports on the initial investigation into how people might engage with SDCSs, which employed a technique known as gaming-simulation (GS).

Section 3.1 presents the motivation for the GS task, with Section 3.2 outlining the specific methodology of this application. The results, which led to the research reported in subsequent chapters, are presented in Section 3.3. Section 3.4 concludes the Chapter with a discussion of the challenges identified through the course of this task, which lead to the proposed methods that are presented in subsequent chapters.

This Chapter draws heavily from the following articles:


3.1 Motivation

Ceteris paribus, data observed from people’s real-world actions are generally accepted as more reliable than stated-choice (SC) data gathered from hypothetical situations for drawing inferences about real-world behaviour. (Ortuzaar and Willumsen 2004) An SC survey may, however, be appealing in situations where:

- new (or radically different) options are to be introduced into an existing marketplace
- a market in an entirely new sort of product or service is to be created
- sufficient information on the attributes of the options is not known or readily inferable
- attributes of different options are highly collinear (such as in the case of quality and price) or have low variability across different options
In such cases, as indeed in the context of SDCSs, revealed-choice data may be either unavailable or unsuitable. (Cherchi and Ortuza 2006)

Whilst the researcher has [near-]complete freedom with respect to the principal design issues of an SC survey – sampling method, definition and presentation of the choice situation, the nature of the choice set of alternative options as defined by the presence/absence of specific options and the options’ attribute levels, whether alternative options carry qualitative labels, the number of replications, any socio-demographic, attitudinal, or debrief questions, etc. – the credibility of any inferences drawn from the survey datasets ultimately rests on whether respondents can be thought to have made choices in a behaviourally-realistic manner. (Richardson 2001) This in turn rests on issues such as whether the design of the choice situation is well-understood by respondents and plausible to them, and whether the ‘response burden’ is commensurate with their ability and willingness to process the information presented to them.

When preparing to investigate novel forms of behaviour, researchers may have little knowledge upon which to base the framing of choice situations that are to be presented to respondents. Qualitative research techniques (e.g. focus groups, in-depth interviews, gaming-simulation, etc.) may be employed in such circumstances to draw inferences as to how people view the substantive situation, which in turn inform how the researcher may design more-structured analytical tools. (Lee-Gosselin 1995, Louviere et al. 2000)

### 3.2 Gaming-simulation methodology

In designing the qualitative research, a number of aspects of how novel mobility services might be viewed by the prospective interviewees were considered. For the interviewees of such a research task, it is quite possible that the car club SDCS participation issues we explore are:

- **Hypothetical** – As with any stated response exercise, we pose hypothetical situations to respondents.

- **Unfamiliar** – For the subset of survey participants who are not (and have never been) SDCS members, joining and using a car club may well be an unfamiliar set of circumstances. It is possible that many potential members are unaware of them, or unfamiliar with the service features. (Jemelin and Louvet, 2007)
- **Interdependent with other aspects of one’s activity-travel pattern** – Research has suggested that car club participation is coupled with other elements of people’s activity-travel patterns. (Cervero et. al., 2006; Dallaire, et. al., 2006, Adamou 2011)

- **Interdependent between different people** – As with car ownership and use, it is plausible that SDCS participation may affect those beyond the immediate driver, such as other household members. Mechanisms for such interdependence might include travel as a passenger in an SDCS vehicle, or SDCS travel which indirectly affects them due to intra-household shared resource behaviour or shared activities.

- **Considered across different timescales** – One’s decision to join or leave an SDCS might be inspired by any of a wide set of potential motivations. (Meijkamp, 2000) In any case, whether to be a car club member or not may well be considered on a different timescale than a decision to use an SDCS vehicle to access a particular activity on a given day.

In view of these issues, gaming-simulation techniques were chosen for the qualitative research. GS techniques in general involve collecting a revealed base of behaviour from interviewees, which then frames a subsequent in-depth interview on the topic of interest. (Lee-Gosselin 1995) In the transport domain, this is typically done with the use of activity-travel diaries. During the in-depth interview, interviewees are then presented with perturbations which either relax or tighten the constraints on their revealed activity-travel behaviour, to which they respond as they see fit (i.e. without the constraint of a ‘choice set’ of defined options, as in an SC survey). They are asked to consider their likely response(s) in the context of their diary period, and interact with instruments which tie their in-game actions to the real-world consequences they would entail. The interactive nature of the interview helps to identify the range of responses to the stimulus, and ensure that interviewees’ intentions are consistent with their spatio-temporal needs and with the major opportunities and constraints in their lives.

The pedigree of the GS methods which were employed is found in research into activity-travel behaviour undertaken in the 1970’s. (Jones 1979) Such procedures were developed to assist in specifying quantitative models for assessing the effects of a given policy proposal. In the taxonomy developed in (Lee-Gosselin 1995), GS techniques are suggested to be appropriate in situations in which it is desired to provide a respondent with ‘freedom to act under future conditions’, as interviewees formulate their responses in an open-ended manner. This contrasts with the more-common (though not completely predominant) methods employed in stated-choice surveys of travel behaviour in which the interviewee is presented with a set of responses, each defined by a small set of attributes, from which he or she is asked to choose.
Notable examples of the use of GS techniques within the transport domain include the market for electric vehicles, changes in transit service provision, and energy shortages. (Turrentine et al. 1992; Doherty et al. 2002; Jones 1979) They were employed on the present research by virtue of their dual functionality. First, they provide the desired grounding of interviewees’ responses in the reality of their recently-experienced diary week, yielding useful insights into the robustness or fragility of their mobility patterns. More importantly, however, they are ideal for framing the subsequent discussion of car club SDCSs within the broader context of people’s overall mobility needs.

During the research development phase, it was decided to use households as the unit of analysis in the GS task, based on the small body of literature documenting the complexity of car club SDCS subscription and associated secondary impacts. (Cervero et al. 2006) Hence, all adult household members completed activity-travel diaries and participated in the in-depth interview.

The key instruments in this task were:

- intake forms, collecting socio-demographic, life trajectory, and other related data
- 7-day activity-travel diaries
- Single-A4-sheet summaries of each interviewee’s activity-travel diary (prepared by the interviewer in advance of the in-depth interview)
- acetate overlays of the summary sheets

Feedback from the first households to complete the survey process was taken into account in revising several of the instruments.

Households were recruited from amongst Imperial College London staff, members of the public, and car club SDCS subscribers, and were provided gift certificates worth £20 for their efforts. Interviewees lived in London’s inner suburbs. The effort for each interviewee household implied a quite small sample size. For these reasons, the sample was not intended to be representative of the population at large. Accepting these limitations, the sampling protocol was designed to encompass different segments of the potential market for car club SDCSs:

- Segment 1: Car club SDCS subscribers (non-car-owning) (3 households took part)
- Segment 2: Non-subscribers, non-car-owning (2)
- Segment 3a: Non-subscribers, car-owning, ‘heavy’ drivers (2)
- Segment 3b: Non-subscribers, car-owning, ‘light’ drivers (1)

Following recruitment of interviewees, an intake interview was held to gather the household’s background information, and to distribute the activity-travel diaries. Then, after household
members completed their diary week, each person’s activity-travel diary was summarised onto a single sheet of paper.

The in-depth interview was held shortly after completion of the diary week. The interview began with a brief review of the interviewees’ activity-travel behaviour during the diary week. The interviewees were then presented with a series of scenarios, each of which consisted of perturbations to their household’s reported activity-travel behaviour during the diary week. The scenarios are generally structural in nature, rather than one-off, and are tailored to each household’s particular circumstances. The scenarios are designed to be clear and plausible, but not necessarily very probable, given each household’s circumstances.

Each scenario was designed to build on the prior one, rather than being a fresh ‘choice situation’ – if, for instance, an interviewee’s car was taken away in one scenario, this continued in future scenarios unless stated otherwise.

The presentation of the SDCS included a brief description of a typical service model presently operating in London. Interviewees were shown a promotional brochure for a London-area car club. The description contains some prospective service features that are in the early deployment stage commercially, but which are expected to become more widespread. For instance, interviewees were told that the SDCS vehicles can be accessed by swiping their existing public transport smart card. Interviewees were told that there is a 1 in 20 chance that a car club vehicle would not be available to reserve from their neighbourhood service point when they want it.

As the interviewees considered each scenario, the interviewer ensured that potential responses they raised were considered within the context of the events, opportunities, and constraints of their diary week. The interviewer facilitated the discussion by presenting each scenario, providing transport information as requested (such as maps), and asking probing questions regarding interviewees’ potential responses.

When the interviewee(s) had settled on a response to a particular scenario, they noted the changes with coloured markers on an acetate overlay over the summary sheet of their weekly activity-travel behaviour. Each member of the household prepared a single acetate for each scenario, which then remained as a physical artefact of their response.

At some point during the course of the scenarios, non-car-club–subscriber interviewees are presented with the opening of a car club SDCS in their neighbourhood, and a brief explanation of the service model. Interviewees who are presently car club SDCS subscribers are presented with at least
one scenario in which the service is unavailable. Non-car-owning interviewees were also presented with at least one scenario in which they were given a car for free by a relative.

The interview then proceeded to an open-ended discussion. This portion of the interview made use of the GS exercise to frame a discussion of interviewees’ broader mobility-related decisions—such as getting licensed to drive, acquiring or disposing of a car, moving home or job, purchasing a public transport season ticket, and subscribing to a car club SDCS. This part of the interview explored how interviewees related day-to-day and broader mobility decisions. This was included in the research design to investigate whether, and if so how, people view linkages between subscribing to a car club SDCS and using an SDCS vehicle. If interviewees did not choose to use an SDCS vehicle at any point, the types of service attributes which might lead them to consider subscribing to such a service were explored.

The interview then concluded with a ‘debrief’ stage, in which the interviewer began to pack up the interview materials whilst discussing the GS exercise, with an aim of setting an informal atmosphere. The interviewer paid particular attention to comments that reveal assumptions interviewees had made about the purpose of the research, or about the choice and meaning of scenarios, and explored with the interviewees any comments they make concerning the authenticity or the reliability of their responses. This portion of the interview explored the transferability of interviewees’ stated choices during the hypothetical exercise to a real-world context.

3.3 Findings from the gaming-simulation task

Interviewees’ responses to the stimuli presented to them in each scenario confirmed findings in the literature that car club SDCSs and other methods of travel may be either substitutes or complements to each other. In several scenarios, for instance, interviewees chose to switch from cumbersome public transport journeys to use of a car club SDCS vehicle. But in other scenarios, where interviewees responded to losing access to a personal car, their coping strategies frequently made use of both the car club SDCS and public transport services.

Interviewees – whether currently SDCS subscribers or not – indicated that they would perform relatively little SDCS travel in nearly every scenario with which they were presented. The general perception of the car club SDCS was as a ‘gap-filler’ method of transport which interviewees could see themselves using, if at all, only occasionally. Non-car-owning interviewees indicated that they perceived a difficult trade-off between the comparatively-high fixed costs and “bother” of car ownership (residential parking, road tax, purchase cost, etc.) and the comparatively-high usage costs/bother of the car club SDCS (making an advance reservation, travelling to the parking bay,
inspecting the vehicle, paying an hourly rate, keeping to the duration of the reservation, etc.) The car owners in the sample tended to perceive this as a less-difficult trade-off, as the fixed costs of car ownership were perceived to be more than offset by the much higher usage costs of the car club SDCS for the relatively high level of car use which their activity patterns implied.

The points in time at which interviewees perceived themselves to be in a situation of choosing whether to subscribe to an SDCs varied quite strongly with their present level of automobile access. Car-less interviewees tended to view a car club SDCS as a new, low-commitment option with which they could experiment. For them, a car club SDCS was perceived to offer opportunities to access new activities or to access their existing activities more easily, for a moderate monetary cost. They frequently stated that they would use SDCS vehicles to access activities which they perceive to be facilitated by car access, such as bulk grocery shopping. Several had experimented with online grocery shopping, though members of one car-less household indicated that they had found this unsatisfactory and hence were keen to consider other alternatives.

Car-owning interviewees, however, tended to have activity-travel patterns tailored to personal car use. Switching from personal car ownership to holding an SDCS subscription was generally not something they would consider in the short term. One interviewee had voluntarily taken this step several months before the in-depth interview; she had sold her personal car (she had continuously owned cars for several decades) upon subscribing to a car club SDCS. She described her feeling in selling the car as one of ‘letting go of a lifeline’ akin to ‘stepping off the edge of a swimming pool and seeing if there’s water.’ Another interviewee, who had previously owned a car until several months before her interview, spoke of ‘putting off’ activities such as DIY home repairs and shopping at big-box stores as she did not at the time own a car.

Most car-owning interviewees did, however, see a car club SDCS as an option to consider at a future point in time when their personal car ownership might come into question – such as their personal car requiring major repairs, its lease term ending, moving home or job, or life course events in their family. Amongst car-less interviewees who were not inclined to subscribe to an SDCS at present, many saw life course events or career changes as potentially causing them to reconsider.

Interviewees’ responses to scenarios frequently involved adjusting their patterns of activity participation, either in response to the opportunities afforded by the SDCS becoming available to them (‘if I could use a car club I’d visit ______ on Saturday afternoon in Oxford’), or other stimuli (e.g. deciding to work from home several days during the week in response to losing access to a personal car).
Non-subscribers expressed significant discomfort with one particular aspect of the car club service model, the accrual of hourly charges during the full duration of their reservation of a car club SDCS vehicle. In the car club service model, a vehicle must be returned to the same location from which it is taken, and hourly charges accrue during a reservation regardless of whether the vehicle is in transit or parked (or returned prior to the end of the reservation). Interviewees seemed to perceive that their freedom of activity scheduling might be curtailed, and also seemed averse to the feeling of being ‘on the taxi meter’ whilst, for instance, visiting friends (in the words of one respondent).²

There appears to be a tension between the value that people derive from car club usage (mobility), and the essential resource which a car club SDCS provides (car-time, whether mobile or stationary) under the car club service model.

Interviewees uniformly stated that they would make different use of private cars than SDCS vehicles. In some cases these differences were simply the number of journeys by car (for instance, interviewees did not see car club SDCS vehicles as desirable for commuting purposes due to the hourly charges), though more complex responses were also indicated. Several respondents stated that they would combine multiple car-dependent activities (e.g. a visit to a relative and bulk grocery shopping) into a single SDCS use episode, which they would not do if using a personal car.

In the post-GS-exercise segment of the interview, interviewees reported experimenting with different travel options for their day-to-day travel, such as road routes or public transport connections. They generally saw a car club SDCS subscription (provided it did not involve trading a personal car for an SDCS subscription) as an option akin to a monthly public transport pass in terms of the level of personal commitment involved – and having significantly lower personal commitment than that associated with a car or bicycle.

A theme that cut across interviewees was their perception of a dichotomy between ‘public transport people’ and ‘car people’, though they felt that people can and do move between the two states at different stages of their lives. This seemed particularly relevant with respect to one’s knowledge of transport networks and the a priori commitments of time, effort, and/or money associated with certain methods of travel. One participant felt she could never be a car driver as she considered herself too clumsy, while another car-owner remarked that in his view bus riders just ‘jumped on the bus’ as they had appropriate knowledge to navigate the bus network, whilst this was a barrier to him. Another recalled that when family members visit from outside London, they end up taking frequent taxi rides as the out-of-town visitors are not comfortable with public transport. Some participants could see pressures for them to move between these two archetypal states should
certain life course events take place in future (e.g. birth of first child, moving to a new home, career progression, etc.)

Interviewees indicated that their consideration of their travel requirements at such points in life was generally not one of seeking optimality. For instance, one respondent reported that, when moving home within Central London several times during his 20s, he did not concern himself with considering travel options in detail as he felt he could generally rely on good access to the Underground system within Central London.

3.4 Summary and conclusions: Challenges for specifying the quantitative research methods

Taking a broad look at the qualitative findings from the GS task, it was concluded that the results supported, if possible, designing the remainder of this research, the quantitative portions, to be sensitive to two particular aspects of people’s decision-making behaviour.

First, subscribing to a car club SDCS was viewed by interviewees as interacting with other elements in their activity/travel lifestyle. Rather than SDCS journeys simply substituting for journeys by private cars or non-car methods of travel, interviewees frequently opted to also restructure other elements in their activity-travel pattern in response to changes in the availability of a car club SDCS. Hence, it was concluded that the research should incorporate a multi-day activity/travel pattern as the basic unit of analysis, rather than individual journeys as in a traditional analysis of people’s choice of travel mode. In the analysis, people would thus be specified to select a portfolio of travel methods to use for a set of journeys, rather than a single travel method for a single journey.

Second, the trade-offs which interviewees perceived between the [strategic] fixed costs of making methods of travel available and the [tactical] usage costs of travelling led to the conclusion that the analysis should be specified in a way to replicate people potentially taking with the opportunity to consider such trade-offs.

Following on from the GS task, Chapter four presents the concepts which underpin the quantitative analysis reported later in this thesis. Chapters five and six then describe the design and preparation of the two complementary datasets generated for the empirical application of these concepts. Whilst addressing the challenges identified by the GS task in the context of the proposed SC survey design, it was necessary to ensure that the level of respondent burden would not be unreasonably large, and that respondents would perceive a high degree of plausibility in the choice situations. These issues are discussed in Chapter six.
Notes

1 This service model was the subject of a 2008 agreement between the Chicago-area transit operator and a car club. (Hilkevitch, 2008) The car2go system employs a microchip physically secured to members’ driving licences for managing vehicle access. (Diem, 2008)

2 From informal discussions with SDCS subscribers, it appears that some develop complex strategies of adding time ‘buffers’ to their reservations and/or making prospective reservations which they may later cancel. They generally voiced satisfaction with the outcomes of their management of these issues – in particular that their overall transport costs, even with these sorts of coping strategies, are still lower than those of car ownership.
Chapter 4: Analytical framework

This chapter outlines the framework developed for analysing the market for subscription drive-it-yourself car services (SDCSs). The concepts arise, as noted below, from a combination of:

- the findings of the gaming-simulation research task reported in Chapter three
- the findings of the literature review
- theories of choice analysis and spatio-temporal accessibility

Section 4.1 presents an overview of the concepts within the modelling system, Section 4.2 introduces the notation, and Section 4.3 presents the links between the hypothesised strategic and tactical choicemaking. The options considered for the structure of the system are presented in Section 4.4, and then compared in Section 4.5. The proposed model form is presented in Section 4.6, and Section 4.7 introduces the perceived activity set, a concept of personal accessibility which supports the proposed model structure. Generic estimation issues relating to the proposed model form are briefly discussed in Section 4.8, and Section 4.9 summarises the material in this chapter.

4.1 Overview

The classical unit of analysis for studying people’s use of transport modes (‘mode choice’ analysis) is the individual journey. As discussed in the latter portions of Chapter three, the nature of SDCSs does not appear to be amenable to employing the journey as the unit of analysis. Before discussing this in depth, we first discuss a technique which has been developed in recent years for accommodating the potential for linkages between mode choices for multiple journeys.

Models of trip chains, or ‘tours’, treat a sequence of journeys beginning from a person’s home, onwards to one or more out-of-home destinations, and then returning to their home, as linked. This class of models would ensure consistency in the set of modes which a person uses on a given tour; ensuring, for instance, that a person making a home-shop-home tour would be predicted to drive their car for either both or neither of the journeys in this tour.

The cost structure of SDCSs, and how would-be customers perceive it, implies that the trade-offs between fixed and variable costs is relevant to the choice of whether or not to use the service. This presents a challenge to classical mode choice analytical methods, where fixed costs are assumed to be sunk and only variable costs are considered to be relevant. Tour-level analysis is insufficient for accommodating linkages between journeys in this context. Tour-level analysis accounts for possible
dependencies between journeys within any given tour, but no more broadly. (We re-visit the issue of
tour-level analysis in Section 8.2.1.)

It is therefore hypothesised that people have in mind a view of the set of activities (and therefore,
implicitly, the set of journeys) which they wish to be able to access. Trade-offs between fixed costs
and variable costs are then hypothesised to be made in consideration of this set of activities, which
we term the \textit{perceived activity set}. This concept is discussed further in Section 4.7.

We introduce the term \textit{mobility resource} (functionally equivalent to the term \textit{mobility tool} employed
by Beige and Axhausen (2008) or Scott and Axhausen (2006)) to refer to products or services which
enable travel in some way. Examples include personal cars, petrol, driver’s licences, public transport
tickets (whether pay-per-journey or unlimited for a given time period), bicycles, walking shoes,
membership in a frequent-flyer programme, etc. In principle such resources are not restricted to
market products/services and would include knowledge such as that of a network’s
topology/operating characteristics. For instance, using public transport requires both some sort of a
ticket product \textit{and} knowledge of how the system works; driving one’s personal car requires owning
the personal car \textit{and} having a driver’s licence. In the research reported in this thesis, we focus on a
subset of market-traded \textit{mobility resources} which have moderate-to-high fixed costs, thus requiring
a level of commitment, and are durable [though not permanent] in that they can be used over a
large set of journeys:

- Personal cars
- Bicycles
- Public transport season tickets$^3$
- Subscriptions to SDCSs

Items such as shoes are not included in this listing, despite, in this example, their relationship with
walking, due to their near-full penetration across the population; instead it is implicitly assumed that
[able-bodied] people have access to low-cost resources of this sort. Pay-per-use resources such as
petrol and single-use public transport tickets are not included in this listing of resources as they are
available with little commitment at the point of travel, and hence are treated as having purely
variable cost structures. Non-market-traded products are likewise not explicitly analysed, though
this is a simplification in the interests of practicality, rather than a limitation in principle.

In this specification, mobility resources are not exclusive; a person may own none, one, or several.
Different possibilities to take account of this aspect of the context are discussed below. Whilst they
do not all incorporate the concepts of people selecting a ‘portfolio’ of mobility resources, we
introduce this in the notation which follows in the interest of completeness, and describe further, in Section 4.4.1, its redundancy with regard to the ‘MDCEV’ specification.

4.2 Notation

This section outlines the generic specification of the model system; the indices \( i(i = 1, 2, \ldots I), \)
\( r(r = 0, 1, \ldots R), \) \( d(d = 1, 2, \ldots D), \) \( m(m = 1, 2, \ldots M), \) \( j(j = 1, 2, \ldots J) \) represent people, resources, resource portfolios (combinations of resources), modes of travel and journeys, respectively.

\( \rho \) represents the full set of resources, whilst \( \rho_i \) represents the set of resources which are available to person \( i. \) Each person is defined [and, in a cross-sectional dataset, observed] to hold one portfolio of resources from the full set of portfolios \( \Delta; \) \) the set of available portfolios from which person \( i \) chooses is denoted \( \Delta_i. \) Each available portfolio is comprised of a set of resources \( \{r: r \in \rho_i, \rho_i \subseteq \rho\}; \) we note that set \( \Delta_i \) includes a null portfolio which contains only the 0\(^{th} \) element from the set of resources. In the most general form, where a person can choose anywhere from zero to all resources without restrictions, they face a fully-factorial choice set of \( 2^R \) separate and distinct portfolio options.

We denote \( A \) to represent a matrix of dimension \([D, M]\) which maps each portfolio \( d \) onto the set of modes of travel \( \mu_d \) which it enables; \( \mu_d \) is a subset of \( \mu, \) the set of all travel modes. Element \([d, m]\) is a one/zero indicator which takes the value one if the \( d^{th} \) portfolio enables the \( m^{th} \) mode of travel and zero otherwise. All elements in columns of \( A \) corresponding to modes of travel that do not require \textit{a priori} acquisition of resources take the value one. If the \( m^{th} \) mode of travel requires the \( r^{th} \) resource, then the element \( A[d, m] \) is equal to one if and only if \( r \in d. \)

Empirically, matrix \( A \) takes the structure (neglecting resources associated with SDCS subscriptions):
<table>
<thead>
<tr>
<th>Portfolio of resources</th>
<th>Drive Car</th>
<th>Ride Public Transport without paying per journey</th>
<th>Ride Bicycle</th>
<th>Walk</th>
<th>Take Taxicab</th>
<th>Public Transport (PAYG)</th>
<th>Ride as car passenger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Car</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Own Public transport Season Ticket</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Own Bicycle</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Own Car + PT Season Ticket</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Own Car + Bicycle</td>
<td>1</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Own PT Season Ticket + Bicycle</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Own Car + PT Season Ticket + Bicycle</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Own none of these</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.1: Matrix (denoted as A) of the set of means of travel enabled by various resource portfolios

4.3 Linking the strategic and tactical choice dimensions

We term a person’s choice of portfolio as strategic in nature, and the choice of travel mode as tactical, where strategic choices are defined as those which condition a person’s options for making tactical choices.

As was discussed in Chapter two, traditional models of car ownership – a specific type of resource which can facilitate one’s travel – have either been insensitive to a person’s need/desire for travel, or have accommodated it in a simplistic manner.

It is hypothesised here that people choose to hold a travel resource portfolio on the basis of a compensatory weighing between:

1) the (dis)utility (expense, etc.) of acquiring and/or maintaining the resources contained within the portfolio, and

2) the (dis)utility of using the options [travel modes] which it enables in expected tactical choice situations (i.e. a set of travel needs)
This is denoted as:

\[ U_d^i = V_d^{i,\text{non-travel}} + V_d^{i,\text{travel}} + \epsilon_d^i \]  \hspace{1cm} (4.1)

The utility of portfolio \( d \) to person \( i \) is specified as a summation of a systematic component of utility associated with acquiring and/or maintaining portfolio \( d \) \( (V_d^{i,\text{non-travel}}) \), a second systematic component of utility which relates to the use of modes of travel enabled by portfolio \( d \) in some ‘expected tactical choice situations’ \( (V_d^{i,\text{travel}}) \), and an error term \( \epsilon_d^i \).

The particulars of these links between the strategic and tactical choice dimensions are dependent on broader model specification issues. Readers keenly interested in the chosen specification for these links may wish to read Section 4.7 before Sections 4.4 through 4.6.

4.4 Options for model structure

In the most generic form, each person is specified to make the two types of inter-related choices outlined. We observe the first, the portfolio of resources which a person holds, as a status relating to each person – the number of observations of this choice is equal to the number of people in the sample. We observe the second choice, the mode of travel which is used for each journey, once for each journey in the sample. As outlined above, the portfolio which a person holds is specified to set constraints on the set of modes available for them to use.

4.4.1 Multiply discrete-continuous extreme value model

The Multiply Discrete-Continuous Extreme Value [MDCEV] model, developed by Bhat and colleagues, was recently proposed to analyse situations in which a decision-making agent synthesises a portfolio comprising one or more [discrete] elements from a set of elemental alternative options, and simultaneously decides how much (along a scalar continuous dimension) of some ‘budget’ (time, or money, etc.) to allocate to each element in their chosen portfolio. (Bhat 2008) Concave marginal utility functions for each elemental alternative are parameterised.

The decision-making agent is modelled as seeking to maximise utility subject to their fixed budget. They are predicted to choose all alternatives which provide them a value of marginal utility in excess of some threshold value, in quantities which vary according to the marginal utility functions for each alternative. They are predicted to not choose all elemental alternatives for which the maximum marginal utility (which always occurs at zero, due to the concave functional forms) is less than their threshold value. The set of ‘portfolios’ of elements in this specification is, if unrestricted, known as a power set, equal to \( (2^n - 1) \), where \( n \) is the number of elemental alternatives (note there is no
provision for selection of a null portfolio of alternatives, which strictly speaking a power set would include). The ‘portfolio’ which a person chooses is, as it is comprised of all alternatives for which the prediction along the continuous dimension is larger than zero, accommodated implicitly in this specification. Whilst an application with multiple levels of analysis (analogous to the portfolio and mode levels in this context) has yet to be reported in the literature, the possibility exists in principle.

It was decided that the SDCS context was not directly amenable to the MDCEV form, however, as the MDCEV specification relies on a rigidly-fixed budget of some quantity, where the items within that budget are fungible. In addition to the obvious application to monetary budgets, it has been applied in the time-use context where each person has a fixed time budget of exactly 24 hours per day (where the degree of intertemporal fungibility is perhaps more of an issue than in the case of money.) It was not found to be possible to convincingly posit a similar type of budget that would be appropriate to SDCS participation; fixed time budgets, fixed monetary budgets, and fixed mileage budgets were considered but none appeared appropriate to treat as a rigidly-fixed quantity of fungible elements.

Specifying discrete journeys to be the unit in a person’s budget, in some form of a perceived activity set, though perhaps intuitively appealing, would imply a functional form different from the MDCEV; such a model form would be structured as discrete-discrete, akin to the specification described in the next section, rather than discrete-continuous.

**4.4.2. Distinct models of portfolio and mode choice dimensions**

In this specification, a person’s holding of a portfolio of resources would be treated as one observation, whilst each of their uses of a transport mode to perform a journey would be treated as an observation. There would thus be a total of \( f_i + 1 \) observations for each person \( i \); the number of alternatives in the choice set for their portfolio choice would be \( |\Delta_i| \), the corresponding choice set for each of their journeys would be \( \mu_{d_i} \) (representing the set of modes enabled by the portfolio which person \( i \) was observed to hold.)

If the utility functions for the two choice dimensions share no common parameters, and the residual error terms are uncorrelated, estimating them jointly will yield a parameter set which would be identical to that which would be obtained from two independent estimations. The ‘joint’ and ‘independent’ estimations would also yield asymptotically-consistent parameter sets if the ‘true’ parameters for all common attributes across the two choice dimensions (e.g. people’s tastes for travel time at the mode choice and portfolio choice levels) are precisely equally-sized. Comparing the best-fit parameter sets from independently estimating the two choice dimensions would provide
insight into whether people’s tastes when making ‘tactical’ choices are indistinguishable from their
tastes when making ‘strategic’ choices, or whether there are systematic differences.

4.4.3. Models of combinatorial multi-dimensional [portfolio and mode] choice

This specification would treat each observation associated with a given person, whether their
holding of a portfolio of resources or their use of a mode of transport, as a dimension of a single
multi-dimensional choice (of \( f_i \) + 1 dimensions). Thus the \( i \)th person’s observation is a
concatenation of the portfolio of resources they are observed to hold \( (d_i) \) and the \( |f_i| \) modes which
they are observed to use for each of their journeys: \( \{d_i, m_{1_i}, \ldots, m_{j_i}\} \).

The set of alternatives in this specification, which includes this observation plus all other available
options, varies across people with their person-specific \( f_i \) observation. There are \( D_i \ast \prod_{j=1}^{f_i} M_{d_i} \)
distinct alternatives facing person \( i \); the cardinality of this set of alternatives renders traditional
discrete-choice methods impractical as the logit [alternatively probit, etc.] probability calculation
would require [potentially millions of] millions of computations in both the parameter estimation
stage and sample enumeration in the simulation of people’s choices. As an order of magnitude, a
person facing a choice set of ten portfolios, and a choice set of ten means of travel for each of ten
journeys, would be specified to face a combined set of \( 10 \ast 10^{10} = 10^{11} \) distinct options.

These numerical challenges are potentially surmountable; procedures for reducing the number of
computations in the parameter estimation stage by sampling a subset of alternatives from a larger
full set were initially developed in the 1970s and have since progressed. (Ben-Akiva and Lerman
1985) Situations where sample enumeration becomes computationally-burdensome are rarer, as
sample enumeration involves only a small fraction of the number of computations involved in
parameter estimation. In recent years, however, researchers working with what the literature terms
as ‘colossal’ choice sets have experimented with Monte Carlo [stochastic] techniques for performing
sample enumeration which preserve the consistency of simulated choices. (Kikuchi et al. 2003;
Yamamoto et al. 2001)

4.5 Comparison of the ‘distinct’ and ‘combinatorial’ model forms

Appendix A contains a mathematical derivation showing that, whilst the ‘distinct’ and
‘combinatorial’ model forms arise from different concepts, they are mathematically equivalent
under the assumptions that:

A) the utility functions are affine (linear-in-the-parameters) in functional form, and

B) the residual error terms have the iid [independently and identically distributed] property
required for the multinomial logit functional form.
With regard to the latter assumption, it is noted that the systematic utility \( (V_x) \) terms in the utility functions below can in principle be specified to incorporate correlations amongst unobservables, leaving residual error terms having the iid property.

For ease of reference, Figure 4.1 summarises the two model forms.

**Figure 4.1: Summary of 'Distinct' and 'Combinatorial' structures**

### 4.6 Proposed specification

Section 4.4 presented three candidates which were considered for the general form of the models of the SDCS market. The MDCEV structure rejected as inappropriate for application in this substantive context. The two remaining – which we term the ‘distinct’ and ‘combinatorial’ forms – are conceptually different but, under assumptions described in Section 4.5, functionally identical.

The ‘distinct’ model form is empirically more efficient to work with (due to the manageable cardinality of the choice sets), and would not require stochastic sampling of alternatives to generate forecasts; it was therefore decided to proceed with this general specification. It is recognised that the ‘combinatorial’ model form is more general and may offer the possibility of uncovering patterns which are not identifiable with the ‘distinct’ form, though practical considerations\(^2\) led to the decision to employ the ‘distinct’ form in the remainder of this research.

From this point forward, term Strategic Portfolio [StrAP] model is used to describe this specification.

### 4.7 Perceived activity set

Section 4.3 presented the strategic/tactical concepts underpinning this model system, in which strategic choices (acquiring a portfolio of travel resources) are made in part on the basis of how well
each of the strategic options would serve a set of ‘expected tactical situations’ (a set of travel needs). The term **perceived activity set** (PAS) is proposed to represent the ‘expected tactical choice situations.’ It is formally defined as:

\[
\text{the array of activities which a person views, at a particular point in their life, as encompassing their travel needs.}
\]

In principle each activity in a person’s PAS may be defined by any of a number of attributes – timing, location, frequency, activity type, required co-participants, spatio-temporal rigidity, etc. Any practical application, as indeed is the case with the research reported in this thesis, will require judgments on behalf of the researcher as to which attributes are critical.

It is hypothesised that, at times of [potential] flux in the composition of travel resources in a person’s ‘portfolio’, they assess, through some mental process, the match between [the modes of transport enabled by] various portfolios and their travel needs in their PAS. Were this person an **homo economicus**, and their forthcoming pattern of activity participation deterministically known to them without error, their PAS would be an unbiased, error-free set of activities which they are likely to do within some timescale, where the timescale might be related to the functional lifetime of the durable travel resources which they are considering acquiring. **Homo economicus** is one model of human behaviour, and it is entirely plausible, in principle, that that there are systematic distortions from pure rationality in the hypothesised mental processes. To take one example, the term **perceived** activity set is proposed, rather than the more concise **activity set**, to account for the fact that a person’s view of their travel needs is a matter of their perceptions, which may not align fully with an outsider’s view from an ‘objective’ perspective. Likewise it is plausible that a person’s PAS may in some way relate to their past experiences rather than exist as a set of only prospective forthcoming activities.

For use in a compensatory utility-based analytical framework, the performance of the transport modes enabled by a resource portfolio in accessing each activity in a person’s PAS must be aggregated into a single scalar utility value. This implies two dimensions of aggregation: the first across modes of travel to access each activity, and the second across activities.

It is hypothesised that the mental process associated with the first dimension of aggregation is a form of maximisation. In other words, it is specified that a person, in considering a particular portfolio, considers how well (or poorly) the portfolio would perform in accessing a particular activity to be how well the ‘optimal’ mode enabled by this portfolio would perform. This is written formally as:
\[ U_{dji}^{\text{travel}} = \max_{m \in \mu_d} \left( U_{mji}^{\text{travel}} \right) \]  

\[ U_{dji}^{\text{travel}} = \max_{m \in \mu_d} \left( V_{mji}^{\text{travel}} + e_{mji}^{\text{travel}} \right) \]  

\[ V_{dji}^{\text{travel}} = \frac{1}{\lambda_{\text{travel}}} \ln \sum_{m \in \mu_d} e^{V_{mji}^{\text{travel}} + \lambda_{\text{travel}}} \]  

\[ U_{dji}^{\text{travel}} = \left( \frac{1}{\lambda_{\text{travel}}} \ln \sum_{m \in \mu_d} e^{V_{mji}^{\text{travel}} + \lambda_{\text{travel}}} \right) + e_{dji}^{\text{travel}} \]  

Equation 4.2 shows that, in this specification, the utility of portfolio \( d \) to person \( i \) for performing journey \( j_i \) (alternatively: accessing the activity/activities associated with \( j_i \)) is equal to the utility of the ‘optimal’ mode \( m \) enabled by portfolio \( d \) for performing journey \( j_i \).

If person \( i \)'s judgment of how each travel mode \( m \) would perform for journey \( j_i \) were knowable to the outside analyst with certainty, the analyst could proceed by identifying the \( m \)th mode of travel within \( \mu_d \) which provides person \( i \) with the largest value of utility \( \max_{m \in \mu_d} U_{mji}^{\text{travel}} \). The analyst's knowledge is, however, incomplete due to the latency of a person’s PAS, meaning that a systematic component of utility \( [V_{mji}^{\text{travel}}] \) can be inferred by the analyst, but there remains an unobservable error component \( [e_{mji}^{\text{travel}}] \), as shown in Equation 4.3.

Equations 4.4 and 4.5 present the form of the \( V_{dji}^{\text{travel}} \) and \( U_{dji}^{\text{travel}} \) terms. These relationships arise, if the analyst is prepared to assume that the \( e_{mji}^{\text{travel}} \) terms are Gumbel-distributed with variance \( \lambda_{\text{travel}} \), from the properties of the that distribution. (Ben-Akiva and Lerman 1985)

Readers may recognise Equation 4.4 as the logsum form which arises frequently in the case of choice model forms that include ‘nests’ of alternatives; in this class of models the \( \lambda_{\text{travel}} \) term sets the scale of the distributions of the error terms between upper and lower levels. In a simple two-level nested-logit application the upper bound of all \( \lambda \) terms is 1.0, as larger values would imply [illogical] larger within-nest variance than between-nest variance. In this application, however, we have no \textit{a priori} belief about the scale of the variance and hence also have none about the \( \lambda_{\text{travel}} \) term (beyond that it must be positively-signed.)
The second dimension of aggregation – across journeys – is dissimilar in structure to the first. Whereas the aggregation of utility across modes is thought to involve a person ‘identifying’ the utility provided by the ‘optimal’ mode from amongst the available options enabled by a portfolio, the aggregation across the activities within a person’s PAS is based on the notion that, by definition, the person wishes to be able to access each activity within their PAS. Therefore, we specify that this aggregation is summative across all activities within their PAS:

\[
V_{d_{\theta_i}}^i = \sum_{j_i=1}^{J_i} (V_{d_{j_i}}^i + \varepsilon_{d_{j_i}}^i) \tag{4.6}
\]

\[
U_{d_{\theta_i}}^i = \left( \sum_{j_i=1}^{J_i} V_{d_{j_i}}^i \right) + \varepsilon_{d_{\theta_i}}^i \tag{4.7}
\]

Here we introduce the notation \(\theta_i\) to represent the set of activities forming person \(i\)’s PAS. Equation 4.6 leads to Equation 4.7 as the individual \(\varepsilon_{d_{j_i}}^i\) terms are unidentifiable and the \(\varepsilon_{d_{\theta_i}}^i\) represents their sum into a single error term.

The PAS concept incorporates the possibility that a person may place a higher priority on accessing certain activities within their PAS than on accessing others. For instance, it is plausible that a person may place a higher importance on being able to get to an ill relative to assist with their care than on being able to get to the pub to see their friends. The opposite is also plausible, as is a possibility that a person places equal priority on being able to access each of these two prospective activities. Such systematic variations in the ‘importance’ of accessing different sorts of activities can, expanding on Equation 4.7, be represented as:

\[
U_{d_{\theta_i}}^i = \left( \sum_{j_i=1}^{J_i} \gamma_{j_i} V_{d_{j_i}}^i \right) + \varepsilon_{d_{\theta_i}}^i \tag{4.8}
\]

where the \(\gamma_{j_i}\) terms may be specified in any of a number of ways. Specifying the \(\gamma_{j_i}\) ‘importance’ terms to be functions of activity/journey characteristics, where the parameters of these functions are freely-estimated, would in principle allow insights to be drawn regarding how various activity/journey characteristics correlate with people’s holdings of travel resource portfolios.

The mapping of a set of activities into a set of travel needs [prospective journeys] is a non-trivial process. One possibility would be to assume that a person’s view of accessibility to the out-of-home activities in their PAS is based on simple ‘home to out-of-home activity to home’ two-journey tours,
though this is but one possibility. Another possible specification would be to specify this mapping in keeping with real-world patterns of the nature of a person’s tours. The latter specification was chosen here in the interests of pragmatism, though it is recognised further research on this point may yield useful insights into the specifics of how people perceive their mobility needs.

As noted in Chapter two, the PAS differs from many of the earlier notions of personal accessibility in that, while spatiality is a central aspect, it does not map to a continuous two-dimensional spatial surface as with many accessibility metrics (see Section 2.6.) It is also temporally sensitive (to time-of-day and day-of-week/month/year/etc.), in recognition that accessibility may vary substantially with time. It is a personal concept, unlike public transport accessibility levels (PTALs, see TfL 2010), as it depends crucially on a person’s perceived need for travel, rather than a scalar [impersonal] measure of the accessibility from a particular point to others.

The activity repertoire concept is perhaps closest in nature to the PAS, with the critical distinction being that the activity repertoire encompasses all activities of which a person is aware, whilst the definition of the PAS includes relevance: a person must perceive an activity as relevant to them for it to be within their PAS.

Chapters five through eight discuss the empirical application of the PAS theory.

4.8 Normalisation issues of the StraP specification

This section outlines a set of normalisation conditions for the portfolio-level utility functions which are required for successful identification of optimal parameter sets. Section 7.3 presents an analysis of simulated datasets to verify these conditions analytically.

4.8.1. General normalisation conditions

The general form of utility function at the portfolio level is:

\[ U_d^j = V_{d,non-travel}^j + V_{d,travel}^j + \varepsilon_d^j \]  \hspace{1cm} (4.9)

As described in Section 4.7, the \( V_{d,travel}^j \) term, which describes how well the set of transport modes enabled by a particular portfolio would perform over a person’s PAS, can be expanded as:

\[ U_d^j = V_{d,non-travel}^j + \left( \sum_{j_i} \gamma_{ji} \frac{1}{N_{travel}} \ln \sum_{m \in M_d} e^{(V_{mij}^{travel})} \right) + \varepsilon_d^j \]  \hspace{1cm} (4.10)

This general form requires normalisation in three ways.
First, the entire utility function must be normalised to set the scale of the $c_d$ term; this is typically done by fixing an arbitrarily-selected alternative-specific constant [ASC] at a value of zero. In this model specification, this implies fixing an ASC outside the logsum term (i.e. within the $V_{d}^{\text{non-travel}}$ term) to be zero. As will be discussed in Chapter seven, the $V_{d}^{\text{non-travel}}$ term is empirically specified here to have the form:

$$V_{d}^{\text{non-travel}} = \sum_{r=0, \forall r \in d}^{R} V_{r}^{\text{non-travel}}$$  \hspace{1cm} (4.11)$$

$$V_{d}^{\text{non-travel}} = \sum_{r=0, \forall r \in d}^{R} ASC_r + \beta^{\text{holding}} * x_r^{\text{holding}}$$  \hspace{1cm} (4.12)$$

Thus this first normalisation is accommodated by setting one of the $ASC_r$ terms to be zero.

After scaling the $V_{d}^{\text{non-travel}}$ term, a second normalisation is required to set the relative scale of the $V_{d}^{\text{travel}}$ term. This is done by fixing one ASC within the $V_{m_j}^{\text{travel}}$ term to be zero, in an analogous manner to that outlined above for the $V_{d}^{\text{non-travel}}$ term.

The third normalisation arises from $\lambda^{\text{travel}}$ appearing both within the logsum (where it is multiplied with the $V_{m_j}^{\text{travel}}$ term) and outside the logsum (where it is multiplied with the $\gamma_j$ term.) Thus the following equality holds for any constant $k$:

$$\gamma_j * \frac{1}{\lambda^{\text{travel}}} \ln \sum_{m \in \mu_d}^{M} e^{(V_{m_j}^{\text{travel}} * \lambda^{\text{travel}})} = k \gamma_j * \frac{1}{k \lambda^{\text{travel}}} \ln \sum_{m \in \mu_d}^{M} e^{(\frac{V_{m_j}^{\text{travel}}}{k})} = k \gamma_j * \frac{1}{k \lambda^{\text{travel}}} \ln \sum_{m \in \mu_d}^{M} e^{(\frac{V_{m_j}^{\text{travel}}}{k})}$$  \hspace{1cm} (4.13)$$

The $k$ terms outside the logsum cancel, as do those within the logsum. A third form of normalisation is thus required, which can equivalently take any one of the following forms:

1) Fixing at some arbitrary non-zero value any of the $\beta$’s which comprise the $V_{m_j}^{\text{travel}}$ term.
2) Fixing at some arbitrary non-zero value the $\lambda^{\text{travel}}$ term.
3) Fixing at some arbitrary non-zero value any one of the $\gamma_j$ terms.

4.8.2. Treatment of modes of travel common to all portfolios

Section 4.8.1 outlined the generic normalisation conditions for utility functions of the ‘portfolio-level’ model form; this section discusses a further constraint which arises from the empirical structure of the relationship between modes of transport and mobility resources.
As each ‘portfolio’ is a collection of mobility resources, it follows that those means of travel defined to not require holding any such mobility resources will be ‘enabled’ by all portfolios. (This can be seen in Table 4.1; the ‘common’ modes of transport shared by all portfolios in the empirical system are walking, taxi/minicab, and car passenger travel.)

This presents methodological difficulties in that parameters for the ‘common’ means of travel may not be separately identifiable when only people’s ‘portfolio-level’ holdings are used as the observations. To take an example, a person who holds an empty portfolio (holds no mobility resources) may view either walking or taxi travel as an attractive option for performing a journey in their perceived activity set. Thus this person’s observed [null] holdings of mobility resources on its own would provide no information with which to distinguish between the person’s tastes related to walking and those related to taxi travel. Arbitrarily large negative [ALN] taste parameters associated with either of these two modes of travel would yield, provided that parameters associated with the other mode are correctly-identified, identical goodness-of-fit in predicting this person’s observed holdings in comparison to the case where all parameters are correctly identified. This arises from the presence of the logsum term (which is a rough approximation of a maximisation operator), where ALN values are dominated by any other non-ALN values in the set being operated upon, in all cases where the ALN value is not the only element in the set.

If we have observations of two such people (continuing the example from the previous paragraph), however, it becomes possible in principle to identify separate parameters for these two modes of travel (walking and taxi travel in this example), by relying on any relative differences in the accessibility the two people may face. Consider the case where person #1 faces, for some reason, taxi options for their journeys which are very long in duration, whilst person #2 faces walking options for their travels which are very long in duration. An ALN parameter for taxi travel time would not materially affect goodness-of-prediction for person #1 (provided that the walking travel time parameter was not ALN) as the utility of walking would dominate the utility of taking a taxi in the logsum calculation. The opposite would hold for person #2 – but if both people’s records are in the same estimation dataset, both the walk and travel time parameters would in principle be identifiable as the likelihood function (which would be aggregated across both people’s data records) would penalise ALN walk or taxi travel time parameter values.

Data variation of this sort is a slender reed on which to lean to identify people’s [mode-specific] tastes for attributes of ‘common’ modes of travel, such as those described in the preceding example. In real-world situations [particularly intra-urban] itinerary times by various modes of travel are likely to all correlate closely with journey distance and hence with each other. In the case of taking a taxi
and car passenger travel, both of which we have defined to not require an *a priori* commitment (and hence to be common to all portfolios) for the purpose of this research, all journey itineraries for these two modes may be perfectly correlated depending on the method of synthesising them. This issue was explored using a simulated dataset with structure similar to the empirical datasets, as described in Section 7.3; it was found that the structure of the dataset did not permit identification of all taste parameters for ‘common’ modes of travel if they are specified to be alternative-specific.

The available data offer another possibility for distinguishing between people’s tastes for modes of transport which are common to all portfolios, as the travel modes used by people to perform particular journeys are also observed. This would require an assumption that people’s tastes for various attributes associated with travelling (e.g. journey duration, fare/petrol costs, etc.) when making ‘tactical’ mode choice decisions are structured identically as their tastes for the same attributes when making ‘strategic’ mode choice decisions, which is problematic for two reasons.

First, there is an interesting question of whether people’s tastes for attributes associated with travelling are systematically different at the mode choice level versus at the portfolio level of choice-making. There is evidence that people’s choices of a strategic nature may involve rather distinct considerations than those of a more tactical nature. (e.g. Raux 1997) Using both sources of data to identify the same parameters, though, would foreclose the possibility of investigating patterns of similarities and distinctive characteristics of these choice-making processes for any common parameters.

The other issue arises from the fact that the two choice dimensions do not have an obvious ‘currency’ to link them. In any joint estimation, observations from both choice dimensions would enter the likelihood function. If the two choice dimensions are heterogeneous, the researcher must consider whether and how to scale between the observations. The nature of the two choice dimensions (one observed ‘portfolio’ per person; one observed mode of transport per journey) is such that there is no uniquely-correct scaling factor with which to include both types of observations in the likelihood function. The heterogeneity arises from ‘mode choice’ observations being a type of flow, whilst the ‘portfolio choice’ observations are a type of stock. Thus the number of observations at the ‘mode choice’ level scales both with the period of observation and the number of people observed, whilst the number of observations at the ‘portfolio choice’ level scales only with the number of people observed.

One possibility considered for scaling between the two choice dimensions would have been to weight the data records in the likelihood function so that, regardless of the number of observations
in each of the choice dimensions, the aggregate weighting of the two choice dimensions in the likelihood function was equal. The choice of weighting the two dimensions to be equal in the aggregate (i.e. 50%/50%) would have been arbitrary, though, as would any other choice for this quantitative relationship (60%/40%, 10%/90%, etc.). This issue of selecting an arbitrary scale between the aggregate weighting would be immaterial if similar parameter values were identified for a wide range of possible values for this relationship, though this would be a special case which would not hold in general.

Taking these considerations into account, and the results of the analysis of simulated data presented in Chapter seven, it was decided to accommodate the presence of common travel modes in this empirical application by specifying all alternative-specific parameters to be common for modes of travel appearing in every ‘portfolio’.

4.9 Summary

This chapter outlined the notation and proposed StraP model form to be used in the analysis of subscription drive-it-yourself car services. Alternative specifications were presented, along with the rationale for the selection of the StraP model form.

Two principal methodological innovations are proposed. The first is the integration of two levels of choice analysis which are classically analysed separately: strategic choices of the holdings of travel resources (of which car ownership is the most-frequently-studied) and tactical choices associated with the use of a means of travel for a particular journey. The StraP model specified a person’s strategic choice in this regard to be made on the basis of a weighing between the cost of making the strategic choice and the value to be drawn from tactical use to perform a set of travel needs, which is termed a perceived activity set. This concept, a novel view of personal accessibility, is compared with prior notions of personal accessibility.

The second innovation is the treatment of composite multi-element ‘portfolio’ options in choice analysis. The qualitative interviewing task described in Chapter Three found that people seem to perceive an SDCS to be a part of a broader ‘portfolio’ of travel options which, taken as a whole, represent new alternatives to either owning a car or living a car-free lifestyle.

Chapter five details the techniques used to generate the empirical data to support the proposed StraP model specification, and Chapter Six follows on with the results from the empirical application.
Notes

1 Note that public transport is treated throughout this analysis as a single unitary means of travel, disregarding the diversity amongst the various means of travel which fall under the term such as overground rail, bus, and Tube. This simplification is to maintain the tractability of the analyses, some of which increase in computational demands geometrically with the number of portfolio elements, without sacrificing the core functionality of forecasting consumer demand for SDCs. Were the distinctions within the broad category of ‘public transport’ central to the requirements of this analysis, different approaches would have been considered.

2 One such practical consideration was that the computational resources required for the ‘distinct’ model (the less-computationally-intense of the two specifications) were on the order of several weeks per parameter estimation using a desktop PC, and on the order of several days using Imperial’s internal ‘cloud computing’ system (the High Performance Computing Service.)

3 Implicit in this specification is an assumption that a person is only interested in the performance of the optimal mode of transport enabled by a given portfolio to access a given activity, and that the performance of all other [sub-optimal] modes which it enables are not utility-relevant. Relaxing this specification to account for possible ‘insurance effects’ (i.e. utility-relevance of 2nd/3rd/nth-best modes) is possible in principle, though in the interest of pragmatism we leave this as an avenue for further refinement of this specification in future.

4 In this empirical application all car-driving, car-passenger, and taxi-minicab itinerary durations are perfectly correlated, as described in Chapter five.

5 If, for instance, a model system had two types of observations, one related to days and the other to weeks, they could be combined into a single likelihood function by [equivalently] scaling the weight of either ‘day’ observations by seven or the ‘week’ observations by one-seventh. No such intuitive scalar quantity exists to link the mode-choice and portfolio-choice observations.
Chapter 5: Design of E-NTS dataset

This chapter presents the design and descriptive results from the Enriched National Travel Survey [E-NTS] dataset.

This dataset was developed to provide a portion of the data to support empirical application of the analytical system outlined in Chapter four. The E-NTS dataset is designed to work in tandem with the dataset from the Advanced Vehicular/Activity/Travel And Resource [AVATAR] stated-choice survey, which is described in Chapter six.

Section 5.1 presents an overview of the principles which guided the data generation. Section 5.2 describes the motivation for the selected data generation technique. Section 5.3 presents background information regarding the NTS datasets. Section 5.4 introduces the online travel planning services which were used on the web scraping task, which is itself presented in Section5.5. Section 5.6 discusses the preparation of data for the web scraping task; the scheduling is presented in Section 5.7 and Section 5.8 discusses particulars of the public transport itineraries reported by the web scraping task. Implementation is described in Section 5.9, and the fidelity of the output data is discussed in Section 5.10. Section 5.11 discusses the empirical treatment of fixed ownership costs for the mobility resources, and Section 5.12 concludes this chapter.

5.1 Overview

Before discussing the E-NTS dataset in the remainder of this chapter, we first discuss in this section the circumstances leading to the decision to design the two datasets. At appropriate points throughout this and the next chapter, the functional links designed to facilitate joint analysis using the two sources of data are discussed.

In seeking datasets for use in this research, the following attributes were sought:

- Standard attributes for discrete-choice analysis: Information about the travel options which people used, along with similar information regarding the other options which were available to them but which they did not use. At least some observations should be in situations where subscription drive-it-yourself car services [SDCSs] were available for use.

- A suitably-long period of observation, to take into account the relatively infrequent use of SDCSs observed in the literature (see Chapter two.) Many activity-travel diaries observe people’s travel for a single day (though not all: c.f. Hanson and Hanson 1980, Pas 1988, Schonfelder and Axhausen 2002, Roorda et al. 2005). A person’s perceived activity set may
extend well beyond the few activities observed in a single 24-hour snapshot of their activity/travel; this is consistent with the empirical phenomenon of car buyers purchasing larger cars than necessary for their day-to-day needs in order to accommodate the possibility of less-frequent but important wants and needs (e.g. holiday travel). (University of Pennsylvania 1968; Heffner et al. 2005). In the context of SDCSs, where the viability of non-personal-car mobility portfolios depends strongly on the type and frequency of one’s ‘need’ for car usage, we sought a multi-day travel diary.

Other considerations applied, which we do not list here in full as they are more general: more-recent data is preferable to less-recent data, panel data is preferable to cross-sectional data, data collected under a rigorous sampling protocol is preferred, larger-sample data is preferable to smaller-sample data, etc.

Great Britain’s National Travel Survey [NTS] was identified as a promising dataset based on these criteria, in particular as all respondents report their travel for a seven-day period. (Anderson et al. 2008) The NTS has been carried out as a continuous survey since the 1980s, before which the NTS consisted of five periodic cross-sectional surveys at variable intervals, starting in 1965/1966. In recent years annual sample sizes have been in the range of 8,000 fully-participating households across Britain.

Datasets from the NTS were found to have two significant shortcomings for this application, however:

1) The NTS datasets that the UK’s Department for Transport [DfT] releases publicly contain very little spatial information about the origins and destinations of NTS respondents’ journeys. Such data is made available at the ‘Government Office Region’ geographic level, of which there are only nine in England (See Figure 5.1.) A journey beginning and ending within Greater London, for instance, would be coded at this spatial level, with no additional information as to where it began and ended. This is problematic as detailed spatial information is typically used with models of the accessibility provided by transport networks to impute the duration [and cost] of alternative itineraries for completing a journey.

2) No data is collected on SDCS subscription or use, as use of such services remains a very small proportion of travel at present and the NTS is undertaken to capture broad trends in Britons’ patterns of travel. Any travel by an NTS respondent in an SDCS vehicle would have been categorised as travel in a ‘non-household’ vehicle.
It was thought to be possible, in principle, to mitigate the first of these shortcomings. The NTS’ travel diary instrument asks respondents to provide fairly-detailed spatial information for each of their journeys, with the level of spatial detail changing after day six of the seven-day diary.

- **Days one through six:** Respondents are asked to write by hand the ‘village, town, or local area’ of each of their origins and destinations in their paper travel diary. This information is subsequently coded by NTS staff into approximately 4,800 ‘place names’ across Great Britain, of which 466 are within Greater London. Respondents also report ‘long-distance journeys’ of greater than fifty miles for three additional weeks prior to their diary week; the origin/destination locations of these journeys are also at the ‘place name’ level of geography.

- **Day seven:** Respondents are asked to write the full address, including postcode, of the origin and destination of each of their journeys.

It was determined, through discussions with DfT staff, that this ‘spatially-enhanced’ data is collated and maintained by the contractor (NatCen) which performs the NTS for the DfT, though is not released in the interest of maintaining the privacy of respondents. In response to the request to use such data for this research, DfT agreed to provide limited access to this ‘spatially-enhanced’ NTS data.²

This led to the design and preparation of the E-NTS dataset, as discussed in Section 5.5, which offered several advantages over traditional methods of imputing the attributes of ‘unchosen’ journey itineraries.

The second shortcoming – the absence of any information about the availability or use of SDCSs – led to the decision to undertake the AVATAR survey.
5.2 Motivation

Employing data from the labour-intensive National Travel Survey offers the potential for large efficiencies relative to an effort to collect such a dataset for this research. Nevertheless, the dataset required substantial effort to prepare it for use in the strategic portfolio [*Strap*] model system; this section describes such enhancements to the NTS datasets.

Any dataset used in a discrete choice analysis must include some level of knowledge of the characteristics of the options with which the decision-maker is presented. In other words, the researcher requires information about the options from which the person chose, in order to infer relationships in the patterns of those attributes and the choices that were observed. In the case where a researcher is analysing people’s choice of means of transport, travel diaries typically provide a range of information on many relevant attributes of the *observed* travel method that a person used for making a particular journey (journey duration, fare paid, number and nature of the legs of multi-leg journeys, etc.)

Obtaining such information for *unchosen* travel methods in such analyses represents a methodological challenge. For instance, for an observed car-driving journey a researcher may wish to know how competitive alternate travel methods (bus, Tube, walking, etc.) were. The typical source of such information is a computer model which contains an abstract representation of the land-use/transport network. From the given spatial and temporal information relating to a particular journey reported in a travel diary, the relevant attributes of unchosen travel methods may be synthesised.³

Generally-speaking, there is a trade-off between the coverage of such land-use/transport network models and the level of detail at which information is represented. Thus, urban or regional land-use/transport models may have relatively fine granularity in terms of network attributes, but limited information regarding the network beyond the urban region. Conversely, land-use/transport models with larger geographic coverage tend to have quite coarse granularity and much of the variability in input data is homogenised.

This trade-off is due to limited computing power of even the high-end computers on which land-use/transport models are traditionally maintained and operated, which is approached as data capacity and processing needs rise rapidly with coverage and granularity.

Both empirical evidence (discussed in Chapter two) and the results of the gaming-simulation task (discussed in Chapter three) indicate that an important use of shared-car services is to access ‘out-of-town’ activity opportunities and responsibilities. Hence, for the present application it is highly
desirable to work with a land-use/transport model which is both fine-grain in detail and supra-regional in coverage. To circumvent the traditional trade-off between these features, a methodology was employed in which the attributes of unchosen travel methods are synthesised via online travel planning services. Data mining techniques are employed to access, via the web, distributed computing resources with the capacity to store the datasets and process the algorithms of the advanced land-use/transport models.

5.3 The National Travel Survey datasets
The NTS\(^4\) generates a set of data files which are linked through unique ID numbering, several of which are:

- **Household/individual data**: Data gathered at an interview, which applies to all household members (e.g. household income, distance to nearest rail station, etc.) or to particular household members (e.g. holding of a driving licence, employment status, etc.) respectively.

- **Journey data**: Data from the seven-day travel diaries. This data is further broken down into ‘stage’ data where multi-modal journeys may be described as consisting of more than one stage.

- **Long-distance journey data**: Data from an interview in which NTS respondents report all journeys over fifty miles in length for a period outside of their diary week.

The period [outside the diary week] for which respondents report their long-distance journeys was shortened from three weeks to one week after the 2005 edition of the NTS. The most recent data at the time of this data collection effort were from 2006; due to this change in survey protocol data from 2006 were not taken forward.

Households requested to participate in the NTS are sampled from addresses in Britain using a stratified protocol. (NTS 2006 Technical Report) First, a set of postcode sectors\(^5\) are sampled, stratified by region, car ownership level, and population density, from across Britain. Sampled sectors are used for two years in a rolling procedure; half of the sectors are replaced each year.

Each year 22 households are then randomly sampled within each of these postcode sectors. As described in Section 5.7, it was decided that it would be practical for this research to design a sample of 300\(^6\) NTS-participating households living in Greater London. The only information regarding the location of the sampled postcode sectors prior to accessing the spatially-enhanced datasets was their Government Office Region (see Figure 5.1).\(^7\) It was thus decided to sample these 300 households in a stratified manner from the postcode sectors within London in order to ensure geographic diversity. The publicly-available NTS dataset does not indicate which postcode sectors are newly-sampled versus carried over from the previous year, thus the sampling on this research
did not take into consideration the quasi-panel nature of the NTS postcode sector sampling. The 300 households were randomly sampled from the 190 postcode sectors in years 2004 and 2005 with stratification to ensure that either one or two households were selected from each postcode sector.

This led to the following characteristics of the E-NTS sample:

- 300 households (average household size of 2.46)
- 738 people (561 adults age 16+) 8
- 10,428 diary journeys (8,658 on diary days one through six; 1,770 on diary day seven)
- 272 ‘long-distance journeys’ (reported to have taken place during the three weeks prior to each respondents’ diary week)

Figure 5.2 shows the distribution of the number of journeys reported by NTS respondents who fall within the E-NTS sample.

Figure 5.2: Cumulative distribution plot of the number of journeys performed by NTS respondents that are within the E-NTS sample
5.4 Online travel planning services

The online travel planning services which were employed on this research are the UK Department for Transport’s Transport Direct service, and Transport for London’s Journey Planner service, which we denote TD and JP, respectively. Travel itineraries are developed from:

- operators’ schedules for public transport itineraries;
- a combination of historic and real-time data sources for car travel itineraries; and
- fine-grained network geometry for walking and cycling itineraries.

Table 5.1 below compares the key service features of TD and JP. JP provides faster results – an important consideration given that a large number of web page loads would be required – and provides fare information for the major public transport systems in London. TD’s coverage extends across Britain, and provides car itineraries, though [at the time of data collection] no cycling itineraries. There are minor differences in the treatment of public transport modes in each of the two services. A methodology was thus developed using these two complementary services to compile information on alternative travel itineraries for journeys made by Londoners in the National Travel Survey.

<table>
<thead>
<tr>
<th></th>
<th>Transport Direct</th>
<th>Journey Planner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managing entity</td>
<td>UK Department for Transport</td>
<td>Transport for London</td>
</tr>
<tr>
<td>Geographic coverage</td>
<td>Great Britain</td>
<td>Greater London</td>
</tr>
<tr>
<td>Time per web page load</td>
<td>~30 seconds</td>
<td>~5 seconds</td>
</tr>
<tr>
<td>Travel modes for which itineraries are reported (unique modes in italics)</td>
<td>Car, Bus/Coach, Train, Underground/Metro, Ferry, Tram/light rail, Walk</td>
<td>Cycle, Rail, DLR, Tube, Tram, Bus, Coach, River, Walk</td>
</tr>
<tr>
<td>Means for reporting public transport fares</td>
<td>Available ticket options presented together with pay-as-you-go fare cost</td>
<td>London buses: Number of bus stages reported, from which fares can be imputed</td>
</tr>
<tr>
<td></td>
<td>London Tube fares not reported.</td>
<td>Tube: Fare zones covered by an itinerary are reported. Fares can be imputed from this information and time-of-day/day-of-week/traveller demographics</td>
</tr>
</tbody>
</table>

Table 5.1: Comparison of online travel planning services
It is of note that TD and JP do not always report identical itineraries when provided the same input information. For instance, the sample journey itinerary shown in Figure 5.4 is reported by JP to start at 13:03 and terminate at 13:29. Given the same inputs, TD reports a similar walk-Tube-walk itinerary that starts at 13:19 and ends at 13:46. The reported itinerary durations are not always identical either; in this example they happen to vary by one minute. It was decided, in the interests of practicality, not to perform a rigorous comparison of such differences between the two travel planning services, but rather to use the itineraries reported by JP where available, and TD only when needed. It is recognised that it may have been desirable, had more time and resources been available, and had it been deemed central to the analysis, to perform an in-depth comparison of the two services in their particulars.

The most-detailed spatial data for journey origins and destinations collected by the NTS travel diary instrument is at the postcode-level, a much finer scale than even the most detailed spatial zoning system in traditional land-use/transport models (there are an average of roughly 15 postal addresses per postcode in the UK.) Thus postcodes were used as the input data for the online travel planning services, which were accessed via a custom web scripting method.

### 5.5 Web scraping

The process of automating a repetitive task of page requests to a web site is known as *web scraping*. Applications are widespread and generally involve gathering a large amount of information from websites that can be done manually but is labour-intensive. Repeated web scraping of car dealers’ websites, for instance, could be used to track price movements of the used car market. A computer script, typically written in a standard programming language, directs the process. Figure 5.3 shows the workflow for this web scraping task.

This research employed the following components:

1) Publicly-accessible websites which link to databases (the TD and JP websites)
2) Web scraping software (Screen Scraper)
3) Purpose-written script files (Java language)
4) Text-format files of input data (origin location, destination location, time-of-day, day-of-week) for the web scraping
5) A plan for scheduling the web scraping: each request takes a number of seconds, hence it was necessary to develop a plan for completing the c.50,000 requests within the allotted time window (See Section 5.7)
6) Database software for processing the raw output data (Microsoft Excel and SPSS)
Figure 5.4: Screen capture of formatted HTML output from the Journey Planner travel planning service

Figure 5.4 shows a sample journey itinerary recommended by the JP service, with a Tube journey leg (in blue) bracketed by walking journey legs (in green) before and after. The script searches for user-defined patterns in the HTML and saves the desired information.
Web scraping offered several advantages over traditional methods of synthesising journey itineraries using PC-based models of land-use/transport networks:

1) Finer spatial resolution: Public transport journeys, for instance, are routed through the precise stops/stations which TD and JP recommend, including on-street bus stops. Journey origins and destinations are represented at a relatively-fine spatial scale, the postcode level. London’s state-of-the-practice PC-based land-use/transport model has c.1,000 geographic zones covering Greater London; the same area is covered by roughly 300,000 postcodes.

2) Finer temporal resolution: public transport itineraries, for instance, report journey start and end times based on operators’ schedules, and are therefore sensitive to the variations in travel conditions over the course of a day and between weekdays and weekends. The web scraping would thus indicate, at the time-scale of minutes, how close an itinerary would match the observed start and end time-of-day of a journey in the NTS dataset.

3) The online travel planning services are available for public use, requiring only an internet connection. PC-based land-use/transport models generally require proprietary software.

Web scraping, as proposed in this research, also had several shortcomings for the generation of alternate journey itineraries:

1) Reported travel itineraries are sensitive to idiosyncratic incidents occurring on the transport network on the day of the web scraping. This was mitigated by requesting that TD and JP report itineraries for several days ahead rather than the current day.

2) Up-to-date transport network conditions, such as short- and long-term engineering works, are represented at a fine spatial (frequently junction-specific or public transport route-specific) and temporal scale. Whilst this is a desirable feature in analyses of real-time or very recent travel behaviour, this is a drawback when using data several years old, as in this instance.

3) Inability to perform system-wide simulations of network link traffic levels and attendant variations in congestion levels: Traditional techniques for generating journey itineraries include ‘volume-delay functions’ which enable travel conditions to be endogenous to the analytical system through an iterative feedback process, with full system-wide simulation on a timescale of minutes or hours. This was not feasible in this application as it was only possible to access the spatially-enhanced NTS data once in this task. Predicted travel conditions under recurring congestion levels were reported by the web services, but each
journey was treated as a single network loading; there was no mechanism for aggregate loading as in system-wide simulation of traffic levels and congestion.

4) It proved not to be possible to gather information for multi-leg car/public transport itineraries. The nature of the access to the spatially-enhanced NTS datasets was such that it was not feasible to perform spatial analysis with a geographic information system to identify possible multi-leg itineraries incorporating both car and public transport journey legs. This drawback is therefore related to the circumstances associated with limited access to the spatially-enhanced NTS datasets, rather than web scraping per se.

5) Limited information regarding travel conditions on public holidays is provided. London’s public transport services do not operate on Christmas Day, for instance. TD reports journey itineraries for only the current subsequent two months, whilst JP does not report journey itineraries for the holiday period more than several weeks in advance. The selected methodology for this research, therefore, was sensitive to day-of-week variations in travel conditions, but not public holidays.

To summarise, the principal advantages of the web scraping relative to traditional techniques were found to relate to the level of detail at which spatio-temporal information is accommodated, and the practicality of performing this analysis with no-cost software. The main disadvantages were the inability to simulate changes in travel network conditions (e.g. in a case such as the uptake of SDCSs leading to traffic congestion substantially worsening or easing), and the possibility of discrepancies in travel network operating characteristics (e.g. the possibility of substantive changes in traffic congestion levels, public transport service patterns, etc. between the collection of the NTS datasets in 2004-05 and the web scraping in 2009.) On the basis of a weighing of these considerations, it was decided to proceed with the web scraping methodology for this research. The next step was to prepare input data for ease of interface with the TD and JP services.

5.6 Preparation of input data for web scraping

Whilst the online travel planning services accept spatial data in the form of place names or postcodes, it was determined that the use of postcodes was more reliable in generating journey itineraries. Providing that the postcode is formatted correctly, both of the services will accept postcodes as journey origin/destination locations without asking the user for clarification. When using place names (e.g. Piccadilly), the services frequently ask the user to clarify by selecting from a list of possible locations (“Did you mean...?”), complicating the task of performing automated requests.
Place names were converted into postcodes using a hierarchical process based on a set of rules, with the lower-numbered rules in the listing below taking precedence over higher-numbered rules:

1) Each data record for journeys from the seventh day of an NTS respondent’s travel diary indicated both the place names and postcodes of the journey origin and destination. The respondent-reported\textsuperscript{14} postcodes were thus readily-available and were taken forward.

2) Postcodes of respondents’ residences were identified by any member of a household making a journey where they recorded the postcode of their home.

3) If a respondent made a journey where they recorded the postcode of their school/college or their “usual place of work”, this was then mapped onto all other journey origins/destinations at their school/college or usual workplace. There is an implicit assumption that each respondent has a single “usual place of work” and/or school/college.

4) If a respondent travelled with another household member on a given journey, their journeys were determined to have the same origin and destination location. The determination of “travelling together” was made on the basis of two respondents in the same household making journeys with: identical day of the week, journey start time, duration, distance, travel mode, and travel purpose. Complementary travel modes (e.g. car driver and car passenger) and travel purposes (e.g. ‘go to school’ and ‘escort to school’) were taken into account.

5) A lookup table was generated of all postcode(s) associated with each NTS activity purpose at each place name, based on respondents’ seventh-day diary data. If a respondent’s trip end had a combination of activity purpose and place name which could be found in this lookup table, the postcode of the trip end was drawn at random from those in this table.

6) Another lookup table was generated which related each place name in Greater London to the postcode of the nearest facility consistent with each activity purpose. This was done using Ordnance Survey’s Address Layer 2 dataset covering Greater London, which contains 3.5 million point locations and a description of the function of each. Each of the 456 functional categories (dwelling, restaurant, airport, etc.) was associated with activity purposes with which the function is intuitively consistent (e.g. points described in AL2 as “restaurant” were associated with the NTS’ activities of “work”, “eat/drink (alone or at work)”, “eat/drink (other occasions)”, and “escort other” (this includes escort to eat/drink). Each place name was associated with only the nearest location (as measured from the centroid of the place name) as the boundaries of the place names were not available, and
the place names do not in general coincide with standard units of political geography (ward, unitary authority, etc.). If a trip end was identified with a place name within Greater London, the postcode of its spatial location was imputed from this lookup table.

7) The above rule could only be applied for activity locations within Greater London (the extent of the AL2 dataset’s coverage). Thus a lookup table was generated of all postcode(s) associated with NTS activities at each place name outside London, irrespective of activity purpose, based on respondents’ seventh-day diary data. If the place name of a respondent’s trip end located outside London could be found in this lookup table, the postcode of the trip end was drawn at random from those in this table.

8) A third lookup table was developed associating each place name with the postcode nearest to its centroid. The spatial location of each postcode in England, Wales and Scotland was accessed from the Office of National Statistics’ National Statistics Postcode Directory 2008. If a journey’s origin and/or destination was located at a place name which could not be associated with a postcode by any of the above rules, it was allocated a postcode by this method.

The above rules to allocate all journeys two postcodes, one each at the origin and destination, were developed of necessity as a set of heuristics. Were the nature of access to the spatially-enhanced NTS datasets different, it may have been possible to use the seventh-day diary data, where both the postcode and place name of each journey origin and destination were observed, to estimate quantitative models in lieu of these heuristic rules.

For respondents’ long-distance journeys which took place outside of their diary week, respondents only report the date on which each journey took place. The timing of such journeys during the day is not reported, however time-of-day information is of relatively high importance in this context as the service offered by certain travel modes can vary quite markedly throughout the day. In order to impute this data point for each long-distance journey, the dataset was split into 529 smaller datasets, one for each combination of the 23 categories of the NTS’ activity purposes. (23 x 23 = 529)

For each combination of prior- and post-activity type, LDJs which took place during the respondent’s diary week (and therefore had observed time-of-day data) were used to generate a discrete distribution of representative journey starting times. Thus, for instance, one distribution of journey start times was generated for journeys beginning at a person’s home and en route to visit a friend’s home, and a separate, independent, distribution was generated for trips returning from a
friend’s home. Imputed starting times for long-distance journeys not from the diary week were then drawn from these distributions.

### 5.7 Scheduling of the web scraping

It was calculated that a time window of three weeks (15 business days) would be feasible for performing the web scraping task. The scheduling of the web scraping task was based on the following considerations:

- 10 [60] seconds was budgeted for each page request to JP [TD], as it was unclear whether the 5 [30] seconds/page-request in testing would be achievable as a sustained rate.
- Journeys where two NTS respondents travelled together would not require separate processing. Eliminating such ‘duplicate’ journeys from the sample of 10,700 journeys left:
  - 7,312 journeys which both began and ended within Greater London. These could be processed by JP alone, with the exception of a single pass through TD for each journey to generate a car travel itinerary.
  - 1,108 journeys which did not both begin and end within Greater London, thus requiring TD to generate non-car itineraries as well as a car travel itinerary.
- Each of the journeys both beginning and ending within Greater London would be processed as follows:
  - One pass through TD to generate a ‘car travel’ itinerary
  - One pass through JP to generate a ‘walk-only’ itinerary
  - One pass through JP to generate a ‘cycle-only’ itinerary
  - One pass through JP to generate a ‘public transport, no restrictions’ itinerary
  - One pass through JP to generate a ‘public transport, rail only’ itinerary
  - One pass through JP to generate a ‘public transport, bus only’ itinerary
  - One pass through JP to generate a ‘public transport, no National Rail’ itinerary
  - Thus each of the 7,312 journeys was budgeted 120 seconds of processing time, for a total of 0.88M seconds.
- Each of the journeys beginning and/or ending outside Greater London would be processed as follows:
  - One pass through TD to simultaneously generate a ‘car travel’ itinerary and a ‘public transport, no restrictions’ itinerary
  - One pass through TD to generate a ‘public transport, no rail’ itinerary
  - One pass through TD to generate a ‘public transport, no bus’ itinerary
  - Thus each of the 1,108 journeys was budgeted 180 seconds of processing time, for a total of 0.20M seconds.
- Three instances of the web scraping could be run concurrently.
• Eight hours per day would be available for web scraping, thus 3 instances over 15 business days would allow 1.1M seconds of ‘processing time’, sufficient to complete this processing.

5.8 Interpretation of public transport itineraries from the web scraping

TD and JP report their suggestion of an ‘optimal’ public transport itinerary\textsuperscript{14} for any combination of the public transport modes which the user chooses to enable. For instance, a multimodal Tube-bus journey itinerary might be output from JP if all travel methods are enabled, but this itinerary would not be output if the web service user indicates they do not wish to generate an itinerary that includes bus travel. ‘Optimal’ is determined by the algorithms underlying the online travel planning services, in which a journey itinerary is selected based on criteria such as travel time, number of transfers, etc.

The online travel planning services can be thought of as reporting the ‘choice’ of journey itinerary within a ‘nest’\textsuperscript{15} one level below the modal alternatives. The choice within this nest is amongst all competing journey itineraries using any combination of the enabled elemental travel methods. We do not observe each possible itinerary (for which the set is very large); rather JP/TD present only their recommendation for the ‘optimal’ itinerary. If we consider that the choice within this nest is deterministic, we need only observe the ‘optimal’ itinerary – the presence of second-best itineraries in the choice set of competing itineraries would not affect TD/JP’s observed ‘choice’ of recommended itinerary. In other words, we assume that the optimal journey itinerary for each modal alternative that is output from the online travel planning services is the only one a traveller considers when choosing amongst different modal alternatives – all latent second-best journey itineraries by a given means of travel are irrelevant. A similar argument extends to TD’JP’s recommended ‘optimal’ optimal itineraries of the other modes of travel (e.g. car, walking, cycling), where it is assumed that all second-best itineraries (i.e. network routings) are irrelevant.

The procedure outlined in section 5.7 involved gathering multiple public transport itineraries by toggling restrictions in the use of particular means of public transport. The NTS dataset indicates the specific public transport method(s) used to perform each journey, hence it was feasible to identify which of the recommended public transport itineraries matched the observed journey’s characteristics most closely, and treat the other itineraries as available but unchosen options.\textsuperscript{16}
5.9 Implementation of web scraping

The web scraping was undertaken during June and July 2009. The procedures outlined in earlier sections of this chapter were found to be feasible, and the available time window for accessing the spatially-enhanced NTS datasets proved to be adequate.

The descriptive statistics of the empirical allocation of postcodes to place names for journey origins and destinations are:

<table>
<thead>
<tr>
<th>Journey origin locations</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1,764</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1,764</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>1,342</td>
<td>2,480</td>
<td>501</td>
<td>90</td>
<td>5</td>
<td>4,418</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>2,423</td>
<td>996</td>
<td>150</td>
<td>31</td>
<td>2</td>
<td>3,602</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>518</td>
<td>134</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>674</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>90</td>
<td>29</td>
<td>0</td>
<td>26</td>
<td>1</td>
<td>146</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1,764</td>
<td>4,377</td>
<td>3,642</td>
<td>673</td>
<td>148</td>
<td>8</td>
<td>10,612</td>
</tr>
</tbody>
</table>

A. Reported by respondent. These journeys took place on day seven of an NTS respondent’s diary week
B. Home, work or school location. Imputed from an observed postcode of a journey origin or destination on day seven by the respondent (in the case of work or school location) or someone in the respondent’s household (in the case of location of residence)
C. Imputed on basis of place name and activity purpose (a random draw from the set of postcodes of observed NTS respondents’ activity locations of same purpose at same place name)
D. Imputed on basis of place name and activity purpose (deterministically matched to nearest postcode to centroid of place name having a facility location identified as a possible match for the reported activity purpose, with facility locations sourced from Ordnance Survey’s Address Layer 2 dataset
E. Imputed on basis of place name (a random draw from the set of postcodes of observed NTS respondents’ activity locations at same place name)
F. Imputed on basis of place name (deterministically matched to nearest postcode to centroid of place name)

Table 5.2: Spatial match quality for journey origins and destinations

The processing time per page request was found to be very close to that in the testing phase (five and 30 seconds, for each JP and TD page request, respectively).

The spatial information which could compromise the privacy of NTS respondents was removed from the data output from the web scraping task, and the remaining information (the attributes of the journey itineraries) was processed and linked with the publicly-available NTS dataset.

5.10 Data fidelity

The process of allocating postcodes for journey origins and destinations was successful for all but 88 of the 10,700 journeys in the original data sample; these journeys did not have usable ‘place name’ locations for either one or both trip ends. (Thus the total of 10,612 journeys reported in Table 5.2.)

14 journeys were excluded as they were performed via use of minor modes (domestic air travel and private hire bus) which could not be accommodated by TD/JP. Of the 10,598 journeys processed in the web scraping task, at least one itinerary was produced for 99.8% (10,575), with the services
being unable to recognise the origin and/or destination postcode of 23 journeys. Table 5.3 shows the proportion of journeys for which the web scraping task reported itineraries by the various modes of transport. The services reported car travel itineraries for 96.7% (10,224/10,575) of the journeys in the sample. A large proportion of journeys (43%) were reported to not have walking itineraries; this arises as the services do not recommend itineraries for walking longer than 80 minutes\textsuperscript{17}. Where missing, itinerary durations were synthesised for walking, cycling, and using a car. This was done by estimating relationships from cases where itineraries were reported for at least two of these modes, and applying these relationships to synthesise the missing data points. For instance, the duration of cycle itineraries were found to be proportional to car itineraries by an average factor of 1.3 (with $r^2$ of 0.90); this factor was applied to synthesise a cycle itinerary duration for those journeys for which a car itinerary was reported by the web services but a cycle itinerary was not.\textsuperscript{18}

<table>
<thead>
<tr>
<th>Observed mode</th>
<th>Proportion of journeys for which Itineraries not found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car (either driver or passenger, including taxi)</td>
<td>3.3%</td>
</tr>
<tr>
<td>Cycling</td>
<td>9.5%</td>
</tr>
<tr>
<td>Walking</td>
<td>43.0%</td>
</tr>
<tr>
<td>Public transport</td>
<td>4.0%</td>
</tr>
</tbody>
</table>

Table 5.3: Proportion of journeys for which no itineraries for various travel modes were reported by the web scraping task

A substantial proportion (17.2%; 501 of 2,911) of people’s public transport journeys were poorly matched by the web scraping task. For these journeys, none of the several itineraries reported by the web scraping task was judged to be a reasonably close match to that reported by the NTS respondent. In general, with the exception of 17 journeys (3.3% of the 501), public transport itineraries were reported for these journeys by the travel planning services, though they did not closely match the reported itinerary. These 501 journeys were, in comparison to the 83% of public transport journeys which were matched, distinctive in several aspects:

- More likely to be multi-stage journeys: (42.5% v. 21.7%)
- More likely to have been performed via National Rail: (61.5% v. 13.7%)
- Longer duration (68 v. 42 minutes)
- More likely to have been commuting trips (51.5% v. 29.3%)

It was recognised that these distinctive characteristics of the unmatched public transport journeys heighten the potential for biased results in analyses using the web scraping outputs. In the interests
of pragmatism, and based on the view that A) the details of complex public transport itineraries were not central to understanding the market for SDCSs, and B) these journeys represented less than five percent of the data sample, it was decided to treat the self-reported attribute values of these journeys as if they were reported from an itinerary in the web scraping task. Had more time and resources been available, it may have been possible to explore the particulars of these journeys in more detail.

Table 5.4 presents a summary of descriptive statistics from the web scraping task.

<table>
<thead>
<tr>
<th>Characteristics of itineraries output from web scraping</th>
<th>Car</th>
<th>Public transport</th>
<th>Cycle</th>
<th>Walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed method of travel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drive car</td>
<td>n = 3,575</td>
<td>3,419</td>
<td>3,575</td>
<td>3,575</td>
</tr>
<tr>
<td></td>
<td>Mean = 26 mins.</td>
<td>53</td>
<td>33</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>Median = 16 mins.</td>
<td>40</td>
<td>18</td>
<td>70</td>
</tr>
<tr>
<td>Car passenger</td>
<td>2,341</td>
<td>2,258</td>
<td>2,341</td>
<td>2,341</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>54</td>
<td>34</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>37</td>
<td>17</td>
<td>65</td>
</tr>
<tr>
<td>Public transport</td>
<td>2,911</td>
<td>2,911</td>
<td>2,911</td>
<td>2,911</td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>49</td>
<td>41</td>
<td>182</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>44</td>
<td>26</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>157</td>
<td>149</td>
<td>157</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>43</td>
<td>22</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>39</td>
<td>18</td>
<td>63</td>
</tr>
<tr>
<td>Cycle</td>
<td>1,442</td>
<td>1,283</td>
<td>1,442</td>
<td>1,442</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>27</td>
<td>8</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>25</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>Walk</td>
<td>149</td>
<td>149</td>
<td>149</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>44</td>
<td>28</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>43</td>
<td>20</td>
<td>76</td>
</tr>
</tbody>
</table>

Table 5.4: Summary of descriptive statistics for travel time of journey itineraries, by observed method of travel

Figure 5.5 and Figure 5.6 present a comparison of the journey speeds predicted by the web scraping task with the observations\textsuperscript{19} in the NTS dataset for car and public transport journeys, respectively (the two travel modes which might plausibly be expected to have the strongest time-of-day patterns in this regard). The blue (observed) and red (predicted) curves were fit using spline methods.

As can be seen from Figure 5.5, both the observed and predicted travel times are broadly in line with the knowledge base regarding average travel speeds in London\textsuperscript{20} – roughly-speaking within a range between 15 and 20 miles per hour, depending on time-of-day. (TfL 2009) The slowest average
speeds are both observed and predicted to occur in the morning peak period, also in keeping with wider knowledge of London road network conditions. There is a perceptible drop in observed car travel speeds in the late afternoon, which is less apparent in the plot of predicted car travel speeds, though in general it would appear that the time-of-day pattern in average car travel speeds predicted by the web scraping task is reasonable.

**Figure 5.5:** Plot of observed (left) and predicted (right) journey speed by time-of-day for car journeys

The time-of-day pattern in travel speeds of public transport journeys, shown in Figure 5.6, also reveals noteworthy patterns. Whilst car travel speeds tended to be lowest in the morning and afternoon peak periods, the converse was found to occur with public transport travel in the data sample. The precise reasons for this are not directly relevant to this analysis – perhaps this signature may be associated with long-distance [relatively high-speed] rail commuting patterns. The salient question addressed by Figure 5.6 is whether the predicted time-of-day pattern is broadly in line with the observed pattern, which appears so upon visual examination.

**Figure 5.6:** Plot of observed (left) and predicted (right) journey speed by time-of-day for public transport journeys
Figure 5.7 shows a plot of the residuals from the web scraping task, calculated by subtracting self-reported journey duration from predicted journey duration; Figure 5.8 shows the same data with separate plots for each means of transport. Vertical bands are readily apparent – these arise from NTS respondents, in keeping with typical results of self-reported travel diaries, tending to report their journey durations in 5-minute and [especially] 15-minute increments. (Wolf 2000) There are no horizontal bands, however, as the JP and TD services report journey durations rounded to the nearest minute.

Figure 5.7: Plot of residual values from the web scraping task, disaggregated by mode of transport. Data for car passenger journeys are suppressed in the interests of clarity.
Both NTS respondent-reported and the synthesised travel times are always positive. This results in a lower bound for each residual value equal to the self-reported journey duration.

Table 5.5 shows the average residual values, disaggregated by mode of travel and journey duration. Substantial heteroscedasticity is apparent; TD and JP tend to, by and large, report larger-than-observed journey durations for short journeys and smaller-than-observed durations for longer ones. This finding is consistent with the patterns which can be visually observed in Figure 5.7 and Figure 5.8. Though we know relatively little about the travelling experience for journeys reported in the NTS, we may speculate as to several possible causes for this observed pattern.

<table>
<thead>
<tr>
<th>Observed duration</th>
<th>Car</th>
<th>Public transport</th>
<th>Cycle</th>
<th>Walk</th>
<th>Taxi and Minicab</th>
</tr>
</thead>
<tbody>
<tr>
<td>One up to 15 minutes</td>
<td>3 (9)</td>
<td>18 (16)</td>
<td>2 (7)</td>
<td>4 (12)</td>
<td>7 (11)</td>
</tr>
<tr>
<td>16 up to 30 minutes</td>
<td>-1 (12)</td>
<td>12 (16)</td>
<td>-2 (16)</td>
<td>3 (15)</td>
<td>-5 (9)</td>
</tr>
<tr>
<td>31 up to 60 minutes</td>
<td>-6 (15)</td>
<td>9 (19)</td>
<td>2 (12)</td>
<td>-11 (16)</td>
<td>--</td>
</tr>
<tr>
<td>61+ minutes</td>
<td>-20 (33)</td>
<td>-11 (30)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Table 5.5: Average residual errors (predicted minus observed) disaggregated by mode of transport from the web scraping task (standard deviation values in parentheses). Values are in minutes; cells with n<30 are suppressed.
The positive residual values for short-duration journeys are unsurprising, given the lower-bounding on the residual values noted above. The under-estimation of journey durations for longer-duration journeys is less-intuitive. One hypothesis is that there is a sort of self-selection taking place due to the empirical nature of the NTS dataset. Journey durations predicted by the web scraping task are sensitive to recurring travel delays, though not idiosyncratic delays. Idiosyncratic delays will tend to result in longer reported journey durations in the NTS dataset, though. A similar argument exists with regard to unobserved heterogeneity in the pace at which people perform otherwise identical itineraries – for instance one person may view a particular ½-mile walk as a ten-minute journey whilst another may view it as a 15- or 20-minute walk, or such views may depend on real-time circumstances such as fatigue, sobriety, etc. The first of these hypotheses could be tested by investigating whether the pattern of residuals is the same for journey distances, though this data could not be gathered in the web scraping task. Though it was not performed on this research, it is noted that the second hypothesis could be tested by looking at correlations in the patterns of residuals for each person, with statistical adjustment to take into account patterns associated with geographic location.

A second possible self-selection effect may arise due to the disaggregate timing patterns of people’s public transport journeys, as public transport journeys uniquely involve waiting time prior to departing. We may plausibly expect that NTS respondents making public transport journeys tended to schedule the beginning of their journeys to minimise waiting time before boarding. In the NTS dataset, however, we do not observe the actual scheduling constraints governing respondents’ travel; rather we observe the journey’s attributes as the respondent reports them. Respondents in the data sample reported no waiting time at all for fully half of their public transport journeys, though this implausibly-large proportion is clearly an instrument artefact of the self-reported diaries. JP and TD may take account of the issue of waiting time by adding a small amount of ‘buffer time’ to recommended public transport itineraries, to allow users to arrive in good time to their public transport service. They advise users of service frequency, but, with the exception of unavoidable waiting times between two stages of multi-stage public transport journeys, TD and JP do not advise particular waiting times. It is therefore plausible that the travel planning services may add a small amount of ‘buffer time’ to help ensure that users experience public transport services with a low likelihood of disappointment at a longer-than-predicted journey duration. This hypothesis, whilst not test-able with the available data, appears consistent with the large positive residual for short-duration public transport journeys.
There is the possibility for a third self-selection effect, associated with the process of assigning postcodes to each NTS ‘place name’. If we take as an example two adjacent geographic areas representing two place names, each of which covers a circular area of radius \( r \), the distance of all journeys between them will be lower-bounded by zero and upper-bounded by \( 4r \) (twice the diameter of the circles). We may plausibly expect that short-distance journeys between the two place names will have a greater tendency to be walking journeys, whilst longer-distance journeys will have a lesser such tendency. If the process of allocating postcodes to journey origins and destinations prior to the web scraping task was spatially unbiased, we would be systematically introducing bias in the form of over-predicting the distance (and hence duration) of walking journeys and vice versa for non-walking journeys. Indeed the average residual for walking journeys under 15 minutes in duration is positive, though within the range of similar residuals for journeys by other modes. Whilst this is an intriguing hypothesis, it cannot be tested with the data at hand; no access was available to the geographic boundaries of the ‘place names’ in the spatially-enhanced NTS dataset, and it was not possible to perform detailed spatial analysis with the spatially-enhanced data. Had that been possible, the seventh-day diary data, within which journey origins and destinations are known at both the ‘place name’ and postcode level, could have been used to investigate the extent to which this phenomenon may have occurred.

<table>
<thead>
<tr>
<th>Journey origin locations</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1 (15)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>B</td>
<td>--</td>
<td>1 (15)</td>
<td>3 (19)</td>
<td>1 (18)</td>
<td>3 (15)</td>
<td>--</td>
</tr>
<tr>
<td>C</td>
<td>--</td>
<td>4 (19)</td>
<td>3 (20)</td>
<td>3 (19)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>D</td>
<td>--</td>
<td>0 (20)</td>
<td>0 (30)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>E</td>
<td>--</td>
<td>3 (14)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>F</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Table 5.6: Average residual errors (predicted minus observed) disaggregated by source of spatial match for journey from the web scraping task (standard deviation values in parentheses). Values are in minutes; cells with n<30 are suppressed.

Table 5.6 shows the pattern of the residual errors in predictions of journey duration by the quality of the spatial match, which decreases as one moves right and down in the matrix. The residual error in cell [AA] represents the error strictly due to the difference between observed journey durations and those predicted by the web services, as these journeys from the seventh day of NTS respondents’ diaries were subject to no spatial imputation. All other journeys were subject to one of the various methods of spatial imputation as described in Section 5.6.

The trend in average residuals in Table 5.6 is relatively weak; all are positive and on the order of several minutes regardless of the method of spatial imputation. This suggests that the spatial
imputation process did not introduce a substantial amount of bias into this task of synthesising journey durations.

The corresponding trend in the variation (standard deviation values) of the residuals is also relatively weak. Increasing variation as one moves down and right through the matrix would suggest that the ordering of the heuristic rules described in Section 5.6 is appropriate. The empirical data show that some variation is introduced by the methods of spatial imputation, though the ordering is not monotonic in a way which would unequivocally support the ordering of the heuristic rules which was selected ex ante. The main exception to monotonicity was that the variation in the column labelled ‘E’ in the above matrix was found to be smaller than the variation associated with the column labelled ‘D’, suggesting that, in retrospect, it may have been desirable to have switched the ordering assigned to these two rules.²²

Though it was not possible to undertake as part of this research, one can speculate on a hypothetical analysis where each journey in the dataset would be subjected to more than one method of spatial imputation, such that results can then be compared at a disaggregate level. As a final point, it is noteworthy that the magnitude of variation introduced by the spatial imputation methods²³ is generally small compared to the baseline variation introduced by the web scraping task.

As a further test of the structure of the E-NTS dataset, a simple choice model of travellers’ choice of means of travel was prepared. By employing an uncomplicated model specification with known statistical characteristics, this test investigates whether the E-NTS data allow the identification of parameters in line with a priori expectations.

Separate free parameters were estimated for travellers’ marginal (dis)taste for journey time, together with a generic parameter for journey cost and a full set of alternative-specific constants. The results are shown in Table 5.7. All parameters are strongly statistically significant (the smallest t-value was 3.85), and the marginal travel time parameters and the generic marginal journey cost parameter are negative, as expected. The marginal journey time parameters are all of the same order of magnitude, also in keeping with expectations. The implied values-of-time, which are calculated as the ratio of the parameters for marginal journey duration and cost, range from £8 to £13 per hour, which are quite plausible and within the range of values-of-time reported in the literature. (Mackie et al. 2003, Hensher 2001)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Significance level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant for driving</td>
<td></td>
<td>Normalised to zero</td>
</tr>
<tr>
<td>Constant for car passenger travel</td>
<td>0.528</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Constant for cycling</td>
<td>-2.91</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Constant for taxi/minicab travel</td>
<td>-2.51</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Constant for public transport</td>
<td>-0.480</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Constant for walking</td>
<td>1.260</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Parameter for journey cost (£)</td>
<td>-0.171</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Parameter for journey duration, car driving (minutes)</td>
<td>-0.0369</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Parameter for journey duration, car passenger travel (minutes)</td>
<td>-0.0355</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Parameter for journey duration, cycling (minutes)</td>
<td>-0.0360</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Parameter for journey duration, taxi/minicab travel (minutes)</td>
<td>-0.0253</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Parameter for journey duration, public transport (minutes)</td>
<td>-0.0223</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Parameter for journey duration, walking (minutes)</td>
<td>-0.0360</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Number of journeys in sample</td>
<td>10,575</td>
<td></td>
</tr>
<tr>
<td>Null log-likelihood</td>
<td>-30,290.6</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at estimated parameters</td>
<td>-22,646.3</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.7: Results of diagnostic mode choice model run using the E-NTS dataset

This section presented the findings of a number of diagnostic analyses to investigate the fidelity of the E-NTS data generated by the web scraping task. Whilst the results of the model run shown in Table 5.7, and the other data-quality analyses reported in this section, do not provide proof of suitability for proceeding beyond this diagnostic stage, in their combination they appear to support a tentative conclusion of this nature.

It is also noteworthy that best practice in discrete choice analysis is to specify attribute values such as journey duration to be those perceived beforehand by the choice-maker, rather than objective measurements of attribute values actually experienced as an outcome of the choice. (Ben-Akiva and Lerman 1985) Measurement and prediction of journey durations is an area of current research; it is recognised that people may perceive journey durations with greater nuance than the simple point estimates considered in this research. For instance, people may in some circumstances perceive there to be a distribution of possible journey durations by a given mode/route, and there is a growing body of evidence that people may make travel choices on the basis of criteria other than average expected journey duration. For the purpose of this present research the use of point estimates for expected journey durations appears appropriate; we note that one possible extension of this analysis would be to explore patterns of behaviour associated with uncertainty in the [experienced and/or expected] outcomes of travel.
5.11 Treatment of fixed ownership costs of ‘mobility resources’

Employing the E-NTS dataset in the ‘portfolio’-type analytical framework described in Chapter four requires decisions on the part of the researcher as to what to infer for the fixed holding costs of the mobility resources, in particular those of resources that a person does not own. If buying a car, for instance, depends on a comparison between the fixed costs of acquiring/maintaining one and the value which can be derived from driving it, these acquisition/holding costs must be represented in the specification in some way. There is no single correct way to do this; to continue this example the fixed holding costs of a car vary greatly with the type of car as there is much heterogeneity in offers on the car market.

Simpler representations of the holding costs have the advantages of less preparation effort, less-stringent data requirements and greater ease of estimability, whilst more-complex representations may provide better estimation results or further behavioural insights.

In this application it was decided to specify the fixed holding costs in a rather simplistic manner. This does not preclude subtler treatments being the focus of a future line of enquiry, though it was judged that the requisite level-of-effort for a more-refined specification was better directed towards other aspects of this research.

The assumed attributes of each of the mobility resources were:

- Car ownership was assumed to cost £4,000/year in fixed costs (RAC 2010,)
- Bicycle ownership, including accessories, was assumed to cost £150/year
- Public transport season tickets were assumed to cost £100/month for E-NTS respondents living in Inner London, and £150/month for those living in Outer London. (based on a simplification of TfL’s season ticket prices for various zone-to-zone combination.)

5.12 Summary and conclusions

This chapter describes the preparation of the E-NTS dataset. The particular requirements of the SDCS context led to a choice of a novel method of enriching travel survey data, through a technique known as web scraping. The process of planning the effort is described, followed by descriptive results from the data collection. The error structure of the E-NTS dataset is examined and discussed, and a simple diagnostic model run of travel mode choice is shown to produce results in line with expectations.

After undertaking the effort described in this chapter, the E-NTS dataset had the following structure:

- Each person’s observed ‘portfolio’ holdings of ‘travel resources’
• Each person’s observed use of travel modes to compete the journeys they undertook during a seven-day period
• Attributes of the ‘chosen’ and ‘unchosen’ travel modes for each of these journeys

The next chapter presents the design of the AVATAR survey, which was intended to complement the E-NTS dataset by mimicking its structure as closely as possible whilst gathering the missing information on the use of SDCSs.

Notes

1 NTS respondents also report walking journeys differently on days one through six than on day seven. On days one through six, respondents log only walking journeys of one mile or longer, whilst they report all walking journeys 50 yards or longer on day seven. Analyses in this research that are based on NTS journey-level data accommodate this by weighting ‘short walks’ by a factor of seven. Respondents are not asked to report walks of less than 50 yards on any day of their NTS diary.

2 NatCen was able to make available the full spatially-enhanced NTS datasets on their ‘data enclave’ – a PC in a secure physical and network environment, with scheduled and managed access. After processing this data as described in Section 5.6, access was granted to use a randomised dataset limited to spatio-temporal data points (journey origin and destination locations, time-of-day, and day-of-week) on an internet-enabled PC, where the key field linking the randomised data to the rest of the NTS data points did not leave the ‘data enclave’ PC. Following the completion of the ‘web scraping’ task described in Sections 5.5 through 5.9 of this chapter, the randomised data were re-linked with the spatially-enhanced NTS datasets on the ‘data enclave’ PC. All spatially-enhanced (not publicly-available) information was then removed, and a datafile with only the attributes of alternative journey itineraries and a key field linking to the publicly-available NTS dataset was allowed to be removed from the NatCen systems for subsequent use (subject to a suitable data security protocol.)

3 Researchers occasionally employ methods which synthesise journey itineraries by unchosen travel methods on the basis of observed journey(s) within a dataset that have similar spatio-temporal characteristics as the target journey, but were performed by a different method of travel. Such methods were considered for this application but were deemed infeasible as it was thought likely that suitable information could not be inferred from the dataset for a large proportion of journey itineraries, and it was not possible to determine this prior to accessing the spatially-enhanced NTS dataset.

4 This and subsequent references to the National Travel Survey datasets refer to the annual editions of the NTS following the last round of major changes to the survey protocol, which came into effect with the 2002 edition.

5 There are c.12,000 postcode sectors in Britain, each having an average of c.2,000 postal addresses. (ONS 2008)

6 Data from an additional 200 NTS-responding households were identified for possible inclusion in the sample if time on-site at the NatCen office permitted, which would have resulted in a sample size of 500 households. In the event it was found to not be practical to enlarge the sample by making use of this reserve data.

7 DfT was able to make available, after the data collection effort, the London Borough in which each of these 300 households resides, which was used in the analysis tasks.

8 34 [29] of these people [adults] were not observed to make any journeys during their diary weeks

9 TD is required for journeys which do not both begin and end within Greater London, and to generate car travel itineraries for all journeys.

10 Multi-leg journeys incorporating both car and public transport journey stages are empirically observed to be 1.1% of Londoners’ journeys in the 2004/05 NTS.

101
In cases where respondents did not report the full postcode of a journey origin or destination, NTS staff attempted to infer the postcode from any information the respondent did provide.

Empirically, 279 of the combinations of origin activity and destination activity were not observed in the dataset; there were 250 populated combinations.

This analysis employed data from all NTS-responding Londoners in 2004-05 to generate the empirical distributions.

TD/JP in fact report a small number (typically less than five) of recommended travel itineraries which are consistent with the user’s request (e.g. maximum walking times, limits on the number of transfers, etc.) The ‘optimal’ itinerary was selected as that which, given the departure time of the journey, brings the traveller to their destination the earliest, though it is recognised that the selection of this decision rule was somewhat arbitrary, as other possible decision rules also exist.

The term nest is used in its typical sense in discrete choice analysis; see (Train 2003).

The treatment selected to accommodate this data structure was to place each of these public transport options into a ‘nest’ as they were thought to be likely to share unobserved characteristics, with an associated logsum operator. This logsum term was fixed at 1.0 for all analyses reported here as it would add a dimension to an already-complex estimation process, and was deemed to not be critical to the analysis of subscription drive-it-yourself car services.

The researcher must exercise judgment regarding the nature of the choice set of options from which decision-making agents are specified to choose. An issue arises, for instance in the case of an unreasonably-long walking itinerary, whether to specify that such an option is ‘unavailable’ or rather to include it as an available, albeit onerous, option like any other. In the latter case, it would be specified that unreasonably-long walking itineraries are available for consideration by choicemakers, and are simply not chosen. It was decided to leave such options in the model system, rather than restrict their availability by some set of quasi-arbitrary rules, as an initial treatment. The possibility was left open to re-visit this decision at a later point if the empirical analysis appeared to warrant it.

It is recognised that it would have been possible in principle to incorporate a more subtle treatment of imputed cycling speed, by taking account of a person’s age and gender, for instance.

The ‘observed’ distance and duration data used to calculate the average journey speed for each journey is in fact based on self-reports by NTS respondents.

It should be noted that the data in Figure 5.5 and Figure 5.6 represent all journey in the data sample, rather than only journeys beginning and ending within Greater London. This was chosen as the principal logic for undertaking this analysis is to compare observed and predicted journey speeds, and the reference to the knowledge base regarding average car travel speeds in London is simply to place this analysis in a wider context.

In point of fact, there is undoubtedly some degree of error associated with the spatial data being self-reported by NTS respondents (rather than some notion of ‘ground truth’), the ‘cleaning’ of such data by the NTS staff processing the raw diary data, and the allocation of each journey origin and destination to the postcode in which it falls (given that each one of Royal Mail’s postcodes generally represents a spatial area rather than a discrete point in space.) These sources of possible error are not discussed further here, though it is noted that they exist for the entirety of the spatially-enhanced NTS dataset.

Rule ‘D’ (item #6 in the listing in Section 5.6) incorporates information on the spatial location of ‘activity-relevant’ facilities within each place name, whereas Rule ‘E’ (item #7 in that listing) relies on the location of other NTS respondents’ activity locations, regardless of activity purpose, to form the pool of possible postcodes for an activity location from which one was drawn at random. This implies that other people’s observed activity locations, of all sorts of activity purposes, were better predictors of the [unknown] location of an activity with a known purpose than the locations of facilities which were thought to be ‘relevant’ to the known activity purpose.

The reader can perform these calculations by subtracting the 15 minutes of variation in cell [AA] of Table 5.6 from each of the corresponding values in other cells of the matrix.
Chapter 6: Design of AVATAR survey

Chapter four presented the analytical framework for this research, arising from the findings of the gaming-simulation task, which for a number of reasons was found to require novel data to support the empirical study. The early portions of Chapter five describe the considerations which led to the decision to use Britain’s National Travel Survey as the starting point for the development of the empirical data base. This led first to the design and preparation of what is referred to here as the enhanced NTS [e-NTS] dataset, which is described in the latter portions of Chapter five.

This Chapter introduces the second component in the data collection plan, the Advanced Vehicular/Activity/Travel And Resource [AVATAR] stated-choice survey. It is designed to complement the E-NTS dataset by adding observations of how people might engage with subscription drive-it-yourself car services [SDCs], which the E-NTS dataset lacks.

Section 6.1 discusses the unique design criteria which were identified in the gaming-simulation task reported in Chapter three. Section 6.2 presents the proposed avatar device which was employed in the survey design. Section 6.3 describes the synthetic choice context, and Section 6.4 discusses the application of ‘efficient’ survey design principles in this application. Sections 6.5 presents the sampling plan and fieldwork methodology, and Section 6.6 describes revisions to the survey’s instrument package following the initial field testing. The characteristics of the data sample are contained in Section 6.7. Section 6.8 describes the weighting strategy for the resulting dataset, and Section 6.9 presents descriptive results from the survey. Section 6.10 concludes this Chapter.

The material in this Chapter draws heavily from the following articles:


6.1 Design criteria

As outlined in the final section of Chapter three, a principal outcome of the gaming-simulation task was the setting of the design criteria for the quantitative portions of this research, which it was decided from an early date would require a stated-choice [SC] survey of whether and how people might engage with SDCSs.

The two criteria specific to this application, which supplement the standard design requirements of stated-choice surveys, were:

1) Respondents should be provided the opportunity to indicate engagement with SDCSs as part of a package, or ‘portfolio,’ of linked choices related to their mobility

2) Respondents should be provided the opportunity to engage with SDCSs both strategically (i.e. in considering whether to subscribe) and tactically (in considering how/whether to use an SDCS)

6.1.1. Portfolio form

The first challenging aspect of SC design relates to structuring the choice situations such that for the purposes of quantitative analysis people may be considered to choose a portfolio containing zero, one or multiple elemental alternative options from an n-element choice set. (See Chapter four.)

As noted in Chapter two, Wiley and Timmermans (2009) report on a proposed methodology for SC survey design that accommodates portfolio choice explicitly. They provide instances within the transport domain where portfolio choice may be appropriate, such as combined mode-destination choice and the choice of types of activities to which one allocates one’s time. The authors present design principles for designing portfolio SC surveys such that parameters for both own and cross-effects can be estimated.

6.1.2. Strategic-tactical form

The second challenge encountered was a choice situation structure which we have termed ‘strategic-tactical’.

In the context under study, an a priori [strategic] commitment to subscribe to an SDCS must be made in advance of [tactical] use of an SDCS vehicle. The strategic-tactical framework is proposed here as a person’s strategic choice to subscribe to an SDCS may provide them with little or no value aside from facilitating the expected tactical uses.
6.1.3. **Understandability/plausibility of the choice situation**

In considering these criteria for the SC survey instrument, a concern was whether we would be presenting appropriately-designed choice situations to survey respondents. (Richardson 2001) This may be an issue for a number of reasons, two of which we discuss here.

First, respondents may have difficulty stating behaviourally-realistic choices in situations with a high degree of unfamiliarity, or in hypothetical situations which they do not perceive to be applicable to them. (See Section 3.2) For instance, asking a respondent who has never eaten a particular type of ethnic food to choose amongst different hypothetical varieties of that food may generate behaviourally-unrealistic stated choices. As discussed in Section 2.4, a ‘pivoting’ technique is sometimes employed to mitigate any biases due to this effect, whereby the attributes of alternative options are varied in a narrow range around the respondent’s actual behaviour in a real-world situation.

Second, the definition of the choice situation may be more complex than the respondent’s cognitive capabilities can manage, particularly during the limited duration of the SC survey, perhaps due to the amount of information presented to respondents. (Hensher 2006, Richardson 2001, Sammer 2003) Hence any preferences inferred from an analysis of an SC dataset gathered in a hypothetical context where real-world conditions would provide people with greater opportunities to respond to choice situations in their own timeframe may be subject to undesirable bias.

The views of the substantive nature of the SDCS context that were informed by the findings of the gaming-simulation task led to challenges in presenting respondents with synthetic choice situations that they would perceive as applicable to them. The approach selected to address these challenges is described in Sections 6.2 through 6.4, and involves the use of an avatar (a virtual character).

6.2 ‘Avatar’ device

The proposed design involves introducing the survey respondent to an avatar in advance of the stated-choice exercises, and asking the respondent to advise their avatar in each choice situation. The difficulty in presenting respondents with complex choice situations that are plausible given their own life circumstances, without gathering a large set of information in a long-duration survey prior to the stated-choice game, motivated the desire to construct an avatar for each respondent. The avatar’s life circumstances are at the discretion of the survey designer, hence can be structured to suit the substantive issue under study and presented to the respondent as part of the description of the SC game.
Choice-making-by-proxy may be appealing in a stated-choice context for a number of reasons, such as the possibility of reducing data collection costs, though transport researchers are only now beginning to explore the possibilities. There are, however, a small number of studies largely in the domains of psychology and marketing which examined the degree of congruence between decisions made for oneself and those made on behalf of another. (c.f. Kray 2000, Polman 2010, Fischhoff 1992, Hsee and Weber 1997, McCubbin and Weisstub 1998, Stone et al. 2002, Beck et al. 2011.)

Polman (2010) found that personal decision makers tended to exhibit higher degrees of ‘attribute prominence’ in their decision-making than advisors. He interprets this as a greater tendency to seek the ability to ‘justify’ choices made on behalf of others rather than those made for oneself.

McCubbin and Weisstub (1998) analyze decision-making-by-proxy in cases of mentally-ill patients and report that any of a number of various decision-making rules are plausible – the patient’s actual needs (as determined by a professional) or their best interests (as determined by society, family members, etc.) among them. The possibility of employing any of a number of unique decision making strategies when making choices on behalf of others highlights the potential pitfalls in using a constructed avatar in an SC survey to elicit the respondent’s own preferences. (see also Stone et al. 2002.) Some studies have apparently treated data obtained from personal decision makers and advisors interchangeably (Kishi et al. 1988), though more recent research on decision-making-by-proxy has identified the potential for several sources of systematic biases.

The literature suggests that, compared to making choices for oneself, when advising others we may systematically take fewer sources of information into account (Kray 2000, Polman 2010), engage in risk-seeking/-neutral/-adverse behaviour depending on context (Hsee and Weber 1997, Stone et al. 2002, Beisswanger et al. 2003), or simply have flawed information or views regarding the advisee’s preferences (Kray 2000, McCubbin and Weisstub 1998). Perhaps counterintuitively, Kray (2000) found that ‘little [empirical] support was obtained for the argument that advisors simply think less carefully about decisions than personal decision makers.’

One strategy identified by Hsee and Weber (1997) for mitigating potential sources of bias is to design the advisee to be a ‘vivid other’ in the view of the decision maker: ‘this [self-other] discrepancy occurred only if the target of prediction was abstract and vanished if the target was vivid.’ Stone et al. (2002) report similar findings in their study of attitudes to risk, in particular no evidence of systematic biases in self-other choice situations where the advisee is known to the decision maker. Kray (2000), which reports multiple empirical studies) also examines a circumstance in which ‘vividness’ is defined in a way which bears some similarities to the methodology proposed.
for the present research. Kray’s study participants [undergraduate students] completed ‘demographic profiles’ at the beginning of the study (containing year of study, favourite class, etc.). They were then presented with a demographic profile from another participating student whom they were told that they would be advising. The authors report that ‘increasing the concreteness of the other person facilitated perspective-taking, which increases the mental overlap between the self and others.’ Beck and colleagues, in a study of people’s stated car purchasing choices, and their guess as to what another member of their household would choose, report evidence of respondents using their own preferences as an ‘anchor’ from which they then ‘adjust’ (see Tversky and Kahneman 1974) to take into account their perception of the other person’s preferences. (Beck et al. 2011) The authors report, based on their analysis of an instance in which the ‘other’ (a member of the respondent’s household) is arguably quite ‘vivid’ to the choice-maker, their conclusion that ‘proxy responses are a suitable replacement for actual choice information’. Their findings would seem to support the present application, where it is desirable for the respondent not to predict the ‘other’s’ preferences well, but rather to state choices in which their own preferences are faithfully taken into account.

In the context under study, it was attempted to maximise the mental overlap noted by Kray through designing each respondent’s avatar to have similar socio-demographic characteristics as he or she does. The avatar (named Jane/Joe for women and men, respectively, and with an on-screen cartoon illustration) is introduced to the respondent with a virtual handshake. (See Figure 6.1)

![Avatar Introduction](image)

**Figure 6.1: Sample screen introducing the survey respondent to her [his] avatar**

The avatar is designed with the same employment status and within the same age band as the respondent, and to have a similar household structure. A respondent living with his/her partner and/or children is thus presented with an avatar with the same characteristics, and whether an
avatar resides within Inner or Outer London is determined on the basis of which of these areas in which the respondent reports that they live.

On the basis of gender (two categories), age band (three), domiciling with/without one’s partner (two), presence of children in household (two), employment status (two), and location (two³), each respondent (and their avatar) is classed into one of 96 socio-demographic categories. Further, as described in Section 6.3, the choice context for each respondent’s avatar is oriented around frequently-performed activities by people within the same socio-demographic category as the respondent.

Despite these considerations, there are qualitative differences between the circumstances of the surveys in the literature and the avatar methodology we propose here; it is plausible, for instance, that the mechanisms of choice-making-by-proxy may vary with such attributes as whether the other is a real person versus a virtual avatar, whether the other is a family member or not, whether the respondent feels attracted, repelled or neither towards the other, etc. Further, studies evaluating self-other discrepancies in decision making have tended to employ rather simple choice situations, whereas the proposed SC survey instrument is relatively complex.

6.3 Design features of the AVATAR survey

Given the nature of the SDCS context and the unorthodox challenges it presents, we designed the AVATAR survey to be sensitive to the main substantive issues as well as the principles of SC choice situation design.

First, we considered the degree of complexity for the multi-day activity-travel period to present to respondents. This design choice incorporated multiple criteria:

- the theoretically-desired time period (at least multi-activity, and longer time periods being preferred to shorter ones)
- respondent burden (to be minimised),
- the specific characteristics of the SDCS context,
- compatibility with the E-NTS dataset,
- the amount of screen space available on a standard computer monitor to display strategic and tactical choices simultaneously (not to be exceeded)
- the font size of the survey instrument (not to be so small as to be illegible)

It was decided, on the basis of a considered compromise between these criteria, to present respondents with a set of representative activity-travel behaviour which would include five
archetypal out-of-home activities (including one which may be recurring in nature: i.e. commuting to work or school).

The cardinality of the set of out-of-home activities was chosen to be five as a compromise: fewer activities appeared to be an unrealistically small set, but larger numbers of activities increase the number of pieces of information to be processed by respondents geometrically and are difficult to represent on a computer monitor at a reasonable font size. Respondents would be asked to select methods of travel (and to concurrently select any required mobility resources) to ‘solve’ such a representative pattern of activity-travel behaviour on behalf of their avatar.

Designing a survey along these lines would ideally involve respondents completing multi-day activity-travel diaries, and staff processing them, prior to the respondent taking part in the main SC exercise. In this way, the SC instrument would pivot around a trace of each respondent’s recent activity-travel behaviour. It was concluded that the level of effort a methodology of this sort would require – on the part of both respondent and researcher – was infeasible except for rather small samples. Hence alternative design options were sought that could be implemented with a single interview per respondent – which led to the use of the avatar construct.

The trade-offs which interviewees perceived between fixed and usage costs in the gaming-simulation interviews led to our decision to design the SC survey instrument with a strategic-tactical aspect. It was decided that respondents would choose along two dimensions simultaneously – mobility resources and methods of travel. Mobility resources would have attributes which apply in a fixed manner regardless of the level of usage (e.g. costs of owning/maintaining a personal car, costs of subscribing to an SDCS, costs of purchasing a public transport season ticket, etc.) whilst methods of travel would have attributes which apply to usage (e.g. fuel costs, hourly SDCS charges, pay-per-ride public transport fares, etc.).

By using a computer-based instrument, an algorithm would provide respondents with a listing of consistent options along both of these dimensions in real-time. For instance, a respondent wishing to indicate a choice [on behalf of their avatar] of driving a SDCS vehicle to complete one or more journeys in the SC survey would have to first indicate that their avatar should choose to subscribe to an SDCS and bear the fixed costs associated with the subscription. Otherwise, the algorithm would show that the option of driving an SDCS vehicle was not available. At the same time, as shown on the right of Figure 6.2, the cumulative time and monetary cost associated with strategic choices of travel instruments is updated dynamically.
The decision to construct an avatar for the respondent in the SC instrument was made in order to avoid the risks of presenting people with choice situations likely to be qualitatively very different from their own activity-travel behaviour. Respondents are removed from the specifics of their activity-travel pattern, but invited to consider a proxy one drawn from people who are demographically similar to them and live in a similar location. The respondent’s own preferences are sought as they are asked to indicate what they would do if they were in their avatar’s situation. (See Figure 6.3)

The respondent is advised that their avatar is moving to Inner [Outer] London, is considering moving to one of several possible neighbourhoods, and that they will be asked to ‘help Jane [Joe] choose how to get around.’ (See Figure 6.1) Following a brief introduction to the type of choice situation
with which they will be presented, the respondent is invited to take part in four replications of the survey’s main SC exercises (presented as the four neighbourhoods to which their avatar is considering moving).

In the first two replications (the neighbourhoods), respondents are presented with only an SDCS with features of the car club operating model\(^5\). After they completed the second replication, respondents were told that the ‘Versatility’-style SDCS had been introduced commercially in London, and were presented a description of its service features.\(^6\)

The interview concludes, after completion of the four SC replications, with a set of questions forming the debrief stage. The degree of congruence between the choices made by the respondent on behalf of the avatar, and those the respondent would make if choosing for themselves in similar circumstances, is probed using a semi-structured protocol.

### 6.4 Setting of attribute levels

The attributes of each of the mobility resources were fixed for each respondent, and with the exception of season tickets, did not vary across respondents either. They were set for consistency with the pricing levels assumed in the E-NTS (See Section 5.11):

- Car ownership was presented to cost £4,000/year in fixed costs, exclusive of parking and petrol costs\(^7\)
- Bicycle ownership, including accessories, was presented to respondents to cost £150/year
- Public transport season tickets were presented to respondents living in Inner London to cost £100/month, and to those respondents living in Outer London to cost £150/month
- Car club SDCS subscription was presented to cost £50/year, in line with market condition in London.
- ‘Versatility’ SDCS subscription was presented to cost £10/month.

By maintaining these attribute values constant across respondents and SC replications, it would not be possible to identify alternative-specific parameters for the costs of holding the various mobility resources. Additionally, as can be seen from footnotes #5 and #6, the usage attributes of the SDCSs were maintained at a constant rate rather than varied in a systematic fashion. It was felt that the high degree of respondent burden inherent to the strategic/tactical and portfolio features of the survey rendered it prudent not to include further complexity which could be considered to be discretionary. This design decision implied, as discussed in Chapters seven and eight, that the marginal dis-utility associated with the journey duration and usage costs of the SDCSs would not be
separately identifiable from each other, and thus this remains an item to be considered for future research. This was deemed reasonable given the aims of the present research.

Activity-travel patterns are synthesized from the NTS’ London-resident participants (2004/05 edition). The seven-day travel diary data from that survey were segmented along the same demographic characteristics as respondents in our survey, and the five most-frequently-occurring activities in each segment were selected to comprise the avatar’s representative set of activity/travel behaviour.

In order to maximise the ‘vividness’ of the avatar’s circumstances, technical names of people’s activities were changed from generic to specific to be more easily-understood by respondents, and graphics were used to complement the text activity descriptions. (e.g. the name of the ‘visit friend/relative’ category was changed for presentation to respondents to ‘attend[ing] a barbecue at Jane’s [Joe’s] sister’s place.’ See Figure 6.2 regarding the use of graphics to supplement text activity descriptions.)

The combinations of travel times by each mode which are presented to a respondent (e.g. activity X can be accessed either by a Y-minute car journey, or a Z-minute walking journey, etc.) are drawn from the empirical distribution of such data from the E-NTS dataset, to further increase the plausibility of the synthetic choice situation being presented to the respondent.

Random draws were taken from these distributions, at a rate of 25 draws per each of the 96 demographic classes. The selection of which draw to retain for each demographic class was determined by the application of modified ‘efficient’ SC design principles.

It was decided to take a relatively small number of draws in an attempt to balance between the theoretical advantages of design efficiency (in terms of maximising the amount of statistically-relevant information extracted from a given sample size) on the one hand, and, as the empirical distributions might in their ‘tails’ have implausible combinations of travel times, avoiding taking so large a number of draws that the ‘maximally-statistically-efficient’ draws would all tend to be outside the central (and prima facie plausible to respondents) portions of the empirical distributions.

In view of the requirements for the dataset which would be output from the survey, a standard application of efficient survey design principles appeared suboptimal. Whilst classical survey efficiency measures have been developed to design complete datasets from which to efficiently identify a complete set of parameters, the task at hand here is to design part of a dataset from which to efficiently identify part of a set of parameters.
'Efficient’ design techniques in general involve choosing a combination of attribute levels to be presented to respondents such that some metric related to the asymptotic variance-covariance matrix [AVC] (of the parameter estimates) is minimised. D-efficiency is probably the most-widely-used criterion. The determinant of the asymptotic variance-covariance [AVC] matrix is calculated for a large number of candidate designs based on an *a priori* point ‘guess’ of each parameter value. As a determinant is a single scalar value, the determinant from each of the candidate designs can be directly compared and designs which do not have small determinants can be eliminated from consideration. The literature is clear, however, that the determinant of the AVC matrix is not the only possible optimisation criterion which is available to the researcher. Other possibilities with which researchers have experimented include the trace (or, alternatively and more appropriately, the product of the diagonal elements) of the AVC matrix [A-efficiency], the minimum sample size required to achieve identifiability of all parameters with no more than the researcher’s desired maximum level of variance [S-efficiency], or indices where designs which score well are those in which alternatives within each choice set are ‘finely-balanced’ in their systematic utilities [B-efficiency].

Several properties of the matrix determinant appear to have contributed to the wide use of D-efficiency as an efficiency criterion. It accounts for the co-variances of the AVC matrix – the off-diagonal elements – unlike A-efficiency which takes account of only the diagonal elements of the AVC matrix. B-efficient designs which present respondents with choice sets of several nearly-equally-attractive alternatives (e.g. where the choice probabilities based on the priors are roughly equal) do not in general maximise identifiability of parameters; in the case of binary choice the literature reports that this appears to happen in the range of 70%-30% choice probabilities rather than the 50%-50% range representing utility balance. (Kaninen 2002)

S-efficient designs, as opposed to D-efficient ones, tend to have *a priori* predicted, based on the priors] parameter variances which correspond to t-values in a relatively narrow range just meeting the criterion specified by the researcher (e.g. 1.96). A D-efficient design, as it is based on a matrix-wide ‘global’ measure, may well have some elements in its AVC matrix which correspond to t-values much larger than the minimum threshold defined as ‘acceptable’; this is inefficient if the objective is parameter identifiability as a binary yes/no measure and further decreases in t-values beyond this are not considered to be of value. If the objective is to be able to identify all parameters to a minimum degree of confidence, an efficient design in terms of sample size would provide roughly equal information (i.e. t-value) on each of the parameters. The minimum sample size consistent with this objective can then be determined by scaling all elements in the AVC up or down (through
adjusting the sample size) until the smallest [of the roughly-equal] t-value takes a value just larger than the pre-determined threshold for identifiability.

A sample AVC matrix containing arbitrary element values is shown in Table 6.1. We denote A \{A1, A2, A3\} as the subset of parameters which are in principle identifiable from the E-NTS portion of the combined E-NTS/AVATAR estimation dataset, and B \{B1, B2, B3, B4\} as the subset of parameters which are only identifiable from the AVATAR part of the dataset.

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>26.86</td>
<td>22.11</td>
<td>-1.25</td>
<td>-3.53</td>
<td>-13.79</td>
<td>-4.35</td>
<td>-3.50</td>
</tr>
<tr>
<td>A2</td>
<td>22.11</td>
<td>23.37</td>
<td>-1.57</td>
<td>-3.35</td>
<td>-10.68</td>
<td>-7.73</td>
<td>-4.17</td>
</tr>
<tr>
<td>A3</td>
<td>-1.25</td>
<td>-1.57</td>
<td>0.27</td>
<td>0.28</td>
<td>0.80</td>
<td>0.67</td>
<td>0.86</td>
</tr>
<tr>
<td>B1</td>
<td>-3.53</td>
<td>-3.35</td>
<td>0.28</td>
<td>0.56</td>
<td>1.86</td>
<td>0.90</td>
<td>0.99</td>
</tr>
<tr>
<td>B2</td>
<td>-13.79</td>
<td>-10.68</td>
<td>0.80</td>
<td>1.86</td>
<td>11.66</td>
<td>3.50</td>
<td>4.94</td>
</tr>
<tr>
<td>B3</td>
<td>-4.35</td>
<td>-7.73</td>
<td>0.67</td>
<td>0.90</td>
<td>3.50</td>
<td>6.46</td>
<td>3.06</td>
</tr>
<tr>
<td>B4</td>
<td>-3.50</td>
<td>-4.17</td>
<td>0.86</td>
<td>0.99</td>
<td>4.94</td>
<td>3.06</td>
<td>7.20</td>
</tr>
</tbody>
</table>

Table 6.1: Sample of an asymptotic variance/co-variance matrix, containing dummy values

The elements potentially of interest are the black (bottom right, [B by B]) and grey (top right [A by B] and bottom left [B by A]) sections of the matrix in Table 6.1. Elements in the black sub-space are the (co)variance elements of parameter subset B, such as those related to SDCS subscription and usage, which we attempt to identify from the observations in the AVATAR survey dataset. Elements in the grey space are the co-variances between parameters in subset A and B; large co-variances in this space appear undesirable based on the same arguments as those for using D- or S-efficient criteria rather than A-efficiency.

Criteria considered for efficient survey design using the matrix sub-spaces potentially of interest are:

1) **D-efficiency-based**: The determinant of the white matrix sub-space \([A by A]\) subtracted from the determinant of the full matrix \([AB by AB]\). This appears problematic as the dimensions of the two matrix spaces are different, hence the units of the two determinants will be different, and they cannot be subtracted with logical consistency. Raising each of the determinants to the inverse of its dimension (e.g. number of rows or columns) would, however, scale them appropriately for subsequent differencing.

2) **D-efficiency-based**: The summation of the three determinants of the black matrix sub-space and two distinct grey sub-spaces. This is only possible under the special circumstance where \(A = B\); in the more general circumstance where \(A \neq B\) there is no defined determinant for the rectangular [non-square] matrices \(AB\) and \(BA\).
3) **D-efficiency-based:** The product of the elements of a vector of length $AB$ where the elements are the elements [or equivalently half of the elements, given the symmetry], in any arbitrary ordering, from the black and grey spaces.

4) **D-efficiency-based:** The white matrix sub-space [A by A] is replaced with elements of the same value for each candidate design, with one being the obvious [arbitrary] choice. The determinant of the full matrix [$AB$ by $AB$] is then calculated following this ‘filtering’ of the AVC matrix.

5) **S-efficiency-based:** The maximum of the t-values corresponding to the diagonal values in the black matrix sub-space (there being no elements within the grey space which are on the diagonal of the AVC matrix).

6) **S-efficiency-based:** The maximum of the t-values corresponding to all values in the combined black and grey matrix sub-spaces. For off-diagonal elements, where the t-value is not defined, a proxy value could be estimated by taking into account both of the corresponding $i$ and $j$ parameters, perhaps calculated by dividing the square root of the product of the $i$ and $j$ parameters by the $ij$ co-variance.

S-efficiency exhibits demonstrable advantages vis a vis D-efficiency as a criterion for survey design efficiency, provided that the researcher is prepared to neglect the magnitude of the off-diagonal parameter co-variances: it relies on the assumption that the researcher’s objective is identifiability of all parameters of interest in binary yes/no terms, and that minimising the sample size conditional on binary parameter identifiability is the principal concern. Given that the AVATAR survey would require a relatively-large level of expense and effort per respondent, it was decided that in this instance any S-efficiency-based measure can reasonably be preferred over any D-efficiency measure from the listing above. In this case, the choice of efficiency measure is between items #5 and #6, and the planned joint estimation of parameters using both the E-NTS and AVATAR survey datasets led to the choice of item #5 as the criterion for selecting amongst candidate designs of attribute levels.

The process of generating the efficient AVATAR survey design was facilitated by modifying a freeware spreadsheet which was coded for the purpose of preparing D-efficient survey designs (Rose and Bliemer 2010) As the design would be tailored to each respondent’s demographic
qualities, it was not possible (unless the precise demographic composition of the sample were to be known) to estimate the required sample size with certainty (such that the predicted minimum sample size would be accurate to the degree that the parameter priors were correctly-guessed.)

Some insight was obtained by taking the simple arithmetic average of the predicted-required sample size across all 96 classes. This was found to be 204, though with wide variance across the different demographic classes (standard deviation of 281). This is not surprising, as it is quite plausible that people in some particular demographic classes might see very little value to them in subscribing to/using an SDCS, thus leading to the generation of very little information as to the structure of their preferences relating to SDCSs.

The ‘required’ sample size could in principle have been estimated by simulating a population of expected AVATAR survey respondents; this would have been a purely academic calculation, however, as the resources available for field data collection dictated that the maximum feasible sample size would be 75 respondents. It also bears noting that the ‘efficient design’ process was undertaken exclusively for the ‘portfolio’ level of choice-making; the efficiency of the selected design at estimating the ‘mode-choice’ level parameters was not considered. These design compromises were accepted in the interests of pragmatism, and in view that the aims of this research are to develop techniques for understanding SDCS take-up and usage, rather than to develop forecasts for a particular application to a particular level of statistical confidence.

6.5 Sampling and fieldwork methodology

As noted above, resources were available to support a sample size of approximately 75 respondents. It was decided to set the sample frame to be Londoners holding driving licences, and that the sample would be divided into three quotas:

- 25 residents of Greater London who have their own car (they must be the one who drives the car the most)
- 25 residents of Greater London who have a driving licence but do not have their own car (regardless of whether someone else in their household has a car)
- 25 residents of Greater London who are current subscribers to a car club SDCS

The author was responsible for interviewing five respondents from each of these three quotas, for a total of 15 interviews, with the remaining 60 to be performed by SRA Ltd, a firm with expertise in undertaking complex field data collection efforts.
Whilst it was recognised that obtaining a sample that was statistically representative of Londoners at large was not achievable, representative diversity was sought within each of these quotas, with a target of two or more respondents in each quota to come from each cell of a two by two by two cross-tabulation:

- Employed / not in employment
- Inner London / Outer London resident
- Age 16 – 34 / 35+

These criteria proved unattainable given the available resources for recruiting participants; all but four of the respondents were recruited via the interviewers’ personal/professional contacts (a pragmatic rather than optimal recruitment method.) The achieved sample quotas were:

- 22 of the target 25 car club SDCS subscribers
- 18 of the target 25 non-SDCS-subscriber car owners (a total of 25 car owners were interviewed; the remaining seven car owners were also car club SDCS subscribers)
- 32 of the target 25 non-SDCS-subscribing non-car-owners

As the sample was not representative of Londoners, generalising the substantive findings from analysis of the resulting dataset to Londoners at large will require caution, statistical weighting, and the exercise of judgment.

The field data collection effort commenced with a training session held on 10 February 2011, attended by both of the two interviewers from SRA and their line manager. Three meetings were subsequently held with the field interviewers to identify and track issues arising during the data collection task. Field data collection began on 14 February and was completed on 21 March. Data submitted from SRA were reviewed, with approval and acceptance of the data granted on 3 May.

6.6 Revisions following piloting of the SC instrument

In the field testing phase of the data collection, the three trained interviewers performed a total of 11 interviews and met to discuss procedural and data quality issues. The principal findings from the field testing appear to be generalisable and to warrant changes to the instrument package, though due to the limited total sample size it was decided to include all 72 interviews in the sample for the analysis portion of this research.

There were several indications that the choice experiment was sufficiently complex and burdensome that reliability could be compromised if it was continued to be introduced within a fully self-administered computer-aided personal interview (CAPI) format, without verifying that the
respondent had satisfactorily understood the functionality of the key components of the instrument package. A number of respondents, for instance, reported during the debriefing stage of the interview that they had not been aware of one or more of the ‘strategic’ choice options. Several likewise reported confusion regarding the links between the ‘strategic’ and ‘tactical’ choice dimensions, such as the relationship between choosing to purchase a public transport season ticket and choosing to use public transport to access particular activities.

These empirical observations were interpreted as indications that the protocol unduly invited respondents to take [inauthentic] mental shortcuts in their choice-making, and hence two substantial changes were made to it.

The first change was in the method of introducing the respondent to the various parts of the main choice experiment screen. In the field-tested CAPI, respondents were progressively exposed to the layout of the screen, and to text communicating the various functions. The weaknesses of this approach were (a) that the on-screen text was necessarily limited in its length, and (b) that respondents could easily ‘opt out’ of learning particular pieces of information. The protocol was changed to a directed practice, using all parts of the screen in a standardised order. Following a script, the interviewer instructs the respondent in a neutral manner to experiment with each on-screen function, and verifies understanding before the respondent is allowed to proceed to the next function.

The second change was to the layout of the choice experiment screen, to make the linkages between the strategic and tactical choice options much clearer. (See Figure 6.4.) The functional linkage between the two choice dimensions became reflected by physical proximity and vertical alignment on the screen.
A number of other survey design issues that are familiar from other types of SC experiment were addressed in the field test. An example was the possibility that the characterization of relatively ‘green’ modes such as walking, bicycling or carsharing in the choice experiment could give rise to social desirability bias in the advice given to avatars (indeed, bicycling was implausibly popular in the field test).

### 6.7 Characteristics of the AVATAR survey sample

The demographic characteristics of the 72 respondents are:

- 54% men, 46% women (55%/45% of Londoners with a driving licence)
- Age distribution:
  - 1% (1%) aged 16 – 19
  - 31% (20%) aged 20 – 29
  - 35% (26%) aged 30 – 39
  - 17% (21%) aged 40 – 49
  - 10% (16%) aged 50 – 59
  - 4% (10%) aged 60 – 69
  - 3% (7%) aged 70+
- 42% (64%) live with their partner
- Number of children living with the respondent:
  - 81% (69%) do not live with children
  - 15% (14%) live with one child
  - 3% (13%) live with two children
S Le Vine thesis

- 1% (4%) live with three or more children
- 35% [25 respondents] (63%) have a personal car (i.e. NTS definition of ‘main driver’)
- 44% (83%) have at least one car in their household
- 31% [22 respondents] (c.3%14) are a member of a car club
- 71% (73%) are in employment
- Distribution of highest attained level of education
  - Londoners at large:15
    - 4% GCSE (13%) no qualifications
    - 11% A Level (16%) GCSE grades A* - C
    - 7% Diploma (17%) A Level
    - 46% Degree (28%) Degree
    - 29% Masters/PhD (6%) Higher qualifications
    - 3% No answer
- Distribution of annual household income
  - 18% (31%) Up to £25K
  - 25% (34%)£25K to £50K
  - 39% (35%) £50K+
  - 18% No answer
- 60% (37%) live in Inner London

In summary, the sample is most distinctive in comparison with fully-licenced Londoners at large in:

- the proportion living with their partner (42% v. 64%),
- the proportion living with two or more children (4% v. 17%),
- the proportion having a personal car (35% v. 63%)
- the proportion that are members of a car club (31% v. 3%), and
- the proportion holding a degree or higher (75% v. 34%) – with the most-highly-educated being the most over-represented in the sample (29% having a Masters degree or higher v. 6%)

6.8 Weighting of the AVATAR survey dataset

It was decided, given practical limitations on the available resources, to employ a weighting strategy to correct for only the worst of the biases in the data sample’s characteristics. It was thus decided to weight the dataset to account for the over-representation of car club SDCS subscribers and the under-representation of car owners in the sample, relative to Londoners at large having driving licences.
This led to the following weights used in the subsequent analysis:

- Car club SDCS subscribers: 0.11
  
  $$0.11 \times 22 \text{ respondents} = 2.4$$
  
  $$2.4 / 72 \text{ respondents} = 0.03$$

- Non-car-owning and non-SDCS subscriber respondents: 0.76
  
  $$0.76 \times 32 \text{ respondents} = 24.3$$
  
  $$24.3 / 72 = 0.34$$

- Car-owning respondents: 2.52
  
  $$2.52 \times 18 \text{ respondents} = 45.4$$
  
  $$45.4 / 72 = 0.63$$

### 6.9 Descriptive survey results

Table 6.2 presents correlations from the results of the AVATAR survey dataset. The data points fall into four categories:

1) The choices made by respondents in the SC exercises

2) Respondents’ personal characteristics

3) Respondents’ responses to a series of attitudinal questions and questions about how they related to the choice experiment

4) The duration of each interview

A number of the correlations which relate to the design of the AVATAR survey are discussed here, starting with those that tend to support the design decisions made prior to field interviewing.

Choosing to ‘purchase’ a car was strongly-correlated with real-world car ownership, and with living in Outer London, and negatively-correlated with [real-world] choosing to pay more for environmentally-friendly products and propensity to share resources (as captured by willingness to share a table at a busy restaurant.)

Those respondents choosing to ‘purchase’ and use bicycles in the game tended to be young, male, not in employment, and living in Inner London. Thus, the strong correlations with car and bicycle activity in the AVATAR survey appear to be rather plausible, lending some support to the hypothesis that the choices people made [on behalf of their avatar] in the survey are reasonably authentic with respect to choices they would make for themselves in a similar real-world context.

Being employed and having a higher income were found to be positively-correlated with taxi use, also an intuitive finding.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Parental income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children present</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income goal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income strategy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income variation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income adequacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The average interview lasted 36 minutes; the best predictor of a longer-than-average duration was respondent age, with, quite interestingly, a respondent’s stated tendency to plan their personal activities ‘several days in advance’ a close second. Figure 6.5 shows the cumulative distribution of the duration of interviews.

Figure 6.5: Cumulative distribution of the duration of the interviews

The three respondents whose interviews lasted less than 15 minutes were all under age 30; they ‘purchased’ bicycles in 11 of 12 opportunities and ‘used’ those bicycles in 35 of 60 opportunities. The three respondents whose interviews lasted longer than 75 minutes were also distinctive:

- All three were over age 40
- All indicated that their in-game choice-making was relatively ‘close in your (sic) thinking to how you make choices for yourself’; (average six on seven-point Likert scale)
• All indicated a strong propensity to plan their personal activities (seven/seven), share a table at a busy restaurant (six/seven) and willingness-to-experiment (six/seven).
• They chose to travel on public transport most frequently (31 of 60 opportunities), followed by walking (13/60) and cycling (7/60) with all other modes used five or fewer times.

Patterns in the number of mouse click actions performed by each respondent offer further insight into the ways in which respondents interacted with the AVATAR instrument. The highest correlation with the number of mouse clicks was found to be gender; men performed on average 147 whilst women performed 127. The number of mouse clicks also correlated strongly with interview duration, indicating that respondents taking rather long durations to proceed through the interview tended to also interact with the instrument more intensively than average. Interview duration correlated less closely with ‘information-seeking mouse clicks’ than overall mouse clicks, which may be indicative of long-duration respondents engaging in other sorts of ‘discretionary’ mouse clicks, such as changing their preliminary ‘purchase’ and ‘use’ choices before confirming them. A plausible reason for respondents doing so would be to explore how various purchase/use patterns perform in terms of aggregate travel duration, aggregate out-of-pocket travelling costs, and aggregate ‘purchase’ costs. The only way for a respondent to access these aggregate costs for some purchase/use combination of choices (aside from mental arithmetic) was to make a preliminary selection of the pattern, which could be changed if so desired before confirming one’s choices and moving on to the next stage of the interview.

The level of bicycling choices in the survey was negatively-correlated with interview duration. In general, cycling itineraries in the survey design were short-duration, which was paired with no out-of-pocket cost and a relatively low cost of acquisition. Taken together with the implausibly-high mode share of cycling in the AVATAR survey (28%), this may indicate that a substantial portion of respondents followed a choice-making process of minimizing these three quantities and thus avoiding high-cognition consideration of the qualitative aspects of cycling.

In a somewhat unexpected finding, the correlations between subscribing to the ‘Versatility’ SDCS and number of mouse click actions, both overall and information-seeking ones, were moderately negative. This does not however mean that they did not access information about this form of travel (with which they would nearly-certainly have been unfamiliar) as all respondents were presented with two screens of information regarding ‘Versatility’ before they had the opportunity to choose it. These correlations may be explained in part by the negative correlations of respondent age with
both participating in ‘Versatility’ and number of mouse actions to seek information, though findings of this sort would appear to be ripe for further inquiry.

Three of the post-game questions in the survey were intended to provide insights into the performance of the novelties in the AVATAR survey instrument:

- **How similar or different was this game to how you think about getting around?**
  This question was designed to investigate how well the strategic-tactical (i.e. purchase/use) aspect of the survey matched people’s real-world choice-making in this area.

- **You just gave some advice to Jane/Joe. As you thought through Joe/Jane’s choices, how close was your thinking to how you make choices for yourself?**
  This question was designed to investigate similarities/distinctions between respondents’ choice-making processes when choosing for themselves and when choosing on behalf of their avatar.

- **Would you say Jane’s/Joe’s routine of activities is similar or different to yours?**
  This question was designed to investigate how well respondents perceived the match between their lifestyle and that of their avatar to be.

Male respondents reported somewhat higher levels of similarity (an average of five of seven on a Likert scale) between this game and how they think about getting around than women respondents (four/seven); the gender differences in responses were substantially less for the latter two of these questions.

The presence of children in a respondent’s household also correlated negatively with the degree-of-match with how respondents think about getting around (as indeed it was also found to correlate with the latter two of these questions.) Household income correlated positively with this measure, as did the number of information-seeking mouse actions. Of note there was also a moderate negative correlation of degree-of-match how respondents think about getting around with willingness to share a table at a busy restaurant.

The most distinctive pattern associated with the question on similarity in choicemaking for oneself versus on behalf of their avatar related to respondents’ household income – higher-income respondents tended to report higher ‘similarity’ in self-avatar choicemaking than their lower-income counterparts indicated. We note, however, that this question can only identify ‘self-avatar’ differences in choicemaking processes of which respondents are aware and hence are able to identify – any differences in mental processes between choosing for oneself and on behalf of an
avatar that are not consciously identifiable to the respondent would not be captured in these responses.

The strongest relationship with the last of these three questions – which addressed the ‘similarity’ of the respondent’s routine with the avatar’s – was a negative relationship with living with children. In other words, respondents reported that their avatar’s routine of activities was a better match with their own routine than more-formally-educated respondents.

In summary, the manner in which respondents engaged with the AVATAR survey instrument appears to have differed in systematic (though generally not counter-intuitive) ways, which would seem to warrant follow-up studies to support a wider application of these methods. Further, notwithstanding several of the specific patterns of correlation discussed here, the broadly-intuitive patterns described in this section tend to provide support for using the dataset from this survey in the subsequent analyses.

6.10 Summary and conclusions

This Chapter presented the design, fieldwork, and descriptive results of the AVATAR stated-choice survey.

In recognition of the unique criteria for the output dataset, the survey was designed to incorporate several novel features. Amongst these features are the requests that the respondent concurrently considers both strategic and tactical dimensions related to the synthetic choice situation, and that the respondent state their choices in the context of giving advice to an avatar. The decision to employ the avatar device was made in order to address the dilemma of, one the one hand, it being infeasible to gather a large set of information about the respondent prior to the SC exercise, and, on the other, to present choice situations to the respondent which are plausible given their life circumstances but not overly-prescriptive.

Field data collection (n=72) was undertaken in early 2011; the field testing phase led to several revisions to the instrument package. The data sample was identified to be rather distinctive (with respect to London-resident driving-licence-holders at large) in several regards. A strategy was thus developed for weighting the observations to take the worst of these biases (the proportions of the sample that subscribe in a car club SDCS and that own personal cars) into account.

Descriptive analyses of results from the AVATAR survey were found to provide support for taking it forward to substantive analysis, which is discussed in Chapter seven.
Notes

1 In the terminology of (Wiley and Timmermans 2009) own effects refer to parameter vectors found in traditional discrete choice analyses, whilst cross-effects are parameters which capture possible inter-alternative complementarity/substitutability; that, in other words, the attractiveness of an n-alternative portfolio could be more or less attractive than the sum of the attractiveness of the n elements individually.

2 It is plausible that people may draw value from access to travel instruments such as an SDCS subscription aside from pure functionality. For instance, automobiles may serve as ways for the holder to publicly display signals about themselves (see (Kurani et al. 2007) for further discussion of this point). For the purposes of this paper, it is only necessary that functional value as described above is present and plays a significant role, not that such ‘non-functional’ value is absent.

3 Language on survey instruments such as “you are making choices for a friend who is fairly similar to you” (Stone et al. 2002) and “Before you start, please look around and see who sits closest to you. Do not talk to or disturb that person, but look at him/her for a second and remember how he/she looks” (Hsee and Weber 1997) has been employed to define the “vivid” advisee condition, whilst “somebody somewhere in the [same country]” (Hsee and Weber 1997) and “imagine that you have been approached by a typical student who is about to graduate” (Kray 2000) are typical examples of the language used to define the “generic” advisee condition.

4 Respondents living outside of London are presented with an avatar living in Outer London.

5 The car club SDCS was described as:

‘A car club provides by-the-hour hire cars in Jane’s neighbourhood that can be used without paperwork or waiting.
Jane [Joe] would need to reserve a car ahead of time.
Jane can use her mobile or the internet to make a reservation.
It costs £5 per hour, and Jane would be charged from the time she picks the car up until she brings it back.
Jane would be billed each month for her use, and pay £50 per year for membership.
Jane would unlock the car doors by touching a smartcard to the windscreen. Each time she uses it, she must bring the car back to where she started.
Petrol is included in the £5 hourly fee, but not parking.
Other members use the cars at different times, so there is a 1 in 20 chance that a car isn’t available when Jane wants.
Jane is responsible to check for damage before taking a car club car.’

6 The ‘Versatility’ SDCS was described as:

‘A large car hire company has introduced a service they call ‘Versatility’. Versatility cars are parked on street corners around London.
You walk up to a car, with or without a reservation, and then drop it off at any on-street parking space in London.
You only pay until you drop it off at the end of your journey – 20p a minute (parking is free).
Versatility isn’t perfect...
It costs £10 a month to subscribe.
At busy times it can be difficult to find a car nearby, and you may have to wait.
Every so often you might find a car you would like to use in an unclean state or low on fuel.
The cars can’t be taken outside London.’
7 Whilst it is recognised that many types of cars and bicycle existing in the marketplace at different price points with different features, both are treated as homogeneous products in this study. This could be relaxed in future applications, perhaps by segmenting the sample into different classes by income level and targeting the mix of bicycle and car features/prices to each class.

8 A design is not in general efficient except in the exceptional circumstance that the guesses of the priors were precisely correct. In recent research, researchers have begun to experiment with priors which are ‘mixed’ distributions rather than point guesses, but an issue remains, even assuming that the plausible range in which a parameter lies can be ‘correctly’ identified, how to assign a probability density function within this range. (see Bliemer et al. 2008)

9 Other proposed—but-rarely-used measures of survey design efficiency are described in (Kessels et. al. 2006) and (Ibanez et al. 2007), along with their advantages and drawbacks.

10 At first glance the shift from D-efficiency to S-efficiency would seem to be a step backwards in that, as with A-efficiency, S-efficiency does not explicitly account for off-diagonal elements of the AVC matrix. A structural property of an AVC matrix, however, is that the largest [absolute value] in the matrix must lie on the diagonal, hence neglecting the off-diagonal elements is warranted since the S-efficiency criterion is to minimise the largest value on the diagonal and hence in the entire AVC matrix.

11 There does seem to be the possibility that two candidate designs could yield identical S-efficiency measures by their minimum t-values being equal, though intuitively it would seem possible to select one or the other on a systematic basis. For instance let us consider two designs, one where the t-values corresponding to the diagonal elements of the AVC matrix are {1,2,3} and the other where they are {2,2,3}. Intuitively one would select the first design as being more efficient, because it is precisely equal to the second in the identifiability of the least-identifiable parameter, and performs better in identifying the remaining parameters (the ‘non-least-identifiable’ parameters.) This special case does not appear to have been discussed by Rose et al. in the literature proposing S-efficiency, hence would seem to be worth considering exploring as a future enquiry. Perhaps an extension to S-efficiency could be proposed that follows the ‘Premier League principle’ — i.e. in the case of ties in the primary metric, the algorithm then considers secondary, tertiary, etc. metrics (in this case, the second-worst t-value, third-worst t-value, etc.)

12 £5000 was budgeted for data collection; the cost per respondent was therefore anticipated to be £67. 72 interviews were completed successfully, the actual cost per respondent was thus £69. Respondents were provided £15 for their participation (this is included in the cost-per-respondent.)

13 Values in bold font (in parentheses) are the corresponding data for driving licence holders residing in London. Unless otherwise noted, data are from National Travel Survey 2004/05. (2004/05 chosen for consistency with other data used in this research.) For readers having a black and white copy of this document, the AVATAR survey data point is always presented to the left of the Londoners-at-large data point in this listing.

14 Due to the rapid growth in car club membership (there were fewer than 3K car club SDCS subscribers across the UK in 2005), this data point is based on the Oct 2010 subscription figures. (Morgan 2011) This is calculated from 133K members and a London 2010 population of c.7.9M. This corresponds to c.2.1% of adult Londoners, and c.3.3% of the 63% of adult Londoners who hold full drivers’ licences. There may some be double-counting of multi-car-club subscribers due to the methodology in which the data are collected; these figures should thus be treated with some caution.

15 It is noted that qualification categories in the AVATAR survey are not fully consistent with these figures, which are sourced from the 2006 Labour Force Survey, and also that the figures on educational attainment are for all adult Londoners, irrespective of licence-holding.
Chapter 7: Independent analyses of E-NTS and AVATAR datasets

Chapter four described the concepts underlying the proposed strategic portfolio [StraP] system, and Chapters five and six presented the datasets which were generated to support the empirical analysis. This chapter presents results from independently analysing the two datasets. Chapter eight contains results from linking them, which forms the StraP system, and presents the substantive forecasts from its application.

Section 7.1 contains an overview of the empirical application, which is followed in Section 7.2 with a discussion of the model specifications and estimation process. Section 7.3 presents results from estimating the StraP system using a simulated dataset. Section 7.4 contains results from estimation using only the enriched National Travel Survey [E-NTS] dataset; Section 7.5 does the same for the dataset from the Advanced Vehicular/Activity/Travel And Resource [AVATAR] survey. Section 7.6 presents results from estimating the more-traditional models of mode choice, using [separately] both the E-NTS and AVATAR datasets. Section 7.7 then summarises and concludes this chapter.

7.1 Overview

In the broadest of terms, this empirical application consisted of parameterising the StraP model system by identifying parameter sets which provide the best fit to the dependent variables (observed choices) on the basis of independent variables, and then applying this parameterised system to forecast choices in changed circumstances: the entrance of various SDCSs to the marketplace in London.

The E-NTS and AVATAR survey datasets are pooled together in the estimation phase, as they contain complementary information. The E-NTS is the larger of the two datasets, is based on people’s real-world behaviour, and contains observations of some means of travel that are not observed in the AVATAR survey. The AVATAR survey is the source of information on people’s tastes regarding the characteristics of SDCSs, which are not observed in the E-NTS dataset (and generally were not available to E-NTS respondents at the time the dataset was collected.)
Each person’s set of observed holdings is predicted, on the basis of the holding costs of each mobility resource and its value to the person in performing travel associated with their perceived activity set.

The method of travel used for each of a person’s journeys is modelled separately, taking into account that a person’s portfolio of holdings restricts which modes of transport are available for them to perform their travels.

In the application of the StraP system described in Chapter eight, the probabilities of a person holding each particular combination of travel products/services are estimated, and Monte Carlo simulation is used to identify a ‘choice’ for that person. The person’s ‘mode choices’ are then constrained by this set of their holdings.

### 7.2 Model estimation and data structure

This section discusses two issues relating to model estimation and the structure of the E-NTS and AVATAR datasets. Section 7.2.1 presents a brief overview of the complexity of estimating the ‘portfolio’ and ‘mode’ choice levels. Section 7.2.2 investigates patterns of the choice observations in both datasets and the observed relationship between holding and usage. These patterns are of interest as the ‘portfolio’ models aim to replicate the empirical observations; the corresponding results of the degree to which the model is able to replicate the composition of people’s ‘portfolios’ can be found in Sections 8.2 and 8.3.

#### 7.2.1 Complexity of model estimations

The ease of estimating parameters was found to differ markedly for the models of people’s observed use of travel methods and those of people’s observed ‘portfolio’ of holdings. It appears that this is in part due to characteristics of the datasets, and in part due to the relative complexity of the functional forms of the respective models. (See Table 7.1.)

In order to maximise the available effort for studying the principal innovations of this research (identifying possible strategic behaviour and the composition of a ‘portfolio’ from a set of elements) the utility functions for the models of people’s choices of transport modes were specified to be relatively simple in form (linear-in-the-parameters, with simple error structures.)
Table 7.1: Comparison of data and model form characteristics for estimating the ‘mode’ and ‘portfolio’ choice models

<table>
<thead>
<tr>
<th>Form of utility function</th>
<th>Mode Choice Models</th>
<th>Portfolio Choice Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U^i_{mji} = V^i_{mji} + \varepsilon^i_{mji}$</td>
<td>$U^i_d = v^{\text{i,non-travel}}<em>d + \left( \sum</em>{i} \lambda_{ij} \cdot \ln \sum_{m,j} e^{v^{\text{i,travel}}<em>{mji} \cdot x</em>{\text{travel}}} \right) + \varepsilon^d_d$</td>
<td></td>
</tr>
<tr>
<td>Note the $V^i_{mji}$ term is linear-in-the-parameters</td>
<td>Note the $v^{\text{i,non-travel}}<em>d$ and $v^{\text{i,travel}}</em>{mji}$ terms are each linear-in-the-parameters</td>
<td></td>
</tr>
</tbody>
</table>

| Number of observations [E-NTS / AVATAR] | 10,575 / 1,440 | 704 / 288 |

| Order of magnitude of the number of terms in each utility function [E-NTS / AVATAR] | Up to 10 | c.5,000 / c.200 |

| Order of magnitude of duration of estimation | Up to 15 minutes | Up to 72 hours |

7.2.2. Descriptive comparison of observations in the E-NTS and AVATAR survey datasets

A summary of the observed ‘portfolio’ and ‘mode’ choices in the E-NTS and AVATAR survey dataset is shown in Tables 7.2 and 7.3.

The largest difference between the weighted and unweighted observations in the E-NTS is the proportion of walking journeys, due to the reporting of short walking journeys on only the seventh day of respondents’ travel diaries. The sampling for the AVATAR survey was less rigorous than that of the E-NTS, hence the differences between the weighted and unweighted observations are larger.
### E-NTS dataset

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Number (un-weighted)</th>
<th>Number (weighted)</th>
<th>Percent (un-weighted)</th>
<th>Percent (weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>85</td>
<td>85</td>
<td>12%</td>
<td>12%</td>
</tr>
<tr>
<td>B</td>
<td>160</td>
<td>173</td>
<td>23%</td>
<td>25%</td>
</tr>
<tr>
<td>C</td>
<td>101</td>
<td>93</td>
<td>14%</td>
<td>13%</td>
</tr>
<tr>
<td>D</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>4%</td>
</tr>
<tr>
<td>E</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>22%</td>
</tr>
<tr>
<td>AB</td>
<td>55</td>
<td>55</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>AC</td>
<td>60</td>
<td>62</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>AD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>AE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>BC</td>
<td>46</td>
<td>42</td>
<td>7%</td>
<td>6%</td>
</tr>
<tr>
<td>BD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>8%</td>
</tr>
<tr>
<td>BE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>8%</td>
</tr>
<tr>
<td>CD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>8%</td>
</tr>
<tr>
<td>CE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>24%</td>
</tr>
<tr>
<td>DE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>ABC</td>
<td>25</td>
<td>29</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>ABD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2%</td>
</tr>
<tr>
<td>ABE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>ACD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2%</td>
</tr>
<tr>
<td>ACE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>5%</td>
</tr>
<tr>
<td>ADE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>BCD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>4%</td>
</tr>
<tr>
<td>BCE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>10%</td>
</tr>
<tr>
<td>BDE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>CDE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>ABCD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>ABCE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2%</td>
</tr>
<tr>
<td>ABDE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1%</td>
</tr>
<tr>
<td>ACDE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>BCDE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>ABCDE</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0%</td>
</tr>
<tr>
<td>None</td>
<td>172</td>
<td>165</td>
<td>24%</td>
<td>23%</td>
</tr>
<tr>
<td>Total</td>
<td>704</td>
<td>704</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### AVATAR dataset

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Number (un-weighted)</th>
<th>Number (weighted)</th>
<th>Percent (un-weighted)</th>
<th>Percent (weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>13</td>
<td>22</td>
<td>5%</td>
<td>8%</td>
</tr>
<tr>
<td>B</td>
<td>36</td>
<td>33</td>
<td>13%</td>
<td>12%</td>
</tr>
<tr>
<td>C</td>
<td>68</td>
<td>64</td>
<td>24%</td>
<td>22%</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>2</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>E</td>
<td>22</td>
<td>15</td>
<td>8%</td>
<td>5%</td>
</tr>
<tr>
<td>AB</td>
<td>17</td>
<td>21</td>
<td>6%</td>
<td>7%</td>
</tr>
<tr>
<td>AC</td>
<td>9</td>
<td>21</td>
<td>3%</td>
<td>7%</td>
</tr>
<tr>
<td>AD</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>AE</td>
<td>4</td>
<td>10</td>
<td>1%</td>
<td>4%</td>
</tr>
<tr>
<td>BC</td>
<td>14</td>
<td>17</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>BD</td>
<td>8</td>
<td>6</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td>BE</td>
<td>8</td>
<td>12</td>
<td>3%</td>
<td>4%</td>
</tr>
<tr>
<td>CD</td>
<td>8</td>
<td>2</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td>CE</td>
<td>24</td>
<td>25</td>
<td>8%</td>
<td>9%</td>
</tr>
<tr>
<td>DE</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>ABC</td>
<td>4</td>
<td>3</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>ABD</td>
<td>2</td>
<td>0.2</td>
<td>1%</td>
<td>0.1%</td>
</tr>
<tr>
<td>ABE</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>ACD</td>
<td>2</td>
<td>0.2</td>
<td>1%</td>
<td>0.1%</td>
</tr>
<tr>
<td>ACE</td>
<td>5</td>
<td>8</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td>ADE</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>BCD</td>
<td>4</td>
<td>0.4</td>
<td>1%</td>
<td>0.2%</td>
</tr>
<tr>
<td>BCE</td>
<td>10</td>
<td>6</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td>BDE</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>CDE</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>ABCD</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>ABCE</td>
<td>2</td>
<td>0.2</td>
<td>1%</td>
<td>0.1%</td>
</tr>
<tr>
<td>ABDE</td>
<td>1</td>
<td>1</td>
<td>0%</td>
<td>0.3%</td>
</tr>
<tr>
<td>ACDE</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>BCDE</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>ABCDE</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>None</td>
<td>23</td>
<td>19</td>
<td>8%</td>
<td>7%</td>
</tr>
<tr>
<td>Total</td>
<td>288</td>
<td>288</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Summary of observations

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Number (un-weighted)</th>
<th>Number (weighted)</th>
<th>Percent (un-weighted)</th>
<th>Percent (weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>225</td>
<td>231</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>including A</td>
<td>225</td>
<td>231</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>All</td>
<td>286</td>
<td>299</td>
<td>42%</td>
<td>42%</td>
</tr>
<tr>
<td>including B</td>
<td>286</td>
<td>299</td>
<td>42%</td>
<td>42%</td>
</tr>
<tr>
<td>All</td>
<td>232</td>
<td>226</td>
<td>32%</td>
<td>32%</td>
</tr>
<tr>
<td>including C</td>
<td>232</td>
<td>226</td>
<td>32%</td>
<td>32%</td>
</tr>
<tr>
<td>All</td>
<td>29</td>
<td>11</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>including D</td>
<td>29</td>
<td>11</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>All</td>
<td>76</td>
<td>77</td>
<td>26%</td>
<td>26%</td>
</tr>
<tr>
<td>including E</td>
<td>76</td>
<td>77</td>
<td>26%</td>
<td>26%</td>
</tr>
</tbody>
</table>

Codes:  
A: Owning a personal car  
B: Owning a public transport season ticket  
C: Owning a bicycle  
D: Subscribing to a car-club-style SDCS  
E: Subscribing to a ‘Versatility’-style SDCS

Table 7.2: Summary of observed portfolio choices in the E-NTS and AVATAR datasets
<table>
<thead>
<tr>
<th>Mode</th>
<th>Number (un-weighted)</th>
<th>Number (weighted)</th>
<th>Percent (un-weighted)</th>
<th>Percent (weighted)</th>
<th>Number (un-weighted)</th>
<th>Number (weighted)</th>
<th>Percent (un-weighted)</th>
<th>Percent (weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3,575</td>
<td>34%</td>
<td>2,922</td>
<td>28%</td>
<td>164</td>
<td>241</td>
<td>11%</td>
<td>17%</td>
</tr>
<tr>
<td>B</td>
<td>2,911</td>
<td>28%</td>
<td>2,658</td>
<td>25%</td>
<td>404</td>
<td>382</td>
<td>39%</td>
<td>27%</td>
</tr>
<tr>
<td>C</td>
<td>157</td>
<td>1%</td>
<td>147</td>
<td>1%</td>
<td>401</td>
<td>384</td>
<td>28%</td>
<td>27%</td>
</tr>
<tr>
<td>D</td>
<td>1,442</td>
<td>14%</td>
<td>2,961</td>
<td>28%</td>
<td>225</td>
<td>225</td>
<td>16%</td>
<td>14%</td>
</tr>
<tr>
<td>E</td>
<td>149</td>
<td>1%</td>
<td>127</td>
<td>1%</td>
<td>96</td>
<td>105</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>F</td>
<td>2,341</td>
<td>22%</td>
<td>1,758</td>
<td>17%</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>G</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>29</td>
<td>13</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>H</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>121</td>
<td>117</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>Total</td>
<td>10,575</td>
<td>100%</td>
<td>10,575</td>
<td>100%</td>
<td>1,440</td>
<td>1,440</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Codes:  
A: Driving a personal car  
B: Public transport  
C: Riding a bicycle  
D: Walking  
E: Taking a taxi or minicab  
F: Riding as a car passenger  
G: Driving a car-club-style SDCS car  
H: Driving a ‘Versatility’-style SDCS car

Table 7.3: Summary of observed mode choices in the E-NTS and AVATAR datasets

In the AVATAR dataset the relationship between owning mobility resources and using the methods of transport that they enable was defined, and coded in the AVATAR survey software, to not allow exceptions to these relationships (driving a car was defined to require first having purchased a personal car). These relationships are not as rigid in the E-NTS dataset, though, as it arises from real-world behaviour of greater complexity than these simplistic relationships can accommodate. For instance, the functional relationships as defined for the purposes of this research neglect the possibility of driving a household car for which an NTS respondent is not the principal driver.

Of the 10,575 journeys in the E-NTS dataset, 398 (4%) did not conform to these functional relationships:

- 385 of these journeys were car driving journeys
- The remaining 13 of these journeys were cycling journeys where the NTS respondent reported not owning a bicycle.

These 398 journeys were performed by 60 (9%) NTS respondents; their usage of transport modes are thus not fully consistent with their portfolio holdings. (It was decided to allow these records to remain in the sample.)

Whilst the AVATAR survey did not allow respondents to use a mobility-resource-requiring mode of transport without indicating that they would make the required ‘purchase’, respondents were able to ‘purchase’ without indicating that they would ‘use’ the modes it enables. There were found to be
35 of these ‘unused’ mobility resources, representing 8% of all mobility resources ‘purchased’ by AVATAR survey respondents:

- 3 personal car purchases were not used
- 6 unused public transport season tickets
- 8 unused car club subscriptions
- 9 unused bicycles
- 9 unused ‘Versatility’ subscriptions

16 (22%) of AVATAR survey respondents chose at least one of these ‘unused’ mobility resources; with the data at hand it is not knowable whether these respondents were engaging relatively-sophisticated strategies of acquiring mobility resources for purely ‘insurance’ purposes, or whether these ‘unused’ mobility resources are indicative of difficulties in using the survey instrument. Interestingly, the ‘unused’ mobility resources were purchased more frequently by the eight respondents interviewed prior to the survey revisions than the 64 interviewed afterwards.¹

Whilst not undertaken as part of this research, it is recognised that a more subtle treatment (see the discussion in Section 4.1) of the functional relationships between mobility resources and transport methods could well provide further insights into non-traditional access to and use of cars (and bicycles) and respondent interaction with AVATAR-style survey instruments.

Tables 7.4 and 7.5 show the correlations in the observations/stated choices (‘portfolio’ holdings and ‘mode’ usage) from the E-NTS and AVATAR dataset respectively. Though it is recognised that a multivariate analysis of these observations may well identify more complex relationships than this bivariate analysis, it was decided to leave this as an avenue for future research.

<table>
<thead>
<tr>
<th></th>
<th>Car ownership</th>
<th># Car driving journeys</th>
<th>Public transport season ticket ownership</th>
<th># Public transport journeys</th>
<th>Bicycle ownership</th>
<th># Cycling journeys</th>
<th># Walk journeys</th>
<th># Taxi-minicab journeys</th>
<th># Car passenger journeys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car ownership</td>
<td>.71</td>
<td>-.09</td>
<td>-.18</td>
<td>.07</td>
<td>-.03</td>
<td>-.10</td>
<td>-.04</td>
<td>-.22</td>
<td></td>
</tr>
<tr>
<td># Car driving journeys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public transport season ticket ownership</td>
<td>-.17</td>
<td>-.27</td>
<td>.07</td>
<td>-.02</td>
<td>-.10</td>
<td>-.05</td>
<td>-.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Public transport journeys</td>
<td>.49</td>
<td>-.15</td>
<td>-.07</td>
<td>-.11</td>
<td>-.07</td>
<td>-.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bicycle ownership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Cycling journeys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Walk journeys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Taxi-minicab journeys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Car passenger journeys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.4: Correlation matrix of observations (‘portfolio’ holdings and ‘mode’ usage) from the E-NTS dataset
The highest correlation in Table 7.4 is that in the upper-left-most cell in the matrix – between [personal] car ownership and use. This is consistent with the typical finding in analyses of car use that car ownership is generally observed to be a strong correlate of car use (e.g. Simma and Axhauksen 2001, Le Vine et al. 2009.) Interestingly, whilst car ownership is negatively correlated with season ticket ownership, it is more-strongly-negatively associated with public transport usage.

Bicycle ownership and cycling were positively correlated, as one would expect, though this was a weaker relationship than that between public transport season ticket holding and public transport usage. The moderately-negative relationship between public transport activity and walking journeys is not immediately intuitive, and seems to indicate that public transport and walking tend to be substitutes for each other to a larger degree than they are complementary. A second such finding is the relatively weak correlation between cycling and walking.

As with the E-NTS correlations, the highest correlation in the AVATAR observations (see Table 7.5) is that between the number of times each respondent ‘purchased’ a car in the survey, and the number of times they indicated that they would drive a car in the game. On the whole, the correlations between the E-NTS observations and AVATAR survey stated choices (i.e. Tables 7.4 and 7.5) are structured rather similarly; the overall correlation between the cells in the two matrices is 0.74. A distinction which can be seen relates to bicycle ‘activity’ (purchasing/usage), which correlates quite negatively with most other travel choices in the AVATAR survey, though generally weakly in the E-NTS dataset.

The most notable difference relates to the magnitude of the correlations from the AVATAR dataset – these tend to be higher than the corresponding correlations from the E-NTS dataset. The average absolute value of the cells in Table 7.4 is 0.12; the comparable statistic from Table 7.5 (using only the cells which are common between the two tables) is 0.28. Thus the relationships are ‘sharper’, more structured, and more intuitive in the [stated-choice] AVATAR survey dataset than the [revealed-choice] E-NTS dataset. This highlights the attenuated context in which SC respondents acted in comparison to the complexity of the observed real-world phenomenon, and is a point to consider for future applications of similar SC designs.
7.3 Results from analysis of the ‘portfolio choice model’ using simulated choice observations based on the E-NTS dataset

Simulated portfolio-level choices were synthesised using the E-NTS dataset in order to verify the normalisation conditions outlined in Section 4.8.

Before discussing the particular attributes of the simulated datasets, we first introduce the general structure of all analyses based on the E-NTS dataset:

- The people are observed to make a variable number of journeys. In practice this varied from one to 58.
- The observation of whether each person holds each of three ‘portfolio elements’ (owning a car, owning a bicycle, owning a public transport season ticket) leads to a choice set of $3^3 = 8$ portfolio options.
- There are six modes of transport which can be used (car driving, car passenger travel, cycling, public transport, taxi/minicab, and walking). Four are common to all portfolios, two are enabled by holding a particular portfolio element, and the characteristics of one (public transport) are modified (it becomes free at the point of use) by holding a season ticket.

Table 7.6 lists the simulated datasets which were prepared, and the relevant characteristics of each. The simulated choices were generated using the known parameters listed below and attributes from the E-NTS dataset.

The simulated datasets were then used to estimate parameter sets as described at a high level in Table 7.7 and in detail in Table C.1 (the latter of which can be found in Appendix C.) The letter prefix of each of the names indicates which of the simulated datasets were used in the estimation. Unless otherwise noted in Table 7.7, any value denoted as ‘fixed’ was fixed at its ‘target’ value.
<table>
<thead>
<tr>
<th>Synthesised dataset</th>
<th>Portfolio-level ASCs</th>
<th>Mode-choice level ASCs</th>
<th>Monthly cost parameter</th>
<th>Journey cost parameter</th>
<th>Travel time parameters²</th>
<th>Salience (γ) parameters</th>
<th>Logsum (λ_travel) parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>All zeroes</td>
<td>All zeroes</td>
<td>-0.01</td>
<td>-0.1</td>
<td>Shared modes: -0.1</td>
<td>Non-shared modes: -0.05</td>
<td>All 1</td>
</tr>
<tr>
<td>B</td>
<td>All zeroes</td>
<td>All zeroes</td>
<td>-0.01</td>
<td>-0.1</td>
<td>Shared modes: -0.1</td>
<td>Non-shared modes: -0.05</td>
<td>All 1</td>
</tr>
<tr>
<td>C</td>
<td>All zeroes</td>
<td>All zeroes</td>
<td>-0.01</td>
<td>-0.1</td>
<td>Shared modes: -0.1</td>
<td>Non-shared modes: -0.05</td>
<td>All 1</td>
</tr>
<tr>
<td>D</td>
<td>Unique for each mobility resource</td>
<td>All zeroes</td>
<td>-0.01</td>
<td>-0.1</td>
<td>Shared modes: -0.1</td>
<td>Non-shared modes: -0.05</td>
<td>All 1</td>
</tr>
<tr>
<td>E</td>
<td>All zeroes</td>
<td>All zeroes</td>
<td>0</td>
<td>0</td>
<td>All zeroes</td>
<td>All zeroes</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>All zeroes</td>
<td>All zeroes</td>
<td>-0.01</td>
<td>-0.1</td>
<td>Unique for each mode of travel</td>
<td>All 1</td>
<td>1</td>
</tr>
<tr>
<td>G</td>
<td>All zeroes</td>
<td>Unique for each mode of travel</td>
<td>-0.01</td>
<td>-0.1</td>
<td>Shared modes: -0.1</td>
<td>Non-shared modes: -0.05</td>
<td>All 1</td>
</tr>
</tbody>
</table>

Table 7.6: Listing of simulated datasets and characteristics
<table>
<thead>
<tr>
<th>Model run #</th>
<th>Portfolio-level ASCs</th>
<th>Mode-choice-level ASCs</th>
<th>Monthly cost parameter</th>
<th>Journey cost parameter</th>
<th>Travel time parameters</th>
<th>Salience ((y)) parameters</th>
<th>Logsum ((A_{\text{travel}})) parameter</th>
<th>Successful termination of estimation process</th>
<th>Parameters correctly identified?</th>
<th>Log-likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^i)</td>
<td>No</td>
<td>--</td>
<td>-737.6</td>
</tr>
<tr>
<td>A2</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^i), one (y) fixed at 5.0</td>
<td>Yes</td>
<td>Yes</td>
<td>-737.6</td>
</tr>
<tr>
<td>A3</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^i), one (y) fixed at 5.0</td>
<td>Yes</td>
<td>Yes</td>
<td>-737.6</td>
</tr>
<tr>
<td>A4</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-737.6</td>
</tr>
<tr>
<td>A5</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Fixed at 5.0</td>
<td>Yes</td>
<td>No</td>
<td>-756.9</td>
</tr>
<tr>
<td>A6</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Fixed at 5.0</td>
<td>Yes</td>
<td>Yes</td>
<td>-737.6</td>
</tr>
<tr>
<td>A7</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-737.9</td>
</tr>
<tr>
<td>B1</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-684.1</td>
</tr>
<tr>
<td>B2</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-684.1</td>
</tr>
<tr>
<td>B3</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-695.5</td>
</tr>
<tr>
<td>C1</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^i)</td>
<td>Yes</td>
<td>Yes</td>
<td>-866.1</td>
</tr>
<tr>
<td>C2</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^i)</td>
<td>Yes</td>
<td>Yes</td>
<td>-866.1</td>
</tr>
<tr>
<td>C3</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-868.4</td>
</tr>
<tr>
<td>D1</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^h)</td>
<td>Yes</td>
<td>Not fully</td>
<td>-966.2</td>
</tr>
<tr>
<td>D2</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^h)</td>
<td>Yes</td>
<td>Yes</td>
<td>-966.2</td>
</tr>
<tr>
<td>D3</td>
<td>Free(^a)</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-972.9</td>
</tr>
<tr>
<td>D4</td>
<td>Free(^a)</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-917.9</td>
</tr>
<tr>
<td>E1</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^h)</td>
<td>Yes</td>
<td>Yes</td>
<td>-1,333.4</td>
</tr>
<tr>
<td>F1</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^h)</td>
<td>Yes</td>
<td>Yes</td>
<td>-709.8</td>
</tr>
<tr>
<td>F2</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Free(^h)</td>
<td>Yes</td>
<td>Yes</td>
<td>-721.7</td>
</tr>
<tr>
<td>F3</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Yes</td>
<td>Yes</td>
<td>-723.9</td>
</tr>
<tr>
<td>G1</td>
<td>Free(^a)</td>
<td>Free(^c)</td>
<td>Free</td>
<td>Free</td>
<td>Free(^d)</td>
<td>Free</td>
<td>Free(^h)</td>
<td>Yes</td>
<td>Yes</td>
<td>-1,032.7</td>
</tr>
<tr>
<td>G2</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>No</td>
<td>--</td>
<td>-1,026.3</td>
</tr>
<tr>
<td>G3</td>
<td>Fixed</td>
<td>Free(^{12})</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>No</td>
<td>--</td>
<td>-1,026.3</td>
</tr>
</tbody>
</table>

Table 7.7: Results from estimation using simulated data (Codes for subscript letters in endnote #3)

138
Runs with dataset A investigate the normalisation conditions associated with the logsum and salience parameters. Run A1 finds that an ‘un-normalised model’ (as described in section 4.8) is, as expected, unidentifiable. Runs A2 and A4 indicate that with suitable normalisation the ‘target’ parameter set can in fact be identified using only ‘portfolio-level’ choice observations. All ASC values in the ‘target’ parameter set are zeroes, and all are found to have absolute values smaller than 1.0 in both runs. In each of the runs the ordering of the five salience parameters (which had target values from 0.5 to 1.5 in 0.25 increments) was correctly identified.

Runs A2 and A4 ought in theory to yield identical results as they are simply different ways of accommodating the required normalisation of the \( y_{d}^{\text{travel}} \) term that arises from the multiplicative specification:

\[
\gamma_{i} \ast \frac{1}{\lambda_{\text{travel}}} \ln \sum_{m_{d} \in \mu_{d}} e^{\psi_{m_{d} i}^{\text{travel}} \lambda_{\text{travel}}}
\]

This is confirmed as they were found to yield identical parameter sets and goodness-of-fit. Run A3 is identical to A2 with the only difference being normalisation to a different arbitrarily-chosen value. Run A3 also yields identical goodness-of-fit, as expected, and the parameter set differs from that of A2 (and that of A4 as well) by only a constant \( k \) as described in Section 4.8. Run A5 is similarly analogous to Run A4, but the results are not identical: goodness-of-fit is worse and the reported parameter set is different by more than a constant \( k \) from those of Runs A2 and A4.

The parameter set from Run A4 was then factored by \( k = 5 \) to verify that the normalisation condition had been derived correctly. Likelihood remained the same; Run A6 was then run using this parameter set as the starting values. The reported parameter set of A6 is the same as these starting values – with goodness-of-fit identical to A2/A3/A4 (and better than A5).

We would generally expect likelihood to be worse in a restricted case than one with the minimum necessary normalisation. Run A7 is such a restricted case, with both one salience parameter and the logsum term fixed [at their target values], and reports a likelihood value which is marginally worse than A2/3/4/5.

This pattern of results with dataset A (particularly results from Runs A4/5/6) shows that the estimation algorithm finds the presence of multiple local maxima on the likelihood surfaces, which in the case of [at least] Run A5 prevents the identification of the [apparently-] global optimum parameter set. This is unsurprising, given the non-linear utility specifications, in particular the
multiplicative relationship between the salience parameters and all other mode-choice-level parameters.

Runs B1 and B3 report success in identifying the salience \( \gamma \) parameters, in the presence of a fully-specified model form and a restricted model form in which all other parameters are fixed at their ‘target’ values, respectively.\(^4\) C1 and C3 find the same for the logsum parameter which operates across the multiple travel methods enabled by any given portfolio, as do D1 and D3 for alternative-specific constants for each of the mobility resources. One of the three portfolio-level ASCs in D1 is reported to be not statistically-distinguishable from zero, though this was thought to be idiosyncratic rather than systematic as the utility functions are linear with respect to these ASCs. To confirm this intuition, a second dataset having the same characteristics as dataset D but different randomly-drawn error values was estimated. Run D4 is identical to D1 in its specification, but uses this second set of simulated choices; all three portfolio-level ASCs values were successfully recovered in D4, though the ASC for car ownership remains not statistically significant.

Run E1 reports that ‘target values’ of zero can be recovered for all parameters simultaneously; all parameters but two are found to be statistically different than zero, and those two calculated to be significant have positive signs whilst their target values were negative.

Runs F1, F2 and F3 investigate the possibility of recovering marginal travel time parameters, both for means of travel requiring specific mobility resources and those methods of transport common to all portfolios. The data-generation process for dataset F employed different travel time [in minutes] parameter values for each method of travel, ranging in increments of 0.05 from -0.05 to -0.175. Run F1 successfully recovers parameters for those travel modes (driving a personal car and riding a bicycle) requiring mobility resources and a single ‘shared’ travel time parameter for all other means of travel, in the presence of a fully-specified model form. Run F3 does the same when all other parameters are fixed at their target values. All travel time parameters in Run F2 are specified to be alternative-specific, including those from travel modes common to all portfolios. Parameter values are successfully recovered for the ‘mobility-resource-requiring’ methods of travel (cycling and driving a personal car), and the parameter set is reported to be unique (i.e. model is reported to be identifiable), but the values obtained for ‘common’ means of transport do not track the target values (See Table 7.8) In particular, the parameter for travel time whilst travelling by a taxi or minicab differs substantially from its target value.
Runs G1 and G2 investigate the identifiability of ASC parameters at the mode-choice-level. G1 shows that a full parameter set can be recovered in the presence of mode-choice-level ASCs for all modes of travel (including those modes of travel common to all portfolios) in the generation of the estimation dataset; the ‘non-common’ mode-choice-level ASCs are simply shifted away from their target value. G2 shows that mode-choice-level ASCs for all ‘common’ modes of travel cannot be identified simultaneously, even if all other parameters are fixed at their target values; in this case the mode-choice-level ASC for car passenger travel is unidentifiable. Figure 7.1 shows the response of the likelihood function to changing this parameter; it can be seen that the likelihood surface is not globally concave with respect to it.

The surface in Figure 7.1 can be characterised as having three distinct regions. When this parameter is arbitrarily large and positive (labelled ‘E’), it corresponds to car passenger travel is seen by all people to be the optimal method of travel for all journeys, and the $V_{d}^{l_{travel}}$ terms for all portfolios take the same large positive value – thus serving as a sort of portfolio-level ASC common to all portfolios. The likelihood surface is not responsive to changes in this parameter in this region because of this effect.

As we move left along the surface to the sections labelled ‘D’, ‘C’, and ‘B’, it becomes responsive to changes in the value of the parameter. In this region car passenger travel is seen by some people to be the optimal method of travel for some journeys.

Finally, in the third region of interest, labelled ‘A’, the likelihood surface is again unresponsive to changes in the parameter value. This corresponds to ranges of this parameter value where car passenger travel is being seen by all people as a very poorly-performing option for all journeys. In other words, this corresponds to the case where car passenger travel is in essence removed from the system. In this particular case, due to the ‘portfolio’ specification, there is no distinct maximum of the likelihood surface, but rather a ‘plateau’, leading to the parameter being unidentifiable. Whilst

<table>
<thead>
<tr>
<th>Parameter for travel time in minutes, driving a personal car</th>
<th>Target value</th>
<th>Obtained value: (std. error in parentheses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter for travel time in minutes, cycling</td>
<td>-0.1</td>
<td>-0.099 (0.02)</td>
</tr>
<tr>
<td>Parameter for travel time in minutes, car passenger travel</td>
<td>-0.15</td>
<td>-0.130 (0.02)</td>
</tr>
<tr>
<td>Parameter for travel time in minutes, taxi/minicab travel</td>
<td>-0.2</td>
<td>-0.653 (0.40)</td>
</tr>
<tr>
<td>Parameter for travel time in minutes, public transport</td>
<td>-0.25</td>
<td>-0.22 (0.13)</td>
</tr>
<tr>
<td>Parameter for travel time in minutes, walking</td>
<td>-0.3</td>
<td>-0.21 (0.05)</td>
</tr>
</tbody>
</table>

Table 7.8: Comparison of target and obtained parameter values for run F2
we can infer some information about the value of this parameter – that large positive values are implausible – we cannot identify a unique optimum value. We note that the two effects governing the shape of the likelihood function may result in asymmetric responses of the likelihood function. This can be seen in Figure 7.1 as the transition leading from the ‘lower’ plateau of the likelihood values (this transition is labelled ‘B’) is rather more abrupt (i.e. has a second derivative of a higher absolute value) than the transition to the ‘upper’ plateau (labelled ‘D’.)

![Graph showing likelihood function for car passenger travel](image)

**Figure 7.1: Response of likelihood function to varying the mode-choice-level alternative-specific constant for car passenger travel, using dataset G (Run G2)**

Figure 7.2 is the same as Figure 7.1, but the parameter investigated is the mode-choice-level ASC for taxi/minicab travel. Taxi/minicab travel is also common to all portfolios, hence the ‘upper’ and ‘lower’ plateaus occur as with the same parameter for car passenger travel. The difference between the two likelihood surfaces is that a unique optimum can be readily seen for the taxi/minicab parameter, hence it is identifiable. The presence of a unique optimum on this surface, and the lack of one on the surface for car passenger travel, supports the assertion that such parameters are in principle identifiable, though may be not be found to be so in practice due to the correlation structure of the estimation dataset employed.

It is also noteworthy that in both Figures 7.1 and 7.2 the ‘lower’ plateaus are at better likelihood values than the ‘upper’ plateaus. The ‘lower’ plateaus in both of these graphics represent a situation

142
where one of the ‘common’ modes is arbitrarily-unattractive, and essentially removed from the analysis. The $V_{d}^{i,\text{travel}}$ terms for all non-null portfolios (those enabling at least one ‘non-shared’ mode of travel) continue to be differentiated (in utility terms) from each other and from the $V_{d}^{i,\text{travel}}$ term of the null portfolio (which itself does not become ALN due to the presence of other ‘common’ modes with non-ALN parameters.)

The ‘upper’ plateaus represent the opposite situation: where one of the ‘shared’ modes is arbitrarily-attractive and therefore dominates the $V_{d}^{i,\text{travel}}$ terms of all portfolios, both the null portfolio and all non-null ones. All $V_{d}^{i,\text{travel}}$ terms thus become undifferentiated from each other. It is rather unsurprising that empirically this appears to yield worse goodness-of-fit than the case corresponding to the ‘lower’ plateaus.

![Graph showing response of likelihood function to varying the mode-choice-level alternative-specific constant for taxi/minicab travel, using dataset G (Run G2)](image)

Figure 7.2: Response of likelihood function to varying the mode-choice-level alternative-specific constant for taxi/minicab travel, using dataset G (Run G2)

Figure 7.3 shows the response of the likelihood function in reaction to changes in the mode-choice-level ASC for driving a car, which appears in some [but not all] portfolios. The ‘lower’ plateau effect occurs – this corresponds to all people viewing driving a car as highly-undesirable for all journeys, though the logsum calculations for the $V_{d}^{i,\text{travel}}$ term of all portfolios which include car ownership (including the one portfolio which contains car ownership and no other mobility resources) do not
become ALN in response to this ALN parameter. This arises from the presence of the ‘common’ modes of travel (which in this case are seen as more-desirable forms of travel than driving) within the logsum calculation.

There is no ‘upper’ plateau, however, in Figure 7.3. As this parameter becomes arbitrarily large and positive it corresponds to all people seeing car driving as far superior to the other modes of travel. The $V_d^{travel}$ terms for all portfolios that include car ownership are arbitrarily larger than the same terms for portfolios not including car ownership, due to the nature of the logsum calculation. This leads to large differences in utility between these two classes of portfolios (those that include car ownership and those that do not) and hence large impacts on the likelihood function.

Figure 7.3: Response of likelihood function to varying the mode-choice-level alternative-specific constant for driving a personal car, using dataset G (Run G2)

Figure 7.4 shows that the response of the likelihood function with respect to changes in the portfolio-level ASC for owning a personal car, a term that occurs outside of the logsum operators. The concave-definite shape of the likelihood surface is consistent with what would be expected of a parameter which appears linearly in a specification. Of note is the asymmetry in the response on either side of the optimum. This arises as ever-more-negative values for this parameter result in ever-poorer predicted goodness-of-fit for the majority (55%) of people in dataset G who were simulated to own a car, whilst ever-larger positive values likewise result in ever-poorer goodness-of-
fit for the minority of people who were simulated to not own a car. The response on the left-hand-side of the optimum is therefore sharper than that on the right-hand-side.

![Graph showing the response of likelihood function to varying the portfolio-choice-level alternative-specific constant for owning a personal car, using dataset G (Run G2).](image)

**Figure 7.4:** Response of likelihood function to varying the portfolio-choice-level alternative-specific constant for owning a personal car, using dataset G (Run G2)

To complete the discussion of the results from the simulated datasets, we note that Run G3 shows the expected finding that all mode-choice-level ASCs are unidentifiable if not properly normalised by fixing one of them at some arbitrarily-chosen value.

In summary, the evidence from analysis of simulated data indicates the following:

1) Support for the normalisation conditions outlined in section 4.8, as all improperly-normalised runs fail.

2) The E-NTS dataset appeared to have statistical properties suitable for estimation of the StraP model form, as indicated by the successful estimation of runs which were properly-normalised.

3) The presence of local optima on the likelihood surface, arising from the non-linear and multiplicative nature of the utility functions. The evidence for this are the [relatively small] differences in goodness-of-fit and parameter values between parameter sets estimated from mathematically-identical means of normalisation, which vary with the choice of starting values for an estimation.
4) Success, where this is expected, in identifying parameters of the correct sign and order of magnitude as the ‘target’ values from the simulated dataset.

5) Though travel-mode-specific parameters are in principle identifiable with the type of data at hand, it was found that consistent estimates of some parameters for modes of travel common to all portfolios could not be obtained. This supports the decision to, in the empirical analysis, constrain all alternative-specific parameters to be equal for all ‘common’ modes of travel.

Whilst it was recognised that the possibilities for performing statistical analysis of the results from the simulated datasets were not exhausted, and that the possibility existed to prepare and analyse further synthetic datasets with different statistical properties, it was decided that the principal objectives – support for proceeding to empirical analysis with the proposed specification and empirically-collected E-NTS dataset – had been met and that analysis presented after this point would employ exclusively empirical datasets.

7.4 Results from analysis of the ‘portfolio choice model’ using the E-NTS dataset

Results from models using only the E-NTS data are shown in Table 7.9.

Runs 1 and 2 show that goodness-of-fit is quite substantially improved adding the ‘holding cost’ parameter to an ASC-only specification. Interestingly, the ASC for owning a car becomes very large, and in fact very nearly counterbalances (+9.38 v. -9.73) the disutility arising from the expense of owning a car. This is not very surprising, as there is great heterogeneity in the ownership costs of a car, whilst this was treated as a single point value (£4,000/year) in this analysis.

Run 3 is arrived at by adding into the specification information on people’s journeys. All parameters are signed as expected, and with one exception all are statistically-significant. This run provides much better goodness-of-fit than Run 2, which contained no information on people’s journeys. The travel time parameters are ordered, from smallest [least negative] to largest: Ride a bicycle, drive a car, shared modes. This ordering holds constant for Runs 4 and 5 as well. The mode-choice-level ASCs, however, are ordered precisely in reverse in Run 3, though the ordering does not remain the same for Runs 4 and 5. Taken as a whole, the results of Run 3 show that the central hypothesis of this research – that a person’s choice of ‘portfolio’ of mobility resources can be considered as a function of the travel value its constituent elements provide – appears to be supported.

Run 4 is an extension of Run 3 in which the logsum parameter is allowed to be estimated freely; doing so yields a further large increase in goodness-of-fit. Run 5 then releases the ‘salience’
parameters for different journey purposes, all of which are significant with one exception. The salience terms are ordered, from smallest to largest: Escort, social, leisure, shopping/personal-business/other, work/education.

Run 6 is a restricted version of Run 5, where a single generic travel time parameter is estimated rather than mode-specific parameters. Interestingly, this change results in only a small drop in goodness-of-fit relative to Run 5.

Finally, Run 7 has the same specification\(^7\) as Run 5, with the only difference being that the travel-derived utility is divided by the number of journeys each person performed. Thus travel-dependent utility \(V^{t,travel}_d\) is scaled identically for all people in Run 7, whilst in Run 5 it is scaled by the number of journeys a person was observed to perform during their NTS diary week. Run 5 yields a much better goodness-of-fit (-1.121 v. -1.162); therefore these data seem to suggest that the specification of Run 5 is a better approximation of people’s choice-making processes.

Run 8 is identical to the specification of Run 5, with different (but equivalent) normalisation: one of the salience parameters is fixed rather than the \(x^{travel}\) logsum term. As expected, the goodness-of-fit is the same for these two runs, and the value of the salience parameters, the logsum parameter, and the mode-choice-level parameters differ only by an identical constant. It is noteworthy that the estimated significance levels of these parameters do vary [marginally] between these two model runs.
<table>
<thead>
<tr>
<th>Run #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-1,364.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-1,302</td>
<td>-1,236.3</td>
<td>-1,162.3</td>
<td>-1,134.7</td>
<td>-1,121.3</td>
<td>-1,123.1</td>
<td>-1,162.2</td>
<td>-1,121.3</td>
</tr>
<tr>
<td>$\rho^2$ / adjusted $\rho^2$</td>
<td>0.05 / 0.04</td>
<td>0.09 / 0.09</td>
<td>0.15 / 0.14</td>
<td>0.17 / 0.16</td>
<td>0.178 / 0.167</td>
<td>0.177 / 0.167</td>
<td>0.15 / 0.14</td>
<td>0.177 / 0.167</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>99.97</td>
<td>0.34</td>
<td>0.35</td>
<td>0.37</td>
<td>0.38</td>
<td>0.38</td>
<td>0.11</td>
<td>0.38</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a bicycle</td>
<td>-0.75</td>
<td>-0.140 (0.21)</td>
<td>-0.754</td>
<td>-0.766</td>
<td>-0.899</td>
<td>-0.852</td>
<td>-1.18</td>
<td>-0.899</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a car</td>
<td>-0.35</td>
<td>9.38</td>
<td>8.12</td>
<td>8.82</td>
<td>9.44</td>
<td>9.42</td>
<td>11.1</td>
<td>9.44</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a public transport season ticket</td>
<td>-0.30</td>
<td>2.25</td>
<td>1.95</td>
<td>1.50</td>
<td>1.50</td>
<td>1.51</td>
<td>0.947</td>
<td>1.50</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Ride a bicycle</td>
<td>--</td>
<td>--</td>
<td>-2.14</td>
<td>-0.161</td>
<td>-0.481 (0.12)</td>
<td>-0.214</td>
<td>-2.07</td>
<td>-0.161 (0.15)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car</td>
<td>--</td>
<td>--</td>
<td>-0.973</td>
<td>0.112</td>
<td>0.431 (0.09)</td>
<td>0.570</td>
<td>-2.04</td>
<td>0.145</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Shared modes</td>
<td>--</td>
<td>--</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
</tr>
<tr>
<td>Fixed holding costs in GBP per month</td>
<td>--</td>
<td>-0.0292</td>
<td>-0.0313</td>
<td>-0.0338</td>
<td>-0.0355</td>
<td>-0.0355</td>
<td>-0.0376</td>
<td>-0.0355</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>--</td>
<td>--</td>
<td>-0.0396 (0.27)</td>
<td>-0.182</td>
<td>-0.918</td>
<td>-0.972</td>
<td>-8.42</td>
<td>-0.308</td>
</tr>
<tr>
<td>Travel time in minutes (Generic)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>-0.0188</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>-0.021</td>
<td>-0.00458</td>
<td>-0.0184</td>
<td>--</td>
<td>-0.0904 -0.00617</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>-0.014</td>
<td>-0.00323</td>
<td>-0.0129</td>
<td>--</td>
<td>-0.0621 -0.00431</td>
</tr>
<tr>
<td>Travel time in minutes (Shared modes)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>-0.024</td>
<td>-0.00492</td>
<td>-0.0204</td>
<td>--</td>
<td>-0.0980 -0.00683</td>
</tr>
<tr>
<td>Salience parameter ($\gamma$) for ‘escort’ journey purpose</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.0847 (0.27)</td>
<td>0.0716</td>
<td>1.40</td>
<td>0.253 (0.28)</td>
</tr>
<tr>
<td>$\gamma$(leisure)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.194</td>
<td>0.195</td>
<td>0.749</td>
<td>0.577</td>
</tr>
<tr>
<td>$\gamma$(shopping, personal business, and other)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.232</td>
<td>0.213</td>
<td>0.501</td>
<td>0.691</td>
</tr>
<tr>
<td>$\gamma$(social)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.134</td>
<td>0.127</td>
<td>0.678</td>
<td>0.398</td>
</tr>
<tr>
<td>$\gamma$(work and education)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.335</td>
<td>0.315</td>
<td>1*</td>
<td>1*</td>
</tr>
<tr>
<td>Logsum term ($\lambda_{\text{travel}}$)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>4.21</td>
<td>1*</td>
<td>1*</td>
<td>0.186</td>
<td>2.98</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed. * Value is fixed

Table 7.9: Results from parameter estimation of ‘portfolio’ choice using only the E-NTS dataset
7.5 Results from analysis of the ‘portfolio choice model’ using the AVATAR survey dataset

Results from StraP models using only the AVATAR survey data are shown in Table 7.10. The patterns of results are broadly consistent with those from estimation using the E-NTS dataset, with several notable differences.

Before discussing the results, we note that several parameters common to the two estimations cannot be directly compared to draw substantive inferences. This applied to car travel, which was treated differently in the two datasets. The marginal costs of car driving (petrol & parking as defined in this research) were presented explicitly to respondents in the AVATAR survey and are hence included in the specification. These costs are not included in the specification based on the E-NTS, due to a lack of information in the revealed-preference E-NTS dataset. Whilst it would be possible in principle to infer parking costs for unchosen car itineraries on the basis of observed car itineraries, this would be likely to strongly depend on the location of people’s activity locations, which are not known due to the requirements to respect the privacy of NTS respondents.

It was thus decided to specify two separate sets of mode-choice-level parameters for car driving in the two datasets. The parameters from the E-NTS dataset can be interpreted as capturing (in the case of the mode-choice-level parameter for car driving journey duration) petrol costs and (in the case of the mode-choice-level car driving ASC) parking costs. The parameters from the AVATAR survey can be interpreted as not capturing these effects as they are captured by the inclusion of a separate parameter for mode-choice-level journey expense.

The explicit treatment of car driving costs in estimation using the AVATAR dataset leads to an ‘incorrectly’ signed parameter for car driving duration; it is [small though] positively-signed for all runs with the AVATAR dataset. By treating this parameter as distinct from the corresponding parameter for the E-NTS dataset in joint analysis (see Section 8.1), it is possible to disregard this parameter in forecasting applications.⁸

There was a large increase in goodness-of-fit between Runs 1 and 2 when using the E-NTS dataset; this does not occur when the estimation is based on the AVATAR survey dataset. The goodness-of-fit of Run 3, which is the first to include any information on the accessibility to activities in people’s perceived activity sets, is a substantial improvement, however.

The inclusion of the logsum parameter (in Run 4) provides a further large improvement in likelihood. The logsum term is calculated to be much larger here (a value of 22) than when using the E-NTS dataset (a value of 4), showing that the mode-choice-level systematic utility has a lower scale of error for the AVATAR dataset than the E-NTS dataset.

Run 5 is the ‘full’ specification, including the salience parameters. It is noteworthy that the ratio of the largest divided by the smallest salience parameter is rather larger for the AVATAR Run 5 than E-NTS Run 5 (4.0 v. 9.9). It is somewhat worrisome, that the ordering of the magnitude of the salience parameters is quite different between E-NTS Run 5 and AVATAR Run 5, as shown in Table 7.11.⁹
<table>
<thead>
<tr>
<th>Run #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-898.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-708.1</td>
<td>-707.5</td>
<td>-669.9</td>
<td>-646.4</td>
<td>-642.8</td>
<td>-657.3</td>
<td>-644.2</td>
<td>-642.8</td>
</tr>
<tr>
<td>$\rho^2$ / adjusted $\rho^2$</td>
<td>0.21 / 0.21</td>
<td>0.21 / 0.21</td>
<td>0.25 / 0.24</td>
<td>0.280 / 0.261</td>
<td>0.284 / 0.261</td>
<td>0.27 / 0.25</td>
<td>0.283 / 0.259</td>
<td>0.28 / 0.26</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>10.78</td>
<td>0.147</td>
<td>0.164</td>
<td>0.0062</td>
<td>0.035</td>
<td>1.33e-15</td>
<td>0.0007</td>
<td>0.139</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a bicycle</td>
<td>0.0529 (0.74)</td>
<td>0.167 (0.38)</td>
<td>-0.477 (0.46)</td>
<td>-0.311 (0.33)</td>
<td>-0.0720 (0.91)</td>
<td>0.479 (0.15)</td>
<td>-0.0989 (0.91)</td>
<td>-0.0720 (0.91)</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a car</td>
<td>-0.844</td>
<td>0.979 (0.68)</td>
<td>2.91 (0.40)</td>
<td>1.62 (0.49)</td>
<td>2.55 (0.55)</td>
<td>5.00 (0.39)</td>
<td>5.14 (0.15)</td>
<td>2.55 (0.55)</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a public transport season ticket</td>
<td>-0.637</td>
<td>0.0246 (0.98)</td>
<td>-1.90 (0.25)</td>
<td>-0.547 (0.57)</td>
<td>-1.18 (0.27)</td>
<td>-0.529 (0.73)</td>
<td>-0.842 (0.45)</td>
<td>-1.18 (0.27)</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Subscribe to a car club</td>
<td>-3.21</td>
<td>-3.18</td>
<td>-4.09</td>
<td>-4.65</td>
<td>-3.21</td>
<td>-3.12</td>
<td>-3.14</td>
<td>-3.21</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Subscribe to 'Versatility' service</td>
<td>0.136 (0.55)</td>
<td>0.191 (0.42)</td>
<td>-0.803 (0.28)</td>
<td>-0.164 (0.65)</td>
<td>-0.393 (0.52)</td>
<td>0.277 (0.83)</td>
<td>-0.486 (0.57)</td>
<td>-0.393 (0.52)</td>
</tr>
<tr>
<td>Fixed holding costs in GBP per month</td>
<td>--</td>
<td>-0.00547 (0.43)</td>
<td>-0.0156 (0.09)</td>
<td>-0.00969 (0.16)</td>
<td>-0.0133 (0.27)</td>
<td>-0.0204 (0.17)</td>
<td>-0.0198 (0.05)</td>
<td>-0.0133 (0.27)</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>--</td>
<td>--</td>
<td>-0.429 (0.26)</td>
<td>-0.221</td>
<td>-1.16</td>
<td>-0.814 (0.36)</td>
<td>-7.90 (0.18)</td>
<td>-0.303 (0.07)</td>
</tr>
<tr>
<td>Travel time in minutes (Generic)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>-0.4041</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car)</td>
<td>--</td>
<td>--</td>
<td>0.00602 (0.90)</td>
<td>0.00212 (0.72)</td>
<td>0.0143 (0.84)</td>
<td>--</td>
<td>0.484 (0.39)</td>
<td>0.00376 (0.85)</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>--</td>
<td>--</td>
<td>-0.0554 (0.13)</td>
<td>-0.0240</td>
<td>-0.111 (0.08)</td>
<td>--</td>
<td>-0.449 (0.18)</td>
<td>-0.0292 (0.05)</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car club car)</td>
<td>--</td>
<td>--</td>
<td>-0.222</td>
<td>-0.108</td>
<td>-0.256 (0.09)</td>
<td>--</td>
<td>-1.28 (0.20)</td>
<td>-0.0672</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a 'Versatility' car)</td>
<td>--</td>
<td>--</td>
<td>0.0545 (0.57)</td>
<td>0.0148 (0.14)</td>
<td>0.111 (0.23)</td>
<td>--</td>
<td>1.19 (0.31)</td>
<td>0.0292 (0.45)</td>
</tr>
<tr>
<td>Travel time in minutes (Shared modes)</td>
<td>--</td>
<td>--</td>
<td>-0.0189 (0.17)</td>
<td>-0.0155</td>
<td>-0.0178 (0.25)</td>
<td>--</td>
<td>-0.200 (0.07)</td>
<td>-0.0188</td>
</tr>
<tr>
<td>Salience parameter ($\gamma$) for ‘escort’ journey purpose</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>1.70 (0.24)</td>
<td>2.50 (0.44)</td>
<td>4.31 (0.11)</td>
<td>6.49</td>
</tr>
<tr>
<td>$\gamma$ (leisure)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.184 (0.33)</td>
<td>0.389 (0.35)</td>
<td>0.259 (0.54)</td>
<td>0.701</td>
</tr>
<tr>
<td>$\gamma$ (shopping, personal business, and other)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.317 (0.22)</td>
<td>0.327 (0.54)</td>
<td>1*</td>
<td>1.21</td>
</tr>
<tr>
<td>$\gamma$ (social)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.172 (0.41)</td>
<td>-0.0619 (0.86)</td>
<td>0.522 (0.70)</td>
<td>0.654</td>
</tr>
<tr>
<td>$\gamma$ (work and education)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.262 (0.17)</td>
<td>0.379</td>
<td>1.49 (0.05)</td>
<td>1*</td>
</tr>
<tr>
<td>Logsum term ($\lambda$)</td>
<td>--</td>
<td>--</td>
<td>21.9</td>
<td>1*</td>
<td>1*</td>
<td>0.208 (0.16)</td>
<td>3.81</td>
<td></td>
</tr>
</tbody>
</table>

NB: Values in parentheses are $p$-values; values smaller than 0.05 are suppressed. *Value is fixed

Table 7.10: Results from parameter estimation of 'portfolio' choice using only the AVATAR survey dataset
Table 7.11: Comparison of the ordering of salience parameters, Run 5, when using the E-NTS and AVATAR datasets

<table>
<thead>
<tr>
<th></th>
<th>E-NTS dataset</th>
<th></th>
<th>AVATAR dataset</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>Ordering</td>
<td>Value</td>
<td>Ordering</td>
</tr>
<tr>
<td>Salience parameter (y) for</td>
<td>0.253</td>
<td>1</td>
<td>6.49</td>
<td>5</td>
</tr>
<tr>
<td>‘escort’ journey purpose</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y(leisure)</td>
<td>0.577</td>
<td>3</td>
<td>0.701</td>
<td>2</td>
</tr>
<tr>
<td>y(shopping, personal business,</td>
<td>0.691</td>
<td>4</td>
<td>1.21</td>
<td>4</td>
</tr>
<tr>
<td>and other)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y(social)</td>
<td>0.398</td>
<td>2</td>
<td>0.654</td>
<td>1</td>
</tr>
<tr>
<td>y(work and education)</td>
<td>1.0</td>
<td>5</td>
<td>1.0</td>
<td>3</td>
</tr>
</tbody>
</table>

The results of Run 5 (or equivalently Run 8) lead to the following observations about respondents’ tastes for SDCSs:

1) There is large unexplained\(^{10}\) residual distaste (i.e. negative ‘portfolio-level’ ASCs) for subscribing to either form of SDCS (car club or ‘Versatility’), which evaluates in utility terms to roughly £250/month in the case of a car club subscription and £30/month in the case of a ‘Versatility’ subscription. In other words, respondents chose to subscribe to SDCSs less frequently (quite markedly so in the case of a car club subscription) than would be predicted using exclusively the SDCSs quantitative attributes (e.g. their costs and the level of accessibility which they enable.)

This would imply that any estimates of market potential that consider the easily-quantified attributes of SDCSs but neglect ones of a more qualitative nature (e.g. the constraints on the degree of freedom of car access vis a vis car ownership) may be optimistic. (e.g. Schuster et al. 2005)

2) There is positive unexplained preference, at the ‘mode-choice’ level, for travelling in a car club SDCS car, which is equivalent to roughly £5/journey. The opposite is true for travelling in a ‘Versatility’ SDCS car, though the effect is an order of magnitude smaller (£2/journey).

3) The disutility for time spent travelling in a car club SDCS car was found to be large relative to similar disutility for time spent travelling by other modes of travel. It is roughly twice as large as the disutility for time spent cycling, for instance, though as the next paragraph explains the travel time parameters for the SDCSs do not fully capture the dis-utility of time spent travelling by them.

The marginal utility for travel time spent in a ‘Versatility’ SDCS car was found to be positive, which on first glance appears to be ‘incorrectly-signed’. In the interests of simplicity, respondents were told that there was a fixed proportional relationship (20p/minute) between the cost of use and duration travelling with a ‘Versatility’ car. Thus, each minute of
travel in a ‘Versatility’ car does have a dis-utility, though the perfect co-linearity in the attribute data leads to the apparently incorrect sign for the travel time parameter.

These three numbered items can be summarised as follows: All parameters related to SDCSs for which we have a priori expectations have the expected signs. There appears to be large residual dis-taste for car club SDCSs (after the quantitative service attributes captured in this dataset are taken into account) which is in keeping with the low penetration at present of a service which on the basis of its easily-quantifiable attributes would appear to be more attractive to consumers. There are distinctive patterns in the ‘mode-choice-level’ parameters for SDCS use, though the co-linearity of the SDCS travel time and cost variables in the AVATAR design means that the parameters for travel time cannot be interpreted without taking the parameter for per-journey cost into account.

Run 6 with the AVATAR dataset does not provide satisfactory results; the likelihood surface is reported to be essentially flat with respect to the ASC for bicycle travel, and one of the salience parameters is ‘wrongly signed’ – the salience parameter for social travel is reported to be negative.

Run 7 also has problems with estimation, as the likelihood surface is found to be rather (though not completely) flat. As with the E-NTS dataset, Run 7 is based on a specification where the travel-derived utility is divided by the number of journeys each person performed. As is the case with the E-NTS dataset, the goodness-of-fit for this run is worse than that for Run 5, which has the same number of free parameters (though the difference is much smaller for the AVATAR dataset than the E-NTS dataset.) Thus both datasets provide a degree of empirical support for specifying travel-derived utility to scale with the number of journeys a person performs, rather than be equally-scaled for all people. In other words, the data seem to suggest that people, in the process of choosing which mobility resources to own, do not seem to ‘care’ a fixed amount about the capabilities they offer to perform their travels regardless of how frequently they travel. Rather people appear to ‘care’ about each one of their journeys in its own right, and the aggregate utility they implicitly assign to travelling seems to scale with how frequently they travel.

Finally, Run 8 shows that the results of Run 5 can be replicated via a different but identical normalisation method.
7.6 Results from analysis of the ‘mode choice model’ using the E-NTS and AVATAR datasets independently

In comparison with the ‘portfolio choice’ models discussed in Sections 7.3 through 7.5, the ‘mode choice’ models presented in this section have both well-understood statistical properties and straightforward interpretation (c.f. Ben-Akiva and Lerman 1985), and do not break new ground in a methodological sense. The results are however relevant to the broader aims of this research as they represent half of the StraP specification: people’s mode choices are specified to be made within the constraints imposed by their ‘portfolio’ of holdings of mobility resources which are made at some point(s) in advance. They are also of some interest as the knowledge base of travel patterns associated with SDCSs continues to develop.

In all model runs reported below, with the exception of Run 4, people’s choice sets of transport modes are specified to be constrained by their ‘portfolio’ holdings. If, for instance, an NTS respondent did not own a car, their choice set of modes does not include the option to ‘drive a personal car’. Thus these models are designed to identify parameters that apply only to ‘tactical’ [mode-choice-level] choicemaking. Run 4 is specified such that all modes are available to all people in the datasets\textsuperscript{13}, in order to investigate differences between this specification and those of the other runs.

7.6.1. Results from analysis of the ‘mode choice model’ using the E-NTS dataset

Estimation results using only the E-NTS dataset are shown in Table 7.12. Run 1 is an ASC-only model; Run 2 shows a large improvement in likelihood when mode-specific marginal travel time parameters, and a generic fare cost parameter, are introduced. All parameters for which \textit{a priori} expectations correspond to a positive or negative effect are ‘correctly’ signed and significant.

Run 3 shows the worsening in goodness-of-fit when a single generic travel time parameter is specified rather than the alternative-specific parameters of Run 2.

Run 4 shows results without availability conditions in the specification. Provided that one is willing to assume that the inclusion of the availability conditions leads to a quantitative model that is closer to people’s underlying choicemaking processes, the differences between these results and those of Run 2 can be interpreted as the bias introduced by this sort of mis-specification.
Run # | 1 | 2 | 3 | 4
--- | --- | --- | --- | ---
Null log-likelihood | -26,380.5 | -37,778.4 |
Log-likelihood | -20,327.5 | -17,758.5 | -17,949.7 | -25,779.2 |
$\rho^2$ / adjusted $\rho^2$ | 0.23 / 0.23 | 0.33 / 0.33 | 0.32 / 0.32 | 0.32 / 0.32 |
Smallest singular value in Hessian matrix | 4.64 | 0.30 | 0.34 | 1.37 |
ASC Drive a car | 0* | 0* | 0* | 0* |
ASC Ride a bicycle | -3.41 | -3.36 | -3.23 | -3.07 |
ASC Drive a club car | -- | -- | -- | -- |
ASC Drive a ‘Versatility’ car | -- | -- | -- | -- |
ASC Car passenger travel | -1.92 | -1.97 | -1.87 | -0.761 |
ASC Taxi/Minicab travel | -4.61 | -4.05 | -4.11 | -2.91 |
ASC Public transport | -2.31 | -1.95 | -1.43 | -2.60 |
ASC Walk | -1.46 | -0.281 (0.17) | -0.242 | 1.26 |

Fare costs in GBP per journey | -- | -0.140 | -0.0767 | -0.0894 |
Travel time in minutes (Generic) | -- | -- | -0.0282 | -- |
Travel time in minutes (Drive a car) | -- | -0.0396 | -- | 0.00170 (0.13) |
Travel time in minutes (Ride a bicycle) | -- | -0.0347 | -- | -0.00611 |
Travel time in minutes (Drive a car club car) | -- | -- | -- | -- |
Travel time in minutes (Drive a ‘Versatility’ car) | -- | -- | -- | -- |
Travel time in minutes (Car passenger travel) | -- | -0.0352 | -- | 0.00585 |
Travel time in minutes (Taxi/Minicab travel) | -- | -0.0279 | -- | 0.00456 (0.30) |
Travel time in minutes (Public transport) | -- | -0.0224 | -- | 0.0111 |
Travel time in minutes (Walk) | -- | -0.0341 | -- | -0.0296 |

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed. * Value is fixed

Table 7.12: Results from parameter estimation of transport mode choice using only the E-NTS survey dataset

7.6.2. Results from analysis of the ‘mode choice model’ using the AVATAR survey dataset

We note here several patterns which emerge from the mode choice model runs from the AVATAR survey dataset. First, the improvement in fit of the fully-specified model (Run 2) versus the ASC-only model (Run 1) is large.

All parameters obtained from Run 2 are prima facie reasonable and in keeping with expectations.

The results from Run 3 show the drop in goodness-of-fit which arises from constraining all travel time parameters to the same value. It is noteworthy that the parameters for travel time and cost both have larger absolute values in AVATAR Run 3 than the corresponding E-NTS Run 3, though the respective ratios are close in magnitude (£22/hour in the E-NTS dataset v. £23/hour). This suggests those peoples’ underlying data generation processes (i.e. choicemaking processes) inferred from these two runs are similar in structure, but that the errors are not scaled identically (and that there is a larger scale of error in the E-NTS dataset). Interestingly, this relationship does not hold for AVATAR and E-NTS ‘portfolio choice’ Runs 6 (see Tables 7.9 and 7.10), where the analogous ‘value-of-time’ ratios vary more widely.
Run 4 shows the bias which would be introduced into the parameter set of Run 2 by neglecting the separate ‘strategic’ and ‘tactical’ level of choicemaking which are proposed in this research.

<table>
<thead>
<tr>
<th>Run #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-2,015.5</td>
<td>-2,801.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-1,604.5</td>
<td>-1,356.2</td>
<td>-1,415.0</td>
<td>-2,215.0</td>
</tr>
<tr>
<td>$\rho^2$ / adjusted $\rho^2$</td>
<td>0.20 / 0.20</td>
<td>0.33 / 0.32</td>
<td>0.30 / 0.29</td>
<td>0.21 / 0.20</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>5.43</td>
<td>0.41</td>
<td>0.56</td>
<td>0.44</td>
</tr>
<tr>
<td>ASC Drive a car</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
</tr>
<tr>
<td>ASC Drive a bicycle</td>
<td>-0.229 (0.20)</td>
<td>1.02</td>
<td>0.246 (0.20)</td>
<td>1.41</td>
</tr>
<tr>
<td>ASC Drive a car club car</td>
<td>-1.66</td>
<td>1.06 (0.17)</td>
<td>0.667 (0.29)</td>
<td>-0.828 (0.40)</td>
</tr>
<tr>
<td>ASC Drive a ‘Versatility’ car</td>
<td>-0.874</td>
<td>-0.216 (0.57)</td>
<td>-0.0427 (0.87)</td>
<td>-1.20</td>
</tr>
<tr>
<td>ASC Car passenger travel</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>ASC Taxi/Minicab travel</td>
<td>-2.58</td>
<td>-0.775 (0.11)</td>
<td>-1.11</td>
<td>0.509 (0.26)</td>
</tr>
<tr>
<td>ASC Public transport</td>
<td>-1.29</td>
<td>-0.339 (0.22)</td>
<td>0.294 (0.17)</td>
<td>0.813</td>
</tr>
<tr>
<td>ASC Walk</td>
<td>-1.95</td>
<td>0.533 (0.08)</td>
<td>-0.239 (0.30)</td>
<td>2.25</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>--</td>
<td>-0.259</td>
<td>-0.195</td>
<td>-0.277</td>
</tr>
<tr>
<td>Travel time in minutes (Generic)</td>
<td>--</td>
<td>--</td>
<td>-0.0779</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car)</td>
<td>--</td>
<td>-0.0516</td>
<td>--</td>
<td>-0.00620 (0.40)</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>--</td>
<td>-0.100</td>
<td>--</td>
<td>-0.0522</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car club car)</td>
<td>--</td>
<td>-0.0453</td>
<td>--</td>
<td>0.0105 (0.74)</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a ‘Versatility’ car)</td>
<td>--</td>
<td>-0.0419</td>
<td>--</td>
<td>0.0510</td>
</tr>
<tr>
<td>Travel time in minutes (Car passenger travel)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Taxi/Minicab travel)</td>
<td>--</td>
<td>-0.0556</td>
<td>--</td>
<td>0.00595 (0.81)</td>
</tr>
<tr>
<td>Travel time in minutes (Public transport)</td>
<td>--</td>
<td>-0.0430</td>
<td>--</td>
<td>-0.00273 (0.57)</td>
</tr>
<tr>
<td>Travel time in minutes (Walk)</td>
<td>--</td>
<td>-0.0969</td>
<td>--</td>
<td>-0.0829</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed. *Value is fixed

Table 7.13: Results from parameter estimation of transport mode choice using only the AVATAR survey dataset

7.7 Summary

This Chapter presented a collection of results from analysing the E-NTS and AVATAR datasets independently with the StraP model form. Estimated parameter sets for the ‘portfolio’ specification that were inferred from the E-NTS dataset were found to be reasonable. Results of ‘portfolio-level’ model runs from the AVATAR survey were then shown to also be broadly plausible. Finally, standard linear-in-the-parameters models of mode choice were found, for both the E-NTS and AVATAR datasets, to also produce reasonable results.

Chapter eight presents results from linking together the two datasets to form the StraP model specifications used for developing forecasts, together with the substantive forecasting results.
Notes

1 Eight of the 35 ‘unused’ resources were purchased by the earliest eight respondents (a rate of 1:1), whilst the remaining 27 were purchased by the subsequent 64 respondents (a ratio smaller than 1:2).

2 Travel time parameters were specified to be smaller for ‘non-shared’ modes (car driving and cycling) of travel than ‘common’ modes of travel. Each of the ‘non-shared’ modes involve some degree of expense, and hence the travel time parameters are most readily identifiable by specifying there to be a distinct benefit (i.e. less negative travel time parameters) for these modes of travel than for the ‘common’ modes.

3 Codes for subscript letters in Table 7.7:
   a. Portfolio-level ASCs are free for each mobility resource
   b. Portfolio-level ASCs are free for each method of travel enabled by a given portfolio
   c. Mode-choice-level ASCs are free for each method of travel enabled by a given portfolio, but those ASCs for methods of travel common to all portfolios are fixed at zero
   d1. Mode-choice-level ASC are free for each method of travel enabled by a given portfolio, with one such ASC fixed to zero
   d2. Mode-choice-level ASC are free for each method of travel enabled by a given portfolio
   e. All mode-choice-level ASCs are free
   f. Travel time parameters are free and alternative-specific, for all travel methods
   g. Travel time parameters are free and alternative-specific for all travel methods not common to all portfolios. A single common travel time parameter for such travel methods is freely estimated
   h. All salience parameters are free except one, which is fixed to its target value of one
   i. All salience parameters are free

4 Runs B2, C2 and D2 are identically-specified as B1, C1 and D1 respectively but normalised differently. The results in all three cases are, as expected, identical for both methods of normalisations.

5 The logsum operator is not a maximisation operator, thus the likelihood surface will always be responsive to some degree (i.e. have a non-zero partial derivative) to a change in a parameter value, regardless of location on the likelihood surface. This response is vanishingly small in regions such as this; Appendix B presents a derivation of this effect.

6 The term ‘lower plateau’ is proposed to refer to the plateau of the likelihood surface that arises from arbitrarily large and negative values of the parameter under inspection; and vice versa for the term ‘upper plateau’.

7 Note that Run 7 is normalised in a different, but mathematically identical, manner than Run 5 is. Estimation of Run 7 with the same normalisation as Run 5 (fixing the logsum term to 1.0) failed, though it was verified that factoring the parameters such that the logsum term in Run 7 takes the value 1.0 is mathematically identical (i.e. does not change the goodness-of-fit.)

8 An explanation for this ‘incorrectly-signed’ parameter may lie in the design of the AVATAR survey, specifically the decision to truncate itinerary times at 120 minutes in duration.

9 Caution should be exercised in comparing the magnitude of the E-NTS and AVATAR salience parameters in Table 7.11 due to the large difference in the magnitude of the logsum term reported for the two datasets. A direct comparison of the size of the parameters can be made by comparing the E-NTS and AVATAR results of Run 8.

10 Unexplained by the parameters applied to measurable quantities such at costs and itinerary durations.

11 Using the parameter set from Run 5, the dis-utility for each minute of travelling in a ‘Versatility’ SDCS car is calculated as $0.111 + (0.2 \times -1.16) = -0.121$. This value is of the same order of magnitude as those of the other modes of travel. Using the formula for calculating the variance of the sum of two correlated random
variables \( \sigma_{\text{full}} = \sqrt{\sigma_{\text{travel time (mins),car club SDCE}}^2 + \sigma_{\beta_\text{journey}}^2 + 2 \times \sigma_{\text{travel time (mins),car club SDCE} \times \sigma_{\beta_\text{journey}}} \) ,

the standard error, t-value, and p-value of this quantity are thus 0.152, 0.80 and the p-value is thus 0.42.

12 Though we do not report it here, an alternate Run 7 was performed where the salience parameter for travel to/from work or education activities was fixed and the salience parameter for travel to/from shopping, personal business and ‘other’ activities was allowed to vary. The latter parameter (which is the source of the nearly-flat likelihood surface) was found to settle at zero, hence the mathematically-identical normalisation reported here in which it is fixed at 1.0.

13 The one exception to this specification is that car driving was not specified to be available to children. This only occurred in the E-NTS dataset as all respondents to the AVATAR survey were adults.

14 ‘Value-of-time’ ratios are discussed in detail in Section 8.1.4.
Chapter 8: Substantive results

Chapter seven presented a series of diagnostic results of the datasets used in this research. In Chapter eight, we present the substantive findings as they relate to the market for subscription drive-it-yourself car services [SDCs], which concludes the presentation of original research. Chapter nine summarises and concludes this thesis.

Section 8.1 describes the estimation of the parameters used in the Strategic portfolio [StraP] model. Section 8.2 presents the scenarios which were evaluated, together with the results predicted by the StraP model. Section 8.3 discusses patterns in the predicted ‘portfolio’ and ‘mode’ choices, whilst Section 8.4 investigates how the explanatory power from this analysis varies with the amount of travel that is specified to form each person’s unique perceived activity set. Section 8.5 summarises the results presented in this Chapter.

8.1 Development of StraP model used for forecasting applications (from the combined E-NTS/AVATAR survey datasets)

8.1.1. Characteristics of the parameter set arising from combined estimation

The parameters in the combined E-NTS/AVATAR estimation are:

- Those parameters from the E-NTS-only estimation (See Section 7.4)
- Those parameters from the AVATAR-only estimation that are not common with the E-NTS-only estimation. (The exceptions to this are the mode-choice-level alternative-specific constant for car driving and the marginal travel time parameter for car driving, which are estimated separately to the corresponding parameters associated with the E-NTS dataset due to different treatment of car usage costs in the two datasets.)
- A scale parameter, which we term ν, which scales all salience parameters for data sourced from the AVATAR survey (data sourced from the E-NTS dataset are unaffected)
- A second scale parameter, which we term η, which scales the systematic utility for data sourced from the E-NTS dataset (data sourced from the AVATAR survey dataset are unaffected.)

The ν scale term is included in the specification to take account of the different representations of people’s perceived activity set in the two datasets. Whilst the representation of the PAS for each person observed in the NTS dataset is taken to be their observed journeys captured via their seven-day diary, which range between one and 58 journeys, the corresponding representation of the PAS
for AVATAR survey respondents are the sets of activities that they are presented in the stated-choice exercise. The set of journeys in the SC exercise consisted of either ten or 18 journeys, depending on whether the person was employed or in education (in which case their PAS consisted of 18 journeys), or neither (in which case it consisted of 10 journeys). The $\nu$ term captures the difference in scale due to this structural difference between the two datasets.

The $\eta$ term, which allows the scale of all systematic utility terms to vary between the two datasets, is a standard element in analyses combining revealed-choice and stated-choice data, where the differences in data-generating processes may quite plausibly lead to such differences in scale.

For convenience we reproduce Equation 4.1 here as Equation 8.1. Equation 8.2 shows the general form of the ‘portfolio choice’ utility functions:

$$U_d^i = V_d^{i,\text{non-travel}} + V_d^{i,\text{travel}} + \varepsilon_d^i \quad (8.1)$$

$$U_d^i = \sum_{r=0, r \notin d}^R V_r^{i,\text{non-travel}} + \left( \sum_{j=1}^{J_i} \gamma_j \ast \frac{1}{\lambda^{\text{travel}}} \ln \sum_{m \in \mu_d} e^{(\nu_{m|j}^{\text{travel}} + \lambda^{\text{travel}})} \right) + \varepsilon_d^i \quad (8.2)$$

Equations 8.3 and 8.4 show, respectively, the general utility functional form for data records from the E-NTS and AVATAR survey datasets, incorporating the $\eta$ and $\nu$ terms:

$$U_{d,E-NTS}^i = \eta \ast \left( \sum_{r=0, r \notin d}^R V_r^{i,\text{non-travel}} + \left( \sum_{j=1}^{J_i} \gamma_j \ast \frac{1}{\lambda^{\text{travel}}} \ln \sum_{m \in \mu_d} e^{(\nu_{m|j}^{\text{travel}} + \lambda^{\text{travel}})} \right) \right) + \varepsilon_d^i \quad (8.3)$$

$$U_{d,AVATAR}^i = \sum_{r=0, r \notin d}^R V_r^{i,\text{non-travel}} + \left( \sum_{j=1}^{J_i} \nu \ast \gamma_j \ast \frac{1}{\lambda^{\text{travel}}} \ln \sum_{m \in \mu_d} e^{(\nu_{m|j}^{\text{travel}} + \lambda^{\text{travel}})} \right) + \varepsilon_d^i \quad (8.4)$$

### 8.1.2. Description of the parameter set used in the StraP 'portfolio' choice model

Table 8.1 presents the results from joint estimation of the ‘portfolio’ choice model; these results can be compared the results of Runs 5 in Tables 7.9 and 7.10 for the E-NTS and AVATAR datasets, respectively, in isolation. Column ‘A’ shows the outputs from the estimation process, whilst Column ‘B’ shows the outputs after the per-minute SDCS charges (£0.12/minute and £0.20/minute for the car club and ‘Versatility’ service respectively) are taken into account.
All parameters to be used in the forecasting application are in keeping with a priori sign expectations. The mode-choice-level ASC for driving a car club SDCS car is quite large and positive, it is equivalent to a monetary value of £14/journey.

The parameter for car club journey time is correctly-signed (negative) though relatively large in comparison to the same parameter for the other modes of travel. Its value is, for instance, roughly six times the magnitude of the next largest, the travel time parameter for cycling, and when combined with the 'cost' disutility for car club travel becomes larger and negative (its value increases from -0.454 to -0.0570).

The profile of the parameter set associated with 'Versatility' is quite different. The mode-choice-level ASC is small and negative (with a dis-utility equal to £1.46/journey), whilst the travel time parameter is positive. Its 'full' value (-0.06) is correctly-signed once the service’s per-minute monetary cost is taken into account, though this quantity is not statistically significant.

<table>
<thead>
<tr>
<th>Column ‘A’</th>
<th>Column ‘B’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-2.447.7</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-1.957.8</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>0.058</td>
</tr>
<tr>
<td>Scale term (β) for systematic utility (appears only for the E-NTS dataset)</td>
<td>1.30</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a bicycle</td>
<td>-0.343</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a car</td>
<td>6.50</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a public transport season ticket</td>
<td>1.01</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Subscribe to a car club</td>
<td>-4.82</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Subscribe to 'Versatility'</td>
<td>-0.293 (0.51)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Ride a bicycle</td>
<td>-0.431 (0.22)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car (E-NTS)</td>
<td>0.411 (0.11)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car (AVATAR)(a)</td>
<td>1.23(b) (0.14)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car club car</td>
<td>16.3</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a ‘Versatility’ car</td>
<td>-1.66 (0.07)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Shared modes</td>
<td>0</td>
</tr>
<tr>
<td>Fixed holding costs in GBP per month</td>
<td>-0.0253</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>-1.13</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car, E-NTS)</td>
<td>-0.0251</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car, AVATAR)(b)</td>
<td>0.0368 (0.27)</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>-0.0567</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car club car)</td>
<td>-0.454</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a ‘Versatility’ car)</td>
<td>0.169</td>
</tr>
<tr>
<td>Travel time in minutes (Shared modes)</td>
<td>-0.0282</td>
</tr>
<tr>
<td>Scale term (Y) for all Y terms in the AVATAR dataset only</td>
<td>1.67</td>
</tr>
<tr>
<td>Y for ‘escort’ journey purpose</td>
<td>0.101 (0.11)</td>
</tr>
<tr>
<td>Y(leisure)</td>
<td>0.127</td>
</tr>
<tr>
<td>Y(shopping, personal business, and other)</td>
<td>0.170</td>
</tr>
<tr>
<td>Y(social)</td>
<td>0.100</td>
</tr>
<tr>
<td>Y(work and education)</td>
<td>0.216</td>
</tr>
<tr>
<td>Logsum term ((J^{travel}))</td>
<td>1.0(a)</td>
</tr>
</tbody>
</table>

**NB:** Values in parentheses are p-values; values smaller than 0.05 are suppressed.

\(a\) Value is fixed

\(b\) Parameter is output from the estimation process but not used in forecasting applications. See Section 7.5.2.

Table 8.1: Results from parameter estimation of 'portfolio' choice using the combined E-NTS and AVATAR survey datasets.
The salience parameters are reported to be ordered (from largest to smallest):

1) Journeys made to provide access to/from either work or education
2) Journeys made to provide access to/from shopping, personal business and ‘other’ activities
3) Journeys made to provide access to/from leisure activities
4) Journeys made to escort another person to/from any type of activity
5) Journeys made to provide access to/from social activities

The largest of the salience parameters has a value approximately twice the magnitude of the smallest salience parameter. This pattern indicates that the relative importance that people place on accessing different sorts of activities, in their choice of which mobility resources to equip themselves with, varies substantially but within a fairly narrow range.

The \( \nu \) scale term is estimated to be greater than 1, as is the \( \eta \) term. The former implies that the explanatory power of each journey is greater in the AVATAR dataset, which is in keeping with expectations as the AVATAR survey design included fewer journeys than the average weekly diary in the E-NTS dataset. The value of the \( \eta \) term leads to the interpretation that the systematic component of utility is scaled larger (relative to the residual error component) for data records from the E-NTS dataset than those from the AVATAR survey dataset.

The cross-attribute trade-offs implied by the relative magnitudes of the various parameters are of note. A typical cross-attribute trade-off investigated in mode choice analysis is the ‘value of time,’ which is a metric obtained by dividing the parameters for travel time and cost. The calculated VoT using the alternative-specific travel time parameters are:

- Car driving: \((-0.0251 \times 60)/-1.13 = £1.33/\text{hour}\)
- Cycling: £3.01/\text{hour}
- Driving a car club SDCS car: £29.10/\text{hour}
- Driving a ‘Versatility’ car: £3.03/\text{hour}
- Modes common to all portfolios: £1.50/\text{hour}

(Public transport, walking, taxi/minicab travel, and car passenger travel)

There is a marked difference in the calculated VoT values for car club use and the others; the values for all other modes are small by comparison, though as noted in Section 8.1.4 the AVATAR survey was not designed to separately identify alternative-specific travel time and cost parameters.

In the context of a traditional mode choice model, VoT values of £1 – £3/\text{hour} would be so small as to be of ambiguous plausibility (neither obviously plausible nor obviously not so.) It is unclear
whether these values reflect different structures of people’s choice-making at the ‘strategic’ level versus at a ‘tactical’ mode choice level, though we explore this point in some detail in Section 8.1.4.

A further cross-attribute trade-off of note, that between monthly cost for holding mobility resources and per-journey cost for performing journeys. This ratio \[
\frac{\beta^E/journey}{\beta^E/month}
\]
is calculated to be 10.7 when using data from both the E-NTS and AVATAR survey datasets, as in Table 8.1.³ Using only [Run 5 of] the E-NTS dataset, it is calculated to be 5.0; the comparable figure for the AVATAR dataset estimation is 46.0. The \( \beta^E/month \) estimated from the AVATAR dataset is not statistically significant, however, meaning that the last of these ratios should be treated with caution. Indeed, this statistical insignificance is manifested in the ‘joint’ E-NTS/AVATAR ratio being much closer in magnitude to that from the E-NTS estimation.

This [joint E-NTS/AVATAR] ratio appears to be ‘correctly-signed’ but smaller than the range of plausible values. For instance, an adult bus fare and a monthly bus pass in London cost £0.80 and £42 respectively in 2004/05, when the E-NTS dataset was collected. The cross-attribute trade-off would imply that people would consider it economic to purchase a £42 monthly bus pass in order to save £0.80/journey if they foresaw making more than five bus journeys/month, an implausibly low number.

The source of the unreasonable ratio is thought to likely be due to the type of data used to estimate the \( \beta^E/month \) parameter. The monthly fixed cost of owning a car which respondents considered was specified as a constant £4,000 for all respondents; the values for the fixed costs of bicycle ownership and SDCS subscriptions were similarly fixed, and the values for public transport season tickets varied only by whether E-NTS/AVATAR respondents lived in Inner (£100/mo.) or Outer (£150/mo.) London. Empirical evidence to support this hypothesis, that limitations in the data structure are largely responsible for the implausible \( \frac{\beta^E/journey}{\beta^E/month} \) ratio, is the relatively-large portfolio-level ASC for car ownership in the E-NTS dataset (a value of 9.44; see Table 7.7), which takes a much smaller value (2.55, see Table 7.11) in the analysis of the AVATAR dataset – where all respondents were presented with an identical car ownership option.⁴ Evidence against this hypothesis is the insignificance of the \( \beta^E/month \) parameter in the AVATAR survey analysis, compared to the larger value and high level of significance in the E-NTS data analysis.

One may speculate on more-sophisticated treatments for each of these attributes, some which could in principle be performed with the extant E-NTS and AVATAR datasets, though determining
the empirical gains arising from such treatments must remain a matter for future enquiries. As with
the findings from the VoT analysis described in Section 8.1.4, the magnitude of this ratio lends
support for placing limited stock in forecasts of the StraP model outside of a narrow range of
attribute values on either side of the attribute values used in [AVATAR survey] data collection and
parameter estimation.

8.1.3. Description of the parameter set used in the StraP ‘mode’ choice model
The parameter set obtained from joint estimation of the mode choice component of the StraP model
is presented in Table 8.2.

<table>
<thead>
<tr>
<th></th>
<th>Column ‘A’</th>
<th>Column ‘B’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-24,447.2</td>
<td>--</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-17,103.8</td>
<td>--</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>0.19</td>
<td>--</td>
</tr>
<tr>
<td>Scale term (I) for systematic utility (appears only for the E-NTS dataset)</td>
<td>2.56</td>
<td>--</td>
</tr>
<tr>
<td>ASC Drive a car</td>
<td>0a</td>
<td>--</td>
</tr>
<tr>
<td>ASC Ride a bicycle</td>
<td>-1.17</td>
<td>--</td>
</tr>
<tr>
<td>ASC Drive a car club car</td>
<td>-0.593 (0.35)</td>
<td>--</td>
</tr>
<tr>
<td>ASC Drive a ‘Versatility’ car</td>
<td>-0.734</td>
<td>--</td>
</tr>
<tr>
<td>ASC Car passenger travel</td>
<td>-0.828</td>
<td>--</td>
</tr>
<tr>
<td>ASC Taxi/Minicab travel</td>
<td>-1.64</td>
<td>--</td>
</tr>
<tr>
<td>ASC Public transport</td>
<td>-0.849</td>
<td>--</td>
</tr>
<tr>
<td>ASC Walk</td>
<td>-0.537</td>
<td>--</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>-0.0617</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car, E-NTS)</td>
<td>-0.0179</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car, AVATAR)b</td>
<td>0.00382b</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>-0.0166</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car club SDCS car)</td>
<td>-0.00626 (0.70)</td>
<td>-0.0114 (0.48)</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a ‘Versatility’ car)</td>
<td>0.00172 (0.86)</td>
<td>-0.0106 (0.27)</td>
</tr>
<tr>
<td>Travel time in minutes (Car passenger travel)</td>
<td>-0.0155</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Taxi/Minicab travel)</td>
<td>-0.0121</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Public transport)</td>
<td>-0.00919</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Walk)</td>
<td>-0.0132</td>
<td>--</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed.
a Value is fixed
b Parameter is output from the estimation process but not used in forecasting applications. See Section
7.5.2.

Table 8.2: Results from parameter estimation of mode choice using the combined E-NTS and AVATAR survey datasets

All parameters are in keeping with a priori sign expectations. The parameter for ‘Versatility’ SDCS travel time is positive, though following the adjustment described in Section 7.5.2 to take account of the co-linearity between car club journey time and expense, the ‘full’ marginal effect of journey duration using the ‘Versatility’ service is, as expected, negative. These ‘full’ marginal travel time parameters for the SDCS, though ‘correctly-signed’, are however not very significant.

163
8.1.4. Investigation of ‘value-of-time’ estimates

Table 8.3 is a summary of value-of-time calculations for non-SDCS modes using each of the datasets alone and together, from parameter estimates for both the ‘portfolio’ and ‘mode’ choice models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Implied value of time (£/hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-NTS portfolio choice model, Run 4, using ‘shared’ modes travel time parameter</td>
<td>£1.62</td>
</tr>
<tr>
<td>E-NTS portfolio choice model, Run 5, using ‘shared’ modes travel time parameter</td>
<td>£1.33</td>
</tr>
<tr>
<td>E-NTS portfolio choice model, Run 6, using generic travel time parameter</td>
<td>£1.16</td>
</tr>
<tr>
<td>AVATAR portfolio choice model, Run 4, using ‘shared’ modes travel time parameter</td>
<td>£4.21</td>
</tr>
<tr>
<td>AVATAR portfolio choice model, Run 5, using ‘shared’ modes travel time parameter</td>
<td>£4.03</td>
</tr>
<tr>
<td>AVATAR portfolio choice model, Run 6, using generic travel time parameter</td>
<td>£2.96</td>
</tr>
<tr>
<td>E-NTS/AVATAR portfolio choice model, Run 4, using ‘shared’ modes travel time parameter</td>
<td>£1.74</td>
</tr>
<tr>
<td>E-NTS/AVATAR portfolio choice model, Run 5, using ‘shared’ modes travel time parameter</td>
<td>£1.49</td>
</tr>
<tr>
<td>E-NTS/AVATAR portfolio choice model, Run 5, using ‘generic’ modes travel time parameter</td>
<td>£1.85</td>
</tr>
<tr>
<td>E-NTS mode choice model, Run 2, using public transport modes travel time parameter</td>
<td>£9.60</td>
</tr>
<tr>
<td>E-NTS mode choice model, Run 3, using ‘generic’ travel time parameter</td>
<td>£22.06</td>
</tr>
<tr>
<td>AVATAR mode choice model, Run 2, using public transport travel time parameter</td>
<td>£9.96</td>
</tr>
<tr>
<td>AVATAR mode choice model, Run 3, using ‘generic’ travel time parameter</td>
<td>£23.97</td>
</tr>
<tr>
<td>E-NTS/AVATAR mode choice model, Run 2, using public transport travel time parameter</td>
<td>£8.94</td>
</tr>
<tr>
<td>E-NTS/AVATAR mode choice model, Run 3, using ‘generic’ travel time parameter</td>
<td>£4.10</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed.

*This result is based on parameter sets found in Appendix D.

b This result is based on the parameter sets shown in Tables 8.1 and 8.2

Table 8.3: Comparison of ‘values of time’ estimates for non-SDCS modes of transport

The VoTs from the mode choice models are generally found to be in keeping with expectations of the magnitude of this ratio (Mackie et al. 2003), and tend to be larger than the various estimates from the ‘portfolio’ choice models. The current guidance from the UK’s Department for Transport, for instance, advises analysts to use VoT’s which, in 2002 prices, range from a minimum of £4/hour (for people’s non-working time on non-commute journeys) through to £36/hour (for a taxi passenger during working hours) (DfT 2011).

The finding of generally-smaller VoTs for the ‘portfolio’ choice models than the mode choice models is not completely unexpected, though the differences are somewhat sharper than may reasonably have been foreseen. This finding appears to be robust across both the E-NTS [revealed-choice] and
AVATAR [stated-choice] datasets; perhaps a body of literature will develop over time which will settle whether this result is (as it plausibly may be) an artefact of the structure of the datasets, indicative of people engaging subtly different mental processes at strategic and tactical levels of choicemaking, or some combination of the two.

<table>
<thead>
<tr>
<th></th>
<th>Implied value of time (£/hour) using</th>
<th>Implied value of time (£/hour) using</th>
<th>Implied value of time (€/hour) using</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>travel time parameter for car club</td>
<td>travel time parameter for ‘Versatility’ SDCS travel</td>
<td></td>
</tr>
<tr>
<td>AVATAR portfolio choice model, Run 4</td>
<td>£34.32</td>
<td>£7.98</td>
<td></td>
</tr>
<tr>
<td>AVATAR portfolio choice model, Run 5</td>
<td>£18.24</td>
<td>£6.26</td>
<td></td>
</tr>
<tr>
<td>E-NTS/AVATAR portfolio choice model, Run 4</td>
<td>£28.55</td>
<td>£3.60</td>
<td></td>
</tr>
<tr>
<td>E-NTS/AVATAR portfolio choice model, Run 5</td>
<td>£29.10</td>
<td>£3.03</td>
<td></td>
</tr>
<tr>
<td>AVATAR mode choice model, Run 2</td>
<td>£15.49</td>
<td>£21.71</td>
<td></td>
</tr>
<tr>
<td>E-NTS/AVATAR mode choice model, Run 2</td>
<td>£11.09</td>
<td>£10.33</td>
<td></td>
</tr>
</tbody>
</table>

Table 8.4: Comparison of ‘values of time’ estimates for SDCS modes of transport

The VoT estimates shown in Table 8.4 for the SDCS methods of travel are thought to be less reliable, due to the design characteristics of the AVATAR survey (see Section 7.5.2). Recognising this uncertainty in the estimates of the relative importance of SDCS travel time and expenses to consumers, the forecasting application described in Section 8.2 provides, with caution, forecasts within a fairly narrow band of variation in SDCS service attributes from those presented to respondents in the AVATAR survey.

8.2     Forecasting application using the StraP model

This section presents the results from employing the StraP model in forecasting applications.

Section 8.2.1 presents notes on the methods employed to develop the forecasts. Section 8.2.2 lists and describes the scenarios which were analysed. Sections 8.2.3 through 8.2.10 then present and discuss the results predicted by the StraP model for each of the scenarios.

8.2.1. Procedures employed in the development of forecasts

For forecasting purposes, it is specified that the universe of potential SDCS subscribers is restricted to adults holding driving licences.6

The forecasts are developed by the following steps:

1) Introducing the SDCS option(s) into the ‘portfolio-level’ choice sets of licence-holding adults in the E-NTS sample.
2) Simulating, using a Monte Carlo technique, each person’s ‘portfolio-level’ choice of which mobility resources to hold. Aggregate shares of ‘portfolio’ holdings are estimated by summing across each person’s simulated choice. Choices are simulated, rather than merely predicted probabilistically, as step #3 is conditional on each person’s discrete ‘portfolio’ holdings. For the sake of simplicity, once drawn the vector of random draws for each person in the E-NTS dataset is held constant across each scenario reported below. This ensures that all differences between the scenarios arise from differences in the systematic utility calculations; a more statistically-rigorous (though resource-intensive) methodology would be to simulate such draws many times. For reporting purposes below, the ‘portfolio’ choice market shares are presented on the basis of ten simulated choices per person, in order to expose the small effects associated with several of the scenarios (as the number of ‘choice observations’ thus increases from 704 to 7,040).

3) Simulating, for each journey performed by people in the E-NTS dataset, the probability of it being performed by each mode of travel. The set of methods of transport that are available to each person is conditional on the constraints determined by the person’s simulated ‘portfolio-level’ choice. Aggregate modal shares are then estimated again through Monte Carlo simulation.

8.2.2. Descriptions of scenarios

Forecasts are presented for the following scenarios:

0) Neither car club nor ‘Versatility’ SDCSS available
1) Car club SDCSS widely-available in Inner London (and, implicitly, wide consumer awareness)
2) As Scenario #1, with car club SDCSS also widely-available across Outer London as well
3) As Scenario #1, with car club SDCS usage priced 5% less-expensively (£4.75/hour v. £5/hour)
4) As Scenario #1, with car club SDCS journey times reduced by 5% (as a proxy for a policy measure such as allowing SDCS cars to use bus lanes)
5) Car club and ‘Versatility’ SDCSS widely-available across both Inner and Outer London
6) As Scenario #5, but with SDCS usage priced 100x more-expensively (car club usage £500/hour v. £5/hour; ‘Versatility’ usage £20.00/minute v. £0.20/minute)
7) As Scenario #5, with an identical specification of the ‘portfolio’ choice model, but with the mode choice model specification modified by replacing the ASC for cycling usage with the corresponding parameter from the E-NTS parameter set.

Scenario #6 was included in the analysis to investigate the StraP model’s response to implausible attribute values, in order to determine whether the ‘plateau’ effect (see Chapter seven) arises from
car club usage always appearing in ‘portfolio’ options together with other methods of travel. In this case, one would expect this scenario to predict some materially-non-zero level of SDCS subscription, though essentially no SDCS usage.

Scenario #7 was performed for diagnostic purposes, in recognition of the implausibly-high level of cycling (27% mode share) chosen by AVATAR survey respondents. In order to ascertain the sensitivity of the forecasts to this data-collection artefact, the mode-choice-level ASC associated with cycling from the joint E-NTS/AVATAR estimation (see Table 8.2) is replaced with the corresponding parameter from the E-NTS-only estimation (see Table 7.13)\textsuperscript{12}.

The results of the scenarios are generally presented below to the nearest tenth of a percent of market share; this is to expose a number of small effects, not a reflection of such a high degree of confidence in the figures. Indeed, all results below should be interpreted as indicative only, in recognition of the issues discussed earlier.

### 8.2.3. Scenario analysis: Baseline scenario

The baseline scenario, item zero in the listing above, looks at the ability of the StraP model form to replicate E-NTS respondents’ holdings of mobility resources and usage of transport modes. It predicts higher levels of holdings of mobility resources than observed in the E-NTS dataset:

- **Car ownership:** 33% of Londoners observed, 41% predicted
- **Public transport season ticket ownership:** 42% observed, 47% predicted
- **Bicycle ownership:** 32% observed, 36% predicted
- **No holdings:** 23% observed, 18% predicted

The upward bias in predictions of holdings of mobility resources arises from the higher level of [stated-choice] holdings in the AVATAR survey dataset than the E-NTS dataset. The joint estimation (using observations from both datasets) of the parameter set used for forecasting leads to a predicted level of mobility-resource-holding that does not match the observed market shares in the E-NTS dataset.\textsuperscript{13} Though not reported here as it would add time and complexity to the estimation process (due to the higher dimensionality of the parameter space), it would in principle be straightforward to refine these scenario analyses by estimating separate ‘portfolio-choice’ level alternative-specific constants for the E-NTS and AVATAR portions of the estimation dataset. This would result in the predicted shares of ownership of the various mobility resources in this scenario precisely matching the observed shares in the E-NTS dataset.

A similar argument applies to the predicted shares of mode usage, though the bias introduced in E-NTS ‘market shares’ by the joint estimation with the AVATAR survey dataset is smaller:\textsuperscript{14}

- **Car driving:** 28% of Londoners’ observed journeys, 35% predicted
- **Public transport:** 25% observed, 25% predicted
- **Bicycling:** 1% observed, 3% predicted
- **Car passenger:** 17% observed, 20% predicted
S Le Vine thesis

- Walking: 28% observed, 16% predicted
- Taxi/minicab: 1% observed, 1% predicted

The proportion of walking journeys which is predicted is much lower than that observed in the NTS sample, however. This is due to the sampling methodology employed in the NTS – all observed walking journeys under a mile in length are weighted up by a factor of seven to account for the one-day observation period for these journeys. Predicting that these journeys will be performed by walking at any rate less than 100% will mean that the predicted proportion of walking journeys is biased downwards, and that the predicted proportions of all other forms of travel are biased upwards. To account for this artefact of the NTS sampling protocol, observed short-walk journeys are excluded from the presentation of predicted mode choices for this and all subsequent scenarios, leading to the following predictions of ‘residual’ modal share from the baseline scenario:

- Car driving: 34% observed, 37% predicted
- Public transport: 31% observed, 28% predicted
- Bicycling: 2% observed, 3% predicted
- Car passenger: 21% observed, 19% predicted
- Walking: 11% observed, 12% predicted
- Taxi/minicab 2% observed, 1% predicted

The summary of results from the scenarios is presented in Table 8.5.

<table>
<thead>
<tr>
<th>Scenario Number</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted holdings of mobility resources, as a percent of the E-NTS sample of Londoners</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car ownership</td>
<td>40.5%</td>
<td>39.8%</td>
<td>39.1%</td>
<td>39.7%</td>
<td>39.7%</td>
<td>38.9%</td>
<td>40.5%</td>
<td>38.9%</td>
</tr>
<tr>
<td>Public transport season ticket ownership</td>
<td>46.6%</td>
<td>46.4%</td>
<td>46.1%</td>
<td>46.4%</td>
<td>46.4%</td>
<td>46.0%</td>
<td>46.6%</td>
<td>46.0%</td>
</tr>
<tr>
<td>Bicycle ownership</td>
<td>36.2%</td>
<td>35.9%</td>
<td>35.6%</td>
<td>35.9%</td>
<td>35.8%</td>
<td>35.6%</td>
<td>36.2%</td>
<td>35.6%</td>
</tr>
<tr>
<td>Car club SDCS subscription</td>
<td>--</td>
<td>2.7%</td>
<td>5.4%</td>
<td>2.8%</td>
<td>2.8%</td>
<td>5.4%</td>
<td>0.1%</td>
<td>5.4%</td>
</tr>
<tr>
<td>‘Versatility’ SDCS subscription</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>19.9%</td>
<td>18.6%</td>
<td>19.9%</td>
</tr>
<tr>
<td>Predicted usage of methods of transport, as a percent of the E-NTS sample of Londoners’ travel (refer to discussion immediately above this table)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal car driving</td>
<td>37.5%</td>
<td>36.8%</td>
<td>34.6%</td>
<td>36.4%</td>
<td>36.4%</td>
<td>33.3%</td>
<td>37.4%</td>
<td>33.4%</td>
</tr>
<tr>
<td>Personal car passenger</td>
<td>19.1%</td>
<td>18.7%</td>
<td>19.3%</td>
<td>18.7%</td>
<td>18.7%</td>
<td>18.6%</td>
<td>18.6%</td>
<td>18.6%</td>
</tr>
<tr>
<td>Public transport</td>
<td>27.6%</td>
<td>28.2%</td>
<td>28.6%</td>
<td>28.2%</td>
<td>28.2%</td>
<td>27.4%</td>
<td>28.2%</td>
<td>27.7%</td>
</tr>
<tr>
<td>Cycling</td>
<td>2.7%</td>
<td>2.6%</td>
<td>2.6%</td>
<td>2.7%</td>
<td>2.6%</td>
<td>2.6%</td>
<td>2.7%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Walking</td>
<td>11.9%</td>
<td>11.9%</td>
<td>12.3%</td>
<td>12.0%</td>
<td>12.0%</td>
<td>11.9%</td>
<td>11.8%</td>
<td>12.0%</td>
</tr>
<tr>
<td>Taxi/minicab</td>
<td>1.2%</td>
<td>1.3%</td>
<td>1.4%</td>
<td>1.3%</td>
<td>1.3%</td>
<td>1.3%</td>
<td>1.3%</td>
<td>1.3%</td>
</tr>
<tr>
<td>Car club SDCS car driving</td>
<td>--</td>
<td>0.6%</td>
<td>1.3%</td>
<td>0.7%</td>
<td>0.7%</td>
<td>1.0%</td>
<td>0.0%</td>
<td>1.0%</td>
</tr>
<tr>
<td>‘Versatility’ SDCS car driving</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>3.8%</td>
<td>0.0%</td>
<td>3.8%</td>
</tr>
<tr>
<td>All driving</td>
<td>37.5%</td>
<td>37.4%</td>
<td>35.8%</td>
<td>37.1%</td>
<td>37.1%</td>
<td>38.1%</td>
<td>37.4%</td>
<td>38.2%</td>
</tr>
</tbody>
</table>

Table 8.5: Summary of results from baseline scenario and Scenarios #1 through #7
8.2.4. Scenario #1

Scenario #1, in which car club SDCSs are widely-available in Inner London (and consumers are widely aware of them), results in a predicted 2.7% of the E-NTS sample choosing to subscribe to a car club. This corresponds to 9.6% and 17.0% of adults and driving-licence-holders in Inner London, respectively.

The predicted changes in holdings of other mobility resources (as a percentage of all E-NTS respondents) are:

- 1.7% fewer car owners in all of Greater London (39.8% of Londoners predicted v. 40.5% in the baseline scenario)
- 0.4% fewer holders of public transport season tickets (46.4% v. 45.6%)
- 0.7% fewer bicycle owners (35.9% v. 36.2%)
- 0.8% fewer people owning no mobility resources (18.1% v. 18.2%)

The largest share (74.5%) of people predicted to subscribe to a car club SDCS are people who presently do not own personal cars, though the largest single elasticity between car club subscription and holdings of any of the mobility resources is with car ownership. Interestingly, the prediction is for a decrease also in holdings of season tickets and bicycles, with season-ticket-holding experiencing the smallest proportional decrease. Whilst the elasticities are smaller than that with car ownership, it is noteworthy that in this scenario car club subscription is predicted to in the aggregate be a substitute (rather than a complement) to all three other mobility resources, not just car ownership.

The predicted changes in usage of transport modes are:

- 1.8% fewer private car driving journeys (36.8% of Londoners’ journeys predicted v. 37.5% in the baseline scenario)
- 2.1% more public transport journeys (28.2% v. 27.6%)
- 3.2% fewer cycling journeys (2.6% v. 2.7%)
- 2.2% fewer car passenger journeys (18.7% v. 19.1%)
- 0.4% fewer [longer-than-one-mile] walking journeys (11.87% v. 11.92%)
- 3.2% more taxi/minicab journeys (1.3% v. 1.2%)
- 0.6% of journeys performed in car club SDCS cars. This implies roughly four journeys/week (or two two-leg tours) on average by each subscriber.\(^{15}\)
- 0.2% fewer driving journeys (combined private car driving + SDCS usage)

In summary, the introduction of car club SDCSs into the analysis is predicted to lead to more journeys performed by public transport and taxi/minicab services, and fewer by [personal] car driving, cycling, car passenger, and walking.
Figure 8.1: Cumulative distribution of the change in the number of car driving journeys during NTS observation period (7-day diary + 4-week LDJ period) and Scenario #1 by the 19 people predicted to subscribe to a car club SDCS in scenario #1.

Figure 8.1 shows the predicted change in frequency (journeys/week) of all car driving journeys (whether using personal cars or SDCS cars) in the baseline scenario and Scenario #1 for the 19 people predicted to subscribe to a car club in Scenario #1. The model predicts that four of these 19 people will drive less (fewer journeys) than before, one will drive the same amount, and 14 will drive more. But this figure shows that those people predicted to drive less are predicted to drive a lot less (an average of -11 journeys/week), whilst those predicted to drive more are predicted to change their amount of driving journeys to a more modest degree (an average of +3 journeys/week). The net effect on the number of driving journeys is predicted to be small and negative – an average decrease of roughly 5% amongst this group. (Note that the estimated confidence levels for this figure are not presented here.) These predicted patterns – and the net effect in terms of number of driving journeys – are broadly consistent with findings by others (e.g. Martin and Shaheen 2010). Whilst substantively of interest, analysis of the characteristics of these journeys is not reported here.

8.2.5. Scenario #2

Scenario #2 simulates the expansion of car club SDCS services to encompass both Inner and Outer London.
Car club subscription increases from the 2.7% of Londoners predicted in Scenario #1 to 5.4%, though this is less than proportional to the increased population which would be served (Outer London’s population is roughly 50% larger than that of Inner London). Adults in Outer London are more likely to have a driving licence, though they are predicted in this scenario to subscribe to a car club SDCS at a lower rate: 8.2% of licence-holders in Outer London are predicted to subscribe, versus the 17.0% of Inner London licence-holders predicted in Scenario #1.

Introducing car clubs across all of London (including Outer London) is predicted to lead to substantially different patterns of modal shift than those associated with service in only Inner London. Walking and car passenger travel, for instance, are both predicted to increase from introducing car club SDCSSs across all of London, whilst they were predicted to decrease from the introduction of car club SDCSSs into only Inner London.

Car club cars are predicted to be used to perform 1.3% of journeys in this scenario, a rough doubling from that predicted in Scenario #1. Looking closer at the results, the number of car club subscribers is projected to increase by 98.4%, whilst the number of journeys is projected to increase by 101.0% – this implies that the subscribers in Outer London are projected to use the service’s vehicles very marginally more than subscribers in Inner London. This is a rather intuitive finding given the urban geography of Outer London; if anything the difference might have been expected to be larger.

It is also of note that the combined level of driving (private cars + SDCS usage) is markedly lower in this scenario than in either the baseline scenario or Scenario #1. In other words, introducing car clubs to Outer London is predicted to lead to substantially less driving, whilst introducing them only to Inner London is predicted to lead to a much smaller and less clear-cut effect in the number of driving journeys.

**8.2.6. Scenario #3**

Scenario #3 investigates how the StraP model responds to a small reduction (5%) in car club usage fees. There was predicted to be a 4.0% increase in the rate of car club SDCS subscription associated with this change in the service characteristics.

Car club usage is predicted to increase by a larger proportion (6.3%). Thus car club *subscription* is predicted to be relatively inelastic (a calculated elasticity of -0.88) with respect to usage price, though usage is predicted to be elastic (-1.27).

**8.2.7. Scenario #4**

This scenario looks at the StraP model system’s response to a small reduction (5%) in itinerary duration for SDCS journeys.\(^{18}\)
The results of this scenario report that the elasticity of car club subscriptions with respect to travel duration is predicted to be smaller (-0.59) than with respect to usage costs. Car club SDCS usage is, however, predicted to be more elastic with respect to travel duration (1.68) than usage costs.

8.2.8. Scenario #5

This is the first scenario to include, in addition to the car club service model, the one-way-style SDCS which we have termed ‘Versatility’. Though the initial operating area for a ‘Versatility’-style service would likely be limited to portions of central London, both services are made available across London in this scenario. This is in recognition of the fact that with the data at hand it is only known whether each journey performed by E-NTS respondents begins and ends somewhere within Greater London; in this scenario ‘Versatility’ may only be used for journeys both beginning and ending within Greater London.

The predicted impact of the introduction of Versatility on people’s holdings of mobility resources are estimated by comparing the results of this scenario with those of Scenario #2:

- 0.5% fewer car owners (38.9% of Londoners predicted in Scenario #5 v. 39.1% Scenario #2)
- 0.2% fewer public transport season ticket holders (46.0% v. 46.1%)
- 0.2% fewer bicycle owners (35.6% v. 35.7%)
- 0.45% more car club SDCS subscribers (5.42% v. 5.40%)
- 5.4/7.2/11.2% of Londoners/adults/licence-holders subscribing to car club SDCS service, respectively
- 19.9/26.3/41.0% of Londoners/adults/licence-holders subscribing to the ‘Versatility’ service, respectively

As with the introduction of car club SDCSs, the newly-introduced the ‘Versatility’ SDCS is predicted to be a substitute with car, season-ticket, and bicycle ownership, with car ownership being the closest substitute.

It is of interest that the predicted level of subscription to car club SDCSs is very marginally higher in this scenario than Scenario #2, where car clubs also would serve both Inner and Outer London. Thus the two services are found, in this range of subscription levels, to have very weak cross-elasticity. The very small increase in predicted car club SDCS subscription levels associated with the introduction of the ‘Versatility’-style SDCS nevertheless indicates that the two SDCSs are acting as [very weak] complements to each other.
The most striking result of this scenario, though, is the prediction of roughly three-and-a-half times as many ‘Versatility’ subscribers as car club subscribers.

The introduction of Versatility is predicted to have the following impacts on usage of the various methods of transport:

- **3.5% fewer** private car driving journeys (33.3% of Londoners’ journeys predicted in Scenario #5 v. 34.5% in Scenario #2)
- **4.2% fewer** public transport journeys (27.4% v. 28.2%)
- **No change** in the percentage cycling journeys (2.6% in both scenarios)
- **3.7% fewer** private car passenger journeys (18.6% v. 19.3%)
- **2.7% fewer** [longer-than-one-mile] walking journeys (11.9% v. 12.3%)
- **3.1% fewer** taxi/minicab journeys (1.31% v. 1.35%)
- **21.5% fewer** car club SDCS journeys (1.0% v. 1.3%)
- **3.8% of journeys performed** in ‘Versatility’ SDCS cars
- **1.7% more** driving journeys (combined private car driving + SDCS usage)

‘Versatility’ is thus predicted in this scenario to substitute for travel by all other methods of travel, with the exception of cycling. The largest elasticity with any of the non-SDCS methods of travel is predicted to be with public transport. Intriguingly, introducing ‘Versatility’ is predicted to decrease the number of car club journeys by 22% – it would be of much substantive interest to investigate whether this result remains the same if more fine-grained characteristics of spatio-temporal service characteristics (i.e. the availability of SDCS vehicles in different places at various times) are taken into account. (See footnote #9.) Further, in this scenario the total number of driving journeys, including those performed by both private cars and SDCS vehicles, is predicted to increase by roughly 2% despite the decrease in the number of private car journeys.

As a final observation, we note that average ‘Versatility’ subscriber is predicted to perform 3.2 one-way ‘Versatility’ journeys per week in this scenario, whilst car club SDCS subscribers are predicted to perform an average of 3.1 one-way journeys19 via car club SDCS cars in this scenario. Thus [the much-larger pool of] ‘Versatility’ subscribers are predicted to use that SDCS very marginally more intensively than car club subscribers. Whilst intuitive, one may have perhaps expected this effect to be somewhat stronger.
8.2.9. Scenario #6

Scenario #6 was designed to investigate the possibility of the ‘plateau’ effect leading to the prediction of unreasonable forecasts when usage attributes of SDCSs take implausible values (chosen in this instance to be 100x their standard values).

In this set of circumstances, the $V_{di,\text{travel}}$ term for all portfolios including one or both of the SDCSs takes essentially the same value as the $V_{di,\text{travel}}$ for the corresponding portfolio which does not include the SDCS(s). (e.g. $V_{di,\text{travel}}^{\text{own bicycle} + \text{car club subscription}} \approx V_{di,\text{travel}}^{\text{own bicycle}}$). Thus, continuing with this example, the only difference in systematic utility between these two portfolios would arise from the portfolio-level ASC for car club subscription and the disutility associated with the monetary costs of the subscription.

This scenario reports nearly nil predicted car club SDCS subscriptions (0.09%), though it predicts that 18.6% of Londoners would subscribe to the ‘Versatility’ SDCS. No usage is predicted for either SDCS.

The relatively-small negative portfolio-level ASC for ‘Versatility’ subscription (-0.3) leads each alternative that includes a ‘Versatility’ subscription to appear only marginally less-desirable in the choice model than its otherwise-identical counterpart which does not include a ‘Versatility’ subscription. The much larger negative portfolio-level ASC for car club subscription (-4.8), however, results in each portfolio that includes a car club subscription appearing substantially less attractive than its otherwise-identical counterpart portfolio, following a similar argument.

This implausible result (the level of ‘Versatility’ subscriptions) arises from the assumption of independently-distributed error terms across the portfolio options. The ‘plateau’ effect appears to be a property of portfolio-style models with the sort of naive error structure presented in this initial application. (See Section 7.3) This result highlights that relaxing the assumption of IID error terms to account for potential correlated error terms is a logical avenue for future research into portfolio-style choice models involving multiple interacting elemental alternatives.

8.2.10. Scenario #7

This scenario was performed as a diagnostic, to examine the degree to which the results of the prior scenarios may be affected by the implausibly-high mode share of cycling in the AVATAR survey (27%). The holdings of mobility resources are identical to those of Scenario #5; the only change is that this scenario uses the ‘mode choice’ alternative-specific constant for cycling that was estimated exclusively from the E-NTS dataset (rather than the joint E-NTS/AVATAR estimation).
This results in cycling’s mode share predicted to be 2.0%, rather than 2.6% in Scenario #5 (the otherwise-identical scenario) – a reduction of 30%. Usage of the car club SDCS is predicted to remain unchanged, whilst ‘Versatility’ SDCS usage is predicted to decrease by 0.6%, from 3.78% to 3.75%.

In other words, the impact of the high level of cycling in the AVATAR survey on the SDCS forecasts is predicted to be quite small. This appears to be due to the predominance of ‘mode-choice’ observations from the E-NTS in the joint estimation process – there were roughly seven times as many mode-choice observations from the E-NTS dataset than from the AVATAR survey dataset.

### 8.3 Correlation patterns in people’s simulated holding and usage of mobility resources

Section 7.2 investigated patterns in people’s holdings of multiple mobility resources; it was found, for instance, that people owning a personal car were more likely than average to also own a public transport season ticket. (See Table 7.4.) In this section we explore the similarities and differences in the corresponding patterns of predicted choices.

<table>
<thead>
<tr>
<th># Car driving journeys</th>
<th># Public transport season ticket ownership</th>
<th># Public transport journeys</th>
<th># Bicycle ownership</th>
<th># Cycling journeys</th>
<th># Walk journeys</th>
<th># Taxi-minicab journeys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car ownership</td>
<td>.85</td>
<td>.01</td>
<td>-.30</td>
<td>-.05</td>
<td>.17</td>
<td>-.24</td>
</tr>
<tr>
<td>Public transport season ticket ownership</td>
<td>.05</td>
<td>-.22</td>
<td>-.01</td>
<td>-.11</td>
<td>-.09</td>
<td>-.03</td>
</tr>
<tr>
<td># Public transport journeys</td>
<td>.35</td>
<td>-.03</td>
<td>.02</td>
<td>.00</td>
<td>.02</td>
<td></td>
</tr>
</tbody>
</table>

| # Bicycle ownership    | then continuously-varied shading to...   | .03                       | .15                | .25               | .14             |                        |
| # Cycling journeys     | Values -0.5 and lower in bright red     | .56                       | .14                | .00               |                |                        |
| # Walk journeys        | Values of zero in yellow                | .29                       | .03                |                  |                |                        |
| # Taxi-minicab journeys | Values 0.5 and higher in bright green   | .16                       |                    |                  |                |                        |

Table 8.6: Correlation matrix of simulated choices (‘portfolio’ holdings and ‘mode’ usage) from the baseline scenario (#0)

Table 8.6 shows that the StraP model recovers much of the correlation pattern from the observed E-NTS dataset (see Table 7.4). The cells in the two matrices have a correlation coefficient of 0.77.

The simulated choices do not capture the positive correlation between car ownership and public transport season ticket holding, but they do replicate the lack of correlation between car ownership and bicycle ownership and between season ticket holding and bicycle ownership.

The negative relationship between car ownership and public transport usage is well-represented (-0.30 v. -0.18). The largest difference is the degree of correlation between bicycle ownership and cycling; this is predicted to be 0.56, whilst it was observed to be only 0.20. In other words, the model does a relatively poor job of predicting cycle ownership amongst E-NTS respondents who own
a bicycle but used it lightly (or not at all) during their diary week. This is not surprising, as the model hypotheses that ownership of mobility resources is largely a function of a person’s expectations of the value to be gained from using the resource in the future, and in this application the only representation of people’s expected travel was the travel they reported in the NTS.

If it were desired, these correlations could in principle be represented arbitrarily-closely through the use of a more-flexible error structure in the estimation process; drawbacks to this method are that it could be difficult to interpret the additional error terms, and unclear whether they are robust in forecasting applications which would be expected to change the structure of the market for personal mobility.

Table 8.7 shows how the predicted ‘portfolio’ holdings from the baseline scenario align with the E-NTS respondents’ observed portfolios. For instance, the cell at the top-left of the matrix shows that, of the [weighted] 84 people observed to own a car only, 25 were ‘correctly’ predicted to also own a car and no other mobility resources. It shows that the model does best at predicting people’s simple portfolios, and relatively poorer at predicting the precise holdings of those people holding a larger number of mobility resources:

- People observed to hold no mobility resources: 30% ‘correctly’ predicted
- People holding one: 30%
- People holding two: 18%
- People holding all three: 6%

<table>
<thead>
<tr>
<th>Observed 'portfolio' holdings</th>
<th>Predicted 'portfolio' holdings</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>AB</th>
<th>AC</th>
<th>BC</th>
<th>ABC</th>
<th>None</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>25</td>
<td>9</td>
<td>5</td>
<td>19</td>
<td>11</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>84</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>20</td>
<td>39</td>
<td>16</td>
<td>28</td>
<td>13</td>
<td>9</td>
<td>25</td>
<td>174</td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td></td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>15</td>
<td>2</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>53</td>
</tr>
<tr>
<td>AC</td>
<td></td>
<td>16</td>
<td>5</td>
<td>4</td>
<td>15</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>2</td>
<td>10</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>ABC</td>
<td></td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td></td>
<td>13</td>
<td>30</td>
<td>33</td>
<td>7</td>
<td>9</td>
<td>17</td>
<td>5</td>
<td>164</td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td>97</td>
<td>120</td>
<td>87</td>
<td>98</td>
<td>61</td>
<td>72</td>
<td>46</td>
<td>122</td>
<td>704</td>
</tr>
</tbody>
</table>

NB: These figures are based on a single simulated choice per E-NTS respondent, and thus vary marginally from the percentages reported in Section 8.2.3

Codes: A: Owning a personal car
B: Owning a public transport season ticket
C: Owning a bicycle

Table 8.7: Cross-tabulation of observed and predicted (baseline scenario) ‘portfolio’ holdings
### Table 8.8: Correlation matrix of simulated choices (‘portfolio’ holdings and ‘mode’ usage) from Scenario #1

Table 8.8 shows the correlation of car club SDCS activity (both subscription and usage) with other holdings and usage of transport modes. Car club activity is predicted to generally be weakly correlated with the holding other mobility resources and using other methods of travel. The exception to this is a positive association with the amount of walking a person does; the model is predicting that people who subscribe to a car club will tend to walk more frequently than average.

The correlation of car club subscription with car ownership is very weakly negative. Indeed, the model is predicting that there would be some overlap – i.e. people who are both car owners and car club subscribers. (It is predicting that roughly ¼ of subscribers to a car club would also be car owners.) Interestingly, of the 11 stated-choices of car club subscription in the AVATAR survey, there was only one occasion (a frequency of 9%) where a respondent indicated that they would concurrently choose to both own a car and subscribe to a car club SDCS. Of the 22 AVATAR survey respondents who were real-world car club subscribers, seven (32%) were also car owners. In summary, the model is predicting that a substantial minority of car club SDCS subscribers would also be car owners, which is broadly-consistent with empirical observations of existing car club subscribers (Harmer and Cairns 2011).

### Table 8.9: Correlation matrix of simulated choices (‘portfolio’ holdings and ‘mode’ usage) from Scenario #5
Table 8.9 shows the predicted correlations between ‘Versatility’ SDCS activity and other holdings and usage. The pattern here is more distinctive than that found in Table 8.8. People subscribing to the ‘Versatility’-style SDCS are predicted to be marginally more likely than average to also purchase a car and to also subscribe to a car club SDCS.

To summarise the results presented in this section, the model is able to match observed correlations fairly well in holdings of those mobility resources and usage of those methods of transport that were present in the E-NTS dataset, and predicts rather weak correlations of SDCS ‘activity’ (holdings and usage) with non-SDCS ‘activity’.

8.4 Sensitivity of results to representation of peoples’ Perceived Activity Sets

When scoping the design of what would become the E-NTS dataset (see Chapter 4), one criterion was that the dataset should capture as long a period of people’s activity/travel as possible to represent their perceived activity set. Figure 8.2 shows how goodness-of-fit, using the ‘portfolio’ specification, varies with the number of people’s journeys ‘observed’. For instance, when each NTS respondent’s first journey during their diary week is included in the analysis, the likelihood (with the same number of parameters) improves by 42 points versus when none of their journeys are included. Adding people’s second journey into the specification improves likelihood by 8 points, and so on.

![Figure 8.2: Response of likelihood function, using the ‘portfolio’ specification and the E-NTS dataset only, to successive increases in the ‘observed’ number of each journeys performed by each person (i.e. the representation of their perceived activity set)](image-url)
The curve in Figure 8.2 shows strongly-diminishing returns to increases in the representation of people’s PAS around 20 journeys, and very little response at all after roughly 30 — 35 journeys.\textsuperscript{21}

Thus, the empirical evidence appears to indicate that, even excluding the particular ‘gap-filler’ characteristic of people’s use of SDCSs, there are substantial gains in explaining people’s holdings of mobility resources (car/bicycle/public-transport-season-ticket ownership) to be obtained by use of long-duration activity/travel diaries rather than one- or two-day diaries. The growth of ‘gap-filler’ modes such as SDCSs, which seem to afford people with the opportunity to consider re-structuring their holdings of mobility resources, provides a further reason for preferring long-duration travel diaries in future research and applications employing ‘portfolio’ model forms.\textsuperscript{22}

![Figure 8.3: Response of values of the salience (gamma) parameters to increases in the number of people’s journeys which are taken into account, from the same specification as Figure 8.2](image)

Figure 8.3 shows, for the same specification as Figure 8.2, how the salience parameters vary with the number of journeys ‘observed’. The salience parameters tend to decrease inversely with the number of journeys, as would be expected, though there is an unexplained anomaly in a range around 5 — 10 journeys. The relative values of the various salience parameters are also somewhat volatile; at small numbers of ‘observed’ journeys the salience parameter for travel to/from social out-of-home activities is relatively large, whilst at large numbers of observed journeys it becomes small relative to the other salience parameters. The opposite can be observed for the salience parameter for travel to/from shopping, personal-business, and ‘other’ activities.

We also note that the ‘true’ relationship of the magnitude of the salience parameters with the number of journeys would be expected to be inversely proportional (i.e. as more and more journeys are observed, the salience of each journey would be expected to monotonically decrease). The observed trend at large numbers of journeys is flat, however. This likely arises from the truncation associated with the NTS week-long diaries (and similarly the four-week long-distance journey reporting period). As the number of journeys ‘observed’ is increased from 30 to 35, for instance, this only affects the 5% of E-NTS respondents who were observed to perform more than 30 journeys.


8.5 Summary
This Chapter presented empirical results from the StraP model. The parameter sets to be used in the forecasting application were presented first, followed by a discussion of a metric known as the ‘Value of time’, which identified the possibility of (though not definitive proof of) systematic differences in people’s choice-making processes at the ‘strategic’ and ‘tactical’ levels.

A set of scenarios were then presented, which highlighted the complex relationships predicted by the model between people’s holdings and usage of mobility resources.

An analysis of correlation patterns in people’s simulated choices showed that the StraP model does a fair job of capturing the observed correlation patterns, without explicit parameterisation to that effect. Finally, it was shown that enriching each person’s unique perceived activity set beyond the few journeys which would be captured in a one- or two-day travel diary provides a large increase in explanatory power using the StraP specification.

Chapter nine summarises the research reported in this thesis.

Notes

1 The one exception to this statement holding for all parameters is the travel time parameter for car driving which is estimated from the AVATAR survey results. (See Section 7.5.) The corresponding parameter estimated from the E-NTS data records is ‘correctly’ signed and taken forward for the forecasting application.

2 We note that in a traditional mode choice analysis people’s patterns of activity participation are assumed to be fixed; this is a rather strong assumption which appears to run counter to empirical findings regarding the phenomenon known as ‘induced demand’ (Hills 1996, Downs 1992, SACTRA 1994, Noland and Lem 2000) particularly in the context of SDCs. (Adamou 2011.) If this assumption does not hold the resulting parameters (and hence the calculated VoT) may be biased to some degree.

3 Note that this ratio includes an adjustment such that the average of the salience parameters is equal to 1.0 for all runs. This accounts for differences in the scale of the logsum term as described in Chapter four.

4 The attributes of car ownership were presented to AVATAR survey respondents as:

“A car must be taxed, insured and maintained. Joe [Jane] does not expect to have a parking space, so he [she] would park his car on the street at night. Motoring organisations say that owning a mid-sized car in London costs around £4000 per year, excluding petrol and parking.”

5 People may, for instance, plausibly discount future travel time more than costs as monetary costs may be perceived as more tangible.
It is recognised that it is in principle possible for the introduction of SDCSs to lead to a non-licence-holding adult to decide to acquire a licence.

The decision to specify journeys as the unit of analysis for forecasting purposes, rather than tours, was not taken lightly. Tours are a more appropriate unit of analysis for assessing uptake of car club SDCS service models, due to the operational requirement for users to return a vehicle to the same location as it was taken. This is not the case for ‘Versatility’-style SDCSs, however, due to their ‘one-way’ feature. In the interests of pragmatism, journeys were specified to be the unit of analysis for the present forecasting effort. The analysis of transport mode choice at the level of people’s tours involves a number of complexities and is the subject of recent and current research (e.g. Miller et al. 2005, Frank et al. 2008), thus it was decided to allow the detailed analysis of people’s tours in the context of SDCSs to also remain a topic for future research. For the present research, the heuristic specification for the ‘activity time’ portion of car club fees was that one-half would accrue to the ‘to-activity’ journey and one-half to the ‘from-activity’ journey.

It should also be noted that NTS travel diaries do not capture all out-of-home travel performed by respondents; the most significant travel that is not captured are short walks (under 1 mile / 20 minutes). Therefore there will be some upward bias in the imputed activity times prior and subsequent to each NTS activity. To gauge the order of magnitude of this effect, consider that 19% of the journeys that E-NTS respondents performed on day seven of their diaries would not have been reported on days one through six, though it should be kept in mind that these 19% of journeys are distinctive in their characteristics rather than a random sampling. This effect is taken into account in weighting of journeys, but not in the calculation of activity duration.

Though not reported here, it is recognised that the results are idiosyncratic, and that performing multiple draws and averaging the results would eliminate the potential for a single random draw to have a material impact on the predicted results.

Scenarios 3 and 6 were chosen to highlight the elasticities in forecasts, whilst keeping the variation of attributes in a narrow range on either side of their value in the parameter estimation phase. (Refer to Sections 8.1.2 and 8.1.4) All scenarios rely on the following two assumptions:

- SDCS cars are accessible within a 5/10 minute walk in Inner/Outer London respectively.
- Payment for car club SDCS usage accrues for 15 minutes of ‘buffer time’ before and 15 minutes after usage (to mimic the temporally-fixed-advance-reservation feature of car club SDCS usage)

Consumer awareness is noted here to account for the fact that AVATAR survey respondents had information on SDCSs at hand, whilst at present consumer awareness of car-club-style services is quite low.

Whilst it is somewhat unrealistic to expect a ‘Versatility’-style service to be initially available in peripheral parts of London, the E-NTS dataset does not capture whether people’s out-of-home activities taking place in London are located in Inner or Outer London. Thus this scenario can be considered as subject to some optimism bias due to this characteristic.

A review of the parameters associated with cycle ownership and cycling from the E-NTS and AVATAR survey datasets (found in Tables 7.7, 7.9, 7.11, and 7.12) identified the mode-choice-level ASC for cycling as the strongest manifestation of the implausible frequency with which cycling was chosen in the AVATAR survey. Its value when estimated from only the E-NTS [AVATAR] dataset was -3.4 [+1.0].

We note that the market shares discussed in this section refer to each mobility resource independently. The specification (see Chapter seven) includes portfolio-level ASCs for each mobility resource, of which there are three in the E-NTS dataset (car ownership, public transport season ticket ownership, and bicycle ownership.) The properties of the logit model imply that these freely-estimated ASCs will result in ‘modelled’ shares of ownership of these mobility resources that exactly match those of the estimation dataset. An alternative specification could include separate portfolio-level ASCs for each mobility resource; for the E-NTS dataset there would thus be seven \(2^3 - 1\) such ASCs rather than three, and the ‘modelled’ shares of each
combination of mobility resources would precisely match those of the estimation dataset. This latter specification would become less practical as the number of mobility resources in the analysis grows; there would for instance be 31 such ASCs in analyses of the AVATAR survey dataset.

14 The bias introduced into the predicted ‘mode’ choice ‘market shares’ of the E-NTS sample is smaller than that introduced into the predicted ‘portfolio’ choice market shares as the E-NTS ‘mode’ choice sample is a larger proportion of the total (E-NTS + AVATAR) ‘mode’ choice sample used in the joint estimation. The E-NTS dataset provides 88% of the ‘mode’ choice observations \((10,575/(10,575 + 1,440))\), but only 71% of the ‘portfolio’ observations \((704/(704 + 288))\).

15 Though the numbers are not directly comparable, we note here the self-reported car-driving frequency of existing London-based car club members: 11%/13%/38%/38% report driving 3+ times per week/1-2 times per week/more than once per month/less than monthly or never, respectively. (Harmer and Cairns 2011) It would appear that this analysis is predicting somewhat more intensive use per subscriber than the empirical [self-reported] observations of existing car club SDCS subscribers.

16 It is noted that the use of the term ‘week’ in this discussion is concise but not strictly accurate, as the NTS respondents report their journeys for a seven-day period and 50+ mile ‘long-distance’ journeys for a further three weeks.

17 The analysis presented in this paragraph is based on a single simulated choice per E-NTS respondent.

18 Plausible policies for delivering such an improvement in journey time to SDCS cars without concurrently speeding up general-purpose road traffic might include, for instance, allowing access to bus lanes, amongst other measures.

19 This implies 1.6 round-trip two-leg tours, if all of these tours consist of two journeys.

20 There is some third-party data on this point as well, though not directly comparable: 20% of car club subscribers living in London report that their household owns a car. (2009/10 Carplus survey) The figures are not directly comparable as the Carplus survey asks respondents about household car ownership, whilst the results presented here pertain to personal car ownership.

21 77%/95%/99% of NTS respondents reported 20/30/40 or fewer journeys during their diary week, respectively. See Figure 5.2.

22 Though not reported here, this analysis would benefit from investigating how the pattern of goodness-of-fit varies by the number of days for which people’s activity/travel is observed, supplementing the analysis by number of journeys as reported here.
Chapter 9: Summary and conclusions

The research reported in this thesis was motivated by the challenge posed by a new form of personal mobility, car clubs, which we have classed within a more general category termed subscription drive-it-yourself car services [SDCSs]. Cooperative mobility systems of this sort blur the distinctions between public and private mobility networks, bearing some of the hallmarks of each of these classes. SDCSs are one manifestation within the automotive sector of a wider socio-technical trend referred to as servicisation; they have become commercially viable as the costs of the requisite information-technology infrastructure have fallen in recent years. It is suggested that the form of access to automobility that they provide falls on a continuum between opposing poles of personal car ownership and one-off access to the use of a car:

![Diagram of Continuum of Access to Car Use](attachment:image.png)

Understanding servicised personal mobility, and how people interact with such systems, can reasonably be expected to be of growing importance as SDCSs evolve and increase their footprint in the marketplace. The implications of these services are ambiguous and hotly-debated: might they lead to people driving more or less, to increased or decreased accessibility to life opportunities, to more emissions or fewer, greater demands for scarce urban real estate or less, etc. This research attempts to set a framework for analysing people’s patterns of engagement with SDCS services, and provide an initial empirical study to shed light on these questions, though it was not designed to – and cannot – definitively answer them.

The principal challenge to understanding how SDCSs may grow was identified to be the inter-linking between strategic and tactical levels of people’s choice-making when acting as consumers of mobility. By way of contrast, researchers and practitioners are accustomed to analysing issues such as car ownership and usage independently and in sequence.
This research shows that it indeed appears possible to quantitatively detect indications that strategic planning plays a part in people’s actions relating to their mobility, and that the proposed methods have the capacity to not only provide insights into take-up of SDCSs, but also to provide new ways to understand linked issues such as those relating to car ownership. Given the growth in car ownership within the developing world that some are forecasting (Sperling and Gordon 2009) – with broad implications for issues such as economic development, land development patterns, and sustainable management of resources – the development of techniques such as those proposed here appears to be timely. Prudence and great caution are called for in applying these and any other techniques, however; few researchers, for instance, foresaw the recent levelling-off in car ownership within the developed world.

The possibility is raised of wider application of similar techniques beyond people’s holdings of mobility resources (such as car ownership) to related problems such as how people set the spatial and temporal structure for major activities in our lives (housing, socialising, employment, caring, education, shopping, etc.) As we increasingly find new systems and policies on the public agenda which would radically re-structure people’s spatio-temporal accessibility to life opportunities – autonomous vehicle control, carbon pricing or personal budgeting, high-speed rail networks, electric mobility, and to be sure cooperative mobility systems such as SDCSs amongst them – it would seem to be of growing importance, for both public and private-sector actors, to understand people’s adaptation strategies as faithfully as possible.

The remainder of this Chapter looks first at the substantive results reported here, before turning to the methodological novelties of this research. The discussion of these points includes reference to a number of issues raised which remain directions for future research.

## 9.1 Substantive results

The results of the empirical study are thought to be of interest to both researchers and the wider community of practitioners and policy-makers within the transport domain.

The predictions arising from the analysis rely on the slender reeds of the small sample size of the stated-choice sample, its un-representativeness (for which the weighting strategy partially corrected), and the limited statistical significance of several of the parameters. They do nevertheless represent the best guess as to outcomes on the basis of the data at hand, and on the whole are rather plausible. With these caveats in mind, it is noted that the predictions tend to support several commonly-held views pertaining to SDCSs, whilst appearing to run counter to others:
1) An SDCS system offering people the opportunity to drive a car for a one-way trip, at which point their responsibility for it ends, would be predicted to be attractive to a market segment roughly three-and-a-half times larger than can be reached by the traditional car club SDCS operating model.

2) Of the relatively-small segment of people who would subscribe to a car club SDCS, the model would predict that a minority will drive fewer journeys than otherwise and that a majority of them will drive more frequently than they otherwise would. But – those driving fewer journeys would be predicted to on average make many fewer driving journeys and those driving more than before are predicted to each drive relatively infrequently. The net effect, in terms of driving journeys, would be predicted to be very nearly zero, though this result requires further attention to state with more certainty what the model would predict in this regard.

3) The model would predict that introducing a one-way-use SDCS into London would lead to more driving journeys, whilst the introduction of the traditional car club SDCS model would be predicted to less driving. As much of the case for public-sector support (monetary and otherwise) for SDCSs relies on a belief that they lead to less driving and related externalities, this finding is rather striking. This does not imply that novel SDCS service models should be actively discouraged, but rather that outcomes associated with various service models may be quite different and policymakers should thus exercise due caution.

4) Introducing car club SDCSs to Inner London – which is very roughly the area served at present – would be predicted to lead to a small increase in public transport usage, in keeping with arguments by car club SDCS proponents that they are complementary to an efficient public transport network. It would be predicted to lead to small decreases in the number of cycling and [long] walking journeys, however. As with each of these points, these results should be viewed as indicative only and will benefit from comparison with other analytical methods and real-world experience as SDCSs continue to grow. The techniques which lead to these predictions are experimental and, whilst having been subject to the academic peer review process, the predictions should be treated by policy-makers with this in mind. The development of a wider body of literature on the use of such techniques, by both researchers and practitioners, should help to settle the remaining questions regarding them.
Rolling out car club SDCSs across Outer London would be predicted to lead to a rough doubling of the number of subscribers, though underlying this would be a lower rate of participation in Outer London than Inner London. Interestingly, though, subscribers in Outer London would be predicted to use the services very marginally more frequently than subscribers in Inner London. The impacts on usage of other transport methods would also be predicted to be different amongst Outer London subscribers.

In considering whether to access a serviced resource on an as-needed basis, rather than via the rights and responsibilities of ownership, an important factor for consumers to consider is the nature of the access they would be acquiring. There appear to be, at the least, several dimensions of service reliability relevant to SDCSs for consumers to consider, which are substantially different to the stream of service associated with car ownership:

- Can I rely on the SDCS operator continuing indefinitely as an operating business? Will it continue to provide service in locations that work for me? If I am planning to rely on an SDCS operator as part of a ‘portfolio’ of mobility resources, to what degree can I count on the complementary products/services remaining operational?

- How likely is it that this service will continue with its present features – for instance, is there a risk of the pricing model changing to my detriment at some point after I have committed to a subscription?

- To what degree can I expect the operator to be able to resolve unanticipated events (e.g. damage to an SDCS vehicle by a third party during my period of usage) to my satisfaction?

- Will I be able to access an SDCS car when, where, and for the duration I wish, on short notice if necessary? If I am unable to do so, what are my second-best options likely to be? Is the SDCS operator likely to be able to deliver on my reservations once I have made them, and, again, if the answer is ‘no’ what would my options be given that I would have been planning on the car being available to me?

This research merely scratched the surface of the ways in which consumers may develop strategies to accommodate such issues of service [un]reliability – if cooperative mobility systems are to scale up by an order of magnitude or greater, as their proponents desire, enriching our understanding of how people manage these various aspects of service reliability can be expected to be a fruitful direction for further research.
9.2 Methodological contributions

Through its course, this research has proposed several specific methodological innovations.

The most significant perhaps relate to the methodology of the stated-choice survey. Designing an instrument to capture the dynamics between strategic and tactical levels of choice-making was found to be empirically practical here through the device of an ‘avatar’, a virtual character whom respondents were asked to interact with. This is a quite different request of a respondent than asking them to make choices on their own behalf, though it appeared to provide results that are *prima facie* comparable to those from the more traditional way of framing the request.

As analysts can expect to field requests in coming years to make detailed predictions of the impacts of ever-subtler policies, further research into the degree to which respondents to high-cognitive-intensity stated-response surveys (with or without the avatar device) engage in utility-maximising behaviour, or alternatively various forms of heuristic choice-making strategies appears warranted. Particularly relevant questions regard the extent and manner in which such choice-making strategies may extend beyond the realm of the hypothetical ‘stated-response’ context into real-world behaviour. It was found, for instance, that the relationships between holdings of mobility resources and their use were more structured and intuitive in the [hypothetical] stated-choice dataset than the [real-world] revealed-choice dataset. Field testing of the survey instrument revealed the importance of clear visual linking between the strategic and tactical levels of choicemaking on the game board, though again it is noted that the quantitative analysis shows that the strategic-tactical links inferred from the SC dataset were [undesirably] markedly stronger than those from the RC dataset.

The modified-efficient design of the stated-choice survey reported here is innovative in two respects. The first was the development of a new technique to attempt to maximise the information obtained from a stated-choice survey with respect to *some* parameters rather than *all* parameters, which potentially has applications more generally in designing joint stated-choice/revealed-choice analyses. The second relates to the setting of an empirical distribution of attribute levels from which the modified-efficient algorithm draws; this is in contrast to both efficient approaches where the distribution is unconstrained by empirical data and factorial approaches, and is intended to strike a balance between the statistical properties of design-efficiency and the plausibility of empirically-derived attribute levels. This approach was found to be empirically tractable with the data at hand on this research.

The analytical methods in the strategic portfolio [StratP] model form also comprise several novel techniques. The first was the structuring of the options that people face in a given choice context into a set of ‘portfolios’ of the elemental options. In this instance, people were specified to be able
to choose any, all, or none from a set of mobility resources, such as owning a personal car, owning a public transport season ticket, and subscribing to an SDCS. There appear to be unresolved issues with portfolio-style forms of choice models which remain to be addressed in future research, not least of which is the specification of the error structure for such systems of multiple overlapping options. It was found that the StraP system in this application performed better at predicting people’s holdings of mobility resources when they hold relatively few of them; it does poorer at predicting more complex portfolios. It was also found that the likelihood surfaces associated with the portfolio models in this application exhibited large flat regions or plateaus, whereas typical linear-in-the-parameter utility functions have concave likelihood surfaces. These present challenges for estimation, most particularly in cases where all portfolios contain multiple ‘common’ elemental alternatives. One possibility for addressing this issue in future research is to design experiments in which the availability of common elemental alternatives is systematically restricted.

A second novel analytical method was the structuring of the choice situation to have both ‘strategic’ and ‘tactical’ dimensions. A person’s ‘strategic’ choice was specified to set the constraints on their options for their future ‘tactical’ choices. This is similar to traditional techniques in which people’s decisions about their mobility are specified to be sequential, with earlier ones constraining later ones. The key innovation here is that a person’s ‘strategic’ choice was specified to be a function of the value they perceive in having particular options available to them to choose in each of their multiple future ‘tactical’ choice situations.

The analysis proposes a new concept of a person’s level of accessibility to life opportunities, termed the perceived activity set [PAS]. The main distinction between the PAS concept (which bears many similarities to the earlier-proposed activity repertoire) and more traditional measures of accessibility is that it encompasses all activities (at particular places and times) that a person considers relevant to their life. The empirical application reported here has an important shortcoming, however, in that each person’s PAS was specified to be defined by their actual travels during a randomly-observed week. There is thus a substantial potential for endogeneity between this representation of a person’s [latent] PAS and their [observed] choices such as activity locations, frequencies, scheduling, methods of access, etc. This could have important repercussions in, for instance, failing to capture how people might restructure their patterns of activity participation (e.g. choosing, to take one familiar type of personal activity, to re-allocate food-shopping responsibilities amongst family members, or to shop at different grocery stores, perhaps with different frequency or qualities of the purchased food or of the experience) if given the opportunity to use an SDCS, and remains a logical avenue for further research.
Appendix A: Comparison of ‘distinct’ and ‘combinatorial’ model forms

We begin by denoting the utility of portfolio $d$ to person $i$ as:

$$U_d^i = V_d^i + \varepsilon_d^i$$  \hspace{1cm} (A.1)

and the utility of mode $m$ to person $i$ for performing journey $j_i$ as:

$$U_{m_{j_i}}^i = V_{m_{j_i}}^i + \varepsilon_{m_{j_i}}^i$$  \hspace{1cm} (A.2)

The predicted probabilities of person $i$ choosing portfolio $d$ and mode $m$ are, respectively:

$$P_d^i = \frac{e^{V_d^i}}{\sum_{d=1}^{N} e^{V_d^i}}$$  \hspace{1cm} (A.3)

and

$$P_{m_{j_i}}^i = \frac{e^{V_{m_{j_i}}^i}}{\sum_{m_{j_i}=1}^{M_{d_i}} e^{V_{m_{j_i}}^i}}$$  \hspace{1cm} (A.4)

We note that the bounds of the summation in the denominator of equation A.3 are subscripted with $i$ to allow the possibility that some portfolios may not be available to particular people (perhaps due to age or disability, for instance.) Equation A.4 is likewise subscripted with $d_i$ to indicate that the choice set of modes for performing journey $j_i$ is limited to those modes enabled by the portfolio $d_i$ which person $i$ holds at the time of making journey $j_i$. These subscripts are suppressed after this point in the interest of clarity.

By appealing to the iid property, the probability of person $i$ choosing the indexed set of modes $\{m_1, \ldots m_{j_i}\}$ for their set of journeys $\{1, \ldots j_i\}$ can be expressed as:

$$P_{\{m_1, \ldots m_{j_i}\}}^i = \prod_{j_i=1}^{j_i} \frac{e^{V_{m_{j_i}}}}{\sum_{m_{j_i}=1}^{M_{d_i}} e^{V_{m_{j_i}}}}$$  \hspace{1cm} (A.5)

The respective likelihood functions are:

$$L_{bundle \text{ choice dimension}} = \prod_{i=1}^{I} \psi_d^i P_d^i$$  \hspace{1cm} (A.6)

and

$$L_{mode \text{ choice dimension}} = \prod_{i=1}^{I} \prod_{j_i=1}^{j_i} \psi_{m_{j_i}}^i P_{m_{j_i}}^i$$  \hspace{1cm} (A.7)
where $\psi^i_d$ and $\psi^i_{m_j}$ are indicators taking a value of one for the portfolio which person $i$ was observed to hold, and for the travel mode which person $i$ was observed to use for journey $j$, respectively, and zero otherwise.

The likelihood for all observations (portfolio and mode choices) is:

$$L_{\text{both choice dimensions}} = \prod_{i=1}^{I} \psi^i_d \prod_{j=1}^{J_i} \psi^i_{m_j} p^i_{d,m_j}$$

(A.8)

which, by noting that both the $\epsilon^i_d$ and $\epsilon^i_{m_j}$ are [assumed to be independently] distributed with mean zero and equally-scaled variance, and suppressing the $\psi$ terms, we can re-write as:

$$L_{\text{both choice dimensions}} = \prod_{i=1}^{I} e^{\epsilon^i_d \sum_{d=1}^{D} e^{\epsilon^i_d}} \prod_{j=1}^{J_i} e^{\epsilon^i_{m_j} \sum_{m=1}^{M} e^{\epsilon^i_{m_j}}}$$

(A.9)

$$L_{\text{both choice dimensions}} = \Pi_{i=1}^{I} \frac{e^{\epsilon^i_d \sum_{d=1}^{D} e^{\epsilon^i_d}} \prod_{j=1}^{J_i} e^{\epsilon^i_{m_j} \sum_{m=1}^{M} e^{\epsilon^i_{m_j}}}}{\left(e^{\epsilon^i_d} + \sum_{d=1}^{D} e^{\epsilon^i_d}\right) \left(e^{\epsilon^i_{m_j}} + \sum_{m=1}^{M} e^{\epsilon^i_{m_j}}\right)}$$

(A.10)

$$L_{\text{both choice dimensions}} = \Pi_{i=1}^{I} \frac{e^{\epsilon^i_d \sum_{d=1}^{D} e^{\epsilon^i_d}} \prod_{j=1}^{J_i} e^{\epsilon^i_{m_j} \sum_{m=1}^{M} e^{\epsilon^i_{m_j}}}}{\left(e^{\epsilon^i_d} + \sum_{d=1}^{D} e^{\epsilon^i_d}\right) \left(e^{\epsilon^i_{m_j}} + \sum_{m=1}^{M} e^{\epsilon^i_{m_j}}\right)}$$

(A.11)

$$L_{\text{both choice dimensions}} = \Pi_{i=1}^{I} \frac{e^{\epsilon^i_d \sum_{d=1}^{D} e^{\epsilon^i_d}} \prod_{j=1}^{J_i} e^{\epsilon^i_{m_j} \sum_{m=1}^{M} e^{\epsilon^i_{m_j}}}}{\left(e^{\epsilon^i_d} + \sum_{d=1}^{D} e^{\epsilon^i_d}\right) \left(e^{\epsilon^i_{m_j}} + \sum_{m=1}^{M} e^{\epsilon^i_{m_j}}\right)}$$

(A.12)

$$L_{\text{both choice dimensions}} = \Pi_{i=1}^{I} \frac{e^{\epsilon^i_{d,m_j}} \prod_{j=1}^{J_i} e^{\epsilon^i_{m_j}}}{\sum_{d=1}^{D} \sum_{m=1}^{M} e^{\epsilon^i_{d,m_j}} \prod_{j=1}^{J_i} e^{\epsilon^i_{m_j}}}$$

(A.13)

We now turn to the ‘combinatorial’ specification, for which the function:

$$U^i_{d,m_1,...m_j} = V^i_{d,m_1,...m_j} + \epsilon^i_{d,m_1,...m_j}$$

(A.14)

represents the utility to person $i$ of each ‘combination’ (i.e. indexed set) of $\{d, m_1, ..., m_j\}$.

The attendant probability calculation and likelihood functions are:

$$p^i_{d,m_1,...m_j} = \frac{e^{\epsilon^i_{d,m_1,...m_j}}}{\sum_{d,M_1,...M_j} e^{\epsilon^i_{d,v_1,...v_j}}}$$

(A.15)
\[ \mathcal{L}_{\text{combinatorial choice}} = \prod_{i=1}^{l} \psi_{d,m_1,\ldots,m_{j_i}}^i \beta_{d,m_1,\ldots,m_{j_i}}^i \]  \hspace{1cm} (A.16)

\[ = \prod_{i=1}^{l} \frac{v_i^{i}}{e^{d,m_1,\ldots,m_{j_i}}} \frac{e^{\delta,m_1,\ldots,m_{j_i}}}{\sum_{\delta,v_1,\ldots,v_{j_i}}^{d,m_1,\ldots,m_{j_i}} e^{\delta,v_1,\ldots,v_{j_i}}} \]  \hspace{1cm} (A.17)

With reference to the assumed affine form of the utility functions, the utility \(V_{\delta,v_1,\ldots,v_{j_i}}^i\) associated with indexed set \(\{\delta, v_1 \ldots v_{j_i}\}\) can be re-written as:

\[ V_{\delta,v_1,\ldots,v_{j_i}}^i = V_{\delta}^i + V_{v_1}^i + \ldots + V_{v_{j_i}}^i \]  \hspace{1cm} (A.18)

\[ = V_{\delta}^i + \sum_{j=1}^{j_i} V_{v_j}^i \]  \hspace{1cm} (A.19)

Combining equations A.17 and A.19, we can write:

\[ \mathcal{L}_{\text{combinatorial choice}} = \prod_{i=1}^{l} \frac{v_i^{i}}{e^{d,m_1,\ldots,m_{j_i}}} \frac{e^{\delta,v_1,\ldots,v_{j_i}}}{\sum_{\delta,v_1,\ldots,v_{j_i}}^{d,m_1,\ldots,m_{j_i}} e^{\delta,v_1,\ldots,v_{j_i}}} \]  \hspace{1cm} (A.20)

The numerator and the summands in the denominator of equations A.13 and A.20 are identical. As the set indexed by \(\{\delta, v_1 \ldots v_{j_i}\}\) contains all unique combinations of its elements, we can write:

\[ |D, M_1 \ldots M_{j_i}| = |D| \times |M_1| \times \ldots \times |M_{j_i}| \]  \hspace{1cm} (A.21)

The summations in the denominators of equations A.13 and A.20 are of equal cardinality, with identical summands. As all terms in the two equations are equivalent, we conclude that the ‘distinct’ and ‘combinatorial’ model forms are mathematically equivalent under the assumptions noted in Section 4.5.
Appendix B: Derivation of the ‘plateau effect’ described in Section 7.3

The mathematical structure of the StruP model system leads to the potential, due to the logsum form, for near-zero partial derivatives of the log-likelihood function with respect to alternative-specific parameters. The presence of all no-commitment means of travel within each portfolio implies that, for at least some portions of the feasible parameter space, there may be insufficient information in the observed portfolio choices to identify parameters associated with the use of the no-commitment means of travel. The logsums across the modes of travel enabled by each portfolio would be insensitive to small changes in this parameter, as they would all dominate, or be dominated by, other means of travel. Hence the estimation, finding that the partial derivative of the utility functions (and therefore the log-likelihood function) with respect to this parameter is essentially zero, may report that this parameter is large and unidentifiable.

To expose this ‘plateau’ effect, we consider the generic logsum form:

$$LS = \lambda \times \ln \sum_{j \in C_n} e^{\left(\frac{\nu_{\text{travel}}^j}{\lambda}\right)}$$  \hspace{1cm} (B.1)$$

Without loss of generality, we can decompose $\sum_{j \in C_n} e^{\left(\frac{\nu_{\text{travel}}^j}{\lambda}\right)}$ as follows:

$$\sum_{j \in C_n} e^{\left(\frac{\nu_{\text{travel}}^j}{\lambda}\right)} = e^{\left(\frac{\nu_{\text{ni}}}{\lambda}\right)} + \sum_{j \in C_n \forall j \neq i} e^{\left(\frac{\nu_{\text{travel}}^j}{\lambda}\right)}$$  \hspace{1cm} (B.2)$$

we now have:

$$LS = \lambda \times \ln \left(e^{\left(\frac{\nu_{\text{ni}}}{\lambda}\right)} + \sum_{j \in C_n \forall j \neq i} e^{\left(\frac{\nu_{\text{travel}}^j}{\lambda}\right)}\right)$$  \hspace{1cm} (B.3)$$

Using the chain rule, we can calculate the partial derivative of the logsum $[LS]$ with respect to $V_{\text{ni}}$ as:

$$\frac{\partial LS}{\partial V_{\text{ni}}} = \left(\frac{\lambda}{e^{\left(\frac{\nu_{\text{ni}}}{\lambda}\right)} + \sum_{j \in C_n \forall j \neq i} e^{\left(\frac{\nu_{\text{travel}}^j}{\lambda}\right)}}\right) * \left(e^{\left(\frac{\nu_{\text{ni}}}{\lambda}\right)}\right) * \left(\frac{1}{\lambda}\right)$$  \hspace{1cm} (B.4)$$
\[
\frac{\partial L_S}{\partial \nu_{nl}} = \frac{e^{\left(\frac{\nu_{\text{travel}}}{\lambda}\right)}}{e^{\left(\frac{\nu_{nl}}{\lambda}\right)} + \sum_{j \in \mathcal{C}_n \forall j \neq i} e^{\left(\frac{\nu_{\text{travel}}}{\lambda}\right)}}
\]  \hspace{1cm} (B.5)

Equation B.5 asymptotically approaches zero as the ratio
\[
\frac{e^{\left(\frac{\nu_{nl}}{\lambda}\right)}}{\sum_{j \in \mathcal{C}_n \forall j \neq i} e^{\left(\frac{\nu_{\text{travel}}}{\lambda}\right)}}
\]  \hspace{1cm} (B.6)

approaches zero. Points having this characteristic can be found on the log-likelihood surface, thus leading to the ‘plateau effect’.
Appendix C: Detailed specification and results from the analysis of simulated data

<table>
<thead>
<tr>
<th>Run #</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Portfolio Choice level) ASC Own a bicycle</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a car</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a public transport season ticket</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Ride a bicycle</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Car passenger travel</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Taxi/minicab travel</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Public transport</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Walk</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Fixed holding costs in GBP per month</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car)</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.05</td>
<td>0</td>
<td>-0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.05</td>
<td>0</td>
<td>-0.1</td>
<td>-0.05</td>
</tr>
<tr>
<td>Travel time in minutes (Car passenger travel)</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0</td>
<td>-0.15</td>
<td>-0.1</td>
</tr>
<tr>
<td>Travel time in minutes (Taxi/minicab travel)</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0</td>
<td>-0.2</td>
<td>-0.1</td>
</tr>
<tr>
<td>Travel time in minutes (Public transport)</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0</td>
<td>-0.25</td>
<td>-0.1</td>
</tr>
<tr>
<td>Travel time in minutes (Walk)</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0</td>
<td>-0.3</td>
<td>-0.1</td>
</tr>
<tr>
<td>Salience parameter ($\gamma$) for ‘escort’ journey purpose</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\gamma$(leisure)</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\gamma$(shopping, personal business, and other)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\gamma$(social)</td>
<td>1</td>
<td>1.25</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\gamma$(work and education)</td>
<td>1</td>
<td>1.5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Logsum term ($L[\text{travel}]$)</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table C.1: Target parameter values for the simulated datasets
### Table C.2: Obtained parameter values for runs with simulated dataset A

<table>
<thead>
<tr>
<th>Run #</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>A7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log-likelihood</td>
<td>-737.6</td>
<td>-737.6</td>
<td>-737.6</td>
<td>756.9</td>
<td>-737.6</td>
<td>-737.6</td>
</tr>
<tr>
<td></td>
<td>$\rho^2$ / adjusted $\rho^2$</td>
<td>0.45 / 0.44</td>
<td>0.45 / 0.44</td>
<td>0.45 / 0.44</td>
<td>0.44 / 0.43</td>
<td>0.45 / 0.44</td>
<td>0.45 / 0.44</td>
</tr>
<tr>
<td></td>
<td>Smallest singular value in Hessian matrix</td>
<td>0.0034</td>
<td>0.167</td>
<td>0.167</td>
<td>0.167</td>
<td>0.165</td>
<td>0.167</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a bicycle</td>
<td>-0.515 (0.23)</td>
<td>-0.515 (0.23)</td>
<td>-0.515 (0.23)</td>
<td>-0.515 (0.23)</td>
<td>0.422 (0.29)</td>
<td>-0.515 (0.23)</td>
<td>-0.447 (0.28)</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a car</td>
<td>-0.791 (0.37)</td>
<td>-0.791 (0.37)</td>
<td>-0.791 (0.37)</td>
<td>-0.791 (0.37)</td>
<td>0.271 (0.74)</td>
<td>-0.791 (0.37)</td>
<td>-0.647 (0.44)</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a public transport season ticket</td>
<td>-0.249 (0.21)</td>
<td>-0.249 (0.21)</td>
<td>-0.249 (0.21)</td>
<td>-0.249 (0.21)</td>
<td>-0.168 (0.39)</td>
<td>-0.249 (0.21)</td>
<td>-0.239 (0.23)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Ride a bicycle</td>
<td>0.808 (0.13$^*$)</td>
<td>0.568 (0.09)</td>
<td>0.114 (0.09)</td>
<td>0.509 (0.11)</td>
<td>-8.82</td>
<td>0.102 (0.11)</td>
<td>0.582 (0.05)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car</td>
<td>0.393 (0.42$^*$)</td>
<td>0.276 (0.40)</td>
<td>0.0552 (0.40)</td>
<td>0.247 (0.41)</td>
<td>-8.99</td>
<td>0.0495 (0.41)</td>
<td>0.304 (0.30)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Shared modes</td>
<td>0$^*$</td>
<td>0$^*$</td>
<td>0$^*$</td>
<td>0$^*$</td>
<td>0$^*$</td>
<td>0$^*$</td>
<td>0$^*$</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Car passenger travel</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Taxi/minicab travel</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Public transport</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Walk</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Fixed holding costs in GBP per month</td>
<td>-0.00698</td>
<td>-0.00698</td>
<td>-0.00698</td>
<td>-0.00698</td>
<td>-0.00687</td>
<td>-0.00698</td>
<td>-0.00695</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>-0.159</td>
<td>-0.112</td>
<td>-0.0224</td>
<td>-0.100</td>
<td>-0.0155</td>
<td>-0.0201</td>
<td>-0.111</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car)</td>
<td>-0.0297</td>
<td>-0.0209</td>
<td>-0.00418</td>
<td>-0.0187</td>
<td>0.00533</td>
<td>-0.00375</td>
<td>-0.0191</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>-0.0418</td>
<td>-0.0294</td>
<td>-0.00587</td>
<td>-0.0263</td>
<td>-0.00158 (0.46)</td>
<td>-0.00526</td>
<td>-0.0269</td>
</tr>
<tr>
<td>Travel time in minutes (Shared modes)</td>
<td>-0.0903</td>
<td>-0.0635</td>
<td>-0.0127</td>
<td>-0.0569</td>
<td>-1.79</td>
<td>-0.0114</td>
<td>-0.0599</td>
</tr>
<tr>
<td>Travel time in minutes (Car passenger travel)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Taxi/minicab travel)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Public transport)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Walk)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Salience parameter ($\gamma$) for ‘Escort’ journey purpose</td>
<td>0.647</td>
<td>0.920</td>
<td>4.60</td>
<td>1.03</td>
<td>3.57</td>
<td>5.13</td>
<td>0.959</td>
</tr>
<tr>
<td>$\gamma$ (leisure)</td>
<td>0.727</td>
<td>1.03</td>
<td>5.17</td>
<td>1.15</td>
<td>4.39</td>
<td>5.77</td>
<td>1.07</td>
</tr>
<tr>
<td>$\gamma$ (shopping, personal business, and other)</td>
<td>0.703</td>
<td>1.0$^*$</td>
<td>5.0$^*$</td>
<td>1.12</td>
<td>4.34</td>
<td>5.58</td>
<td>1.0$^*$</td>
</tr>
<tr>
<td>$\gamma$ (social)</td>
<td>0.737</td>
<td>1.05</td>
<td>5.24</td>
<td>1.17</td>
<td>4.56</td>
<td>5.85</td>
<td>1.10</td>
</tr>
<tr>
<td>$\gamma$ (work and education)</td>
<td>0.636</td>
<td>0.904</td>
<td>4.52</td>
<td>1.01</td>
<td>3.94</td>
<td>5.05</td>
<td>0.940</td>
</tr>
<tr>
<td>Logsum term ($\lambda^{\text{travel}}$)</td>
<td>0.630</td>
<td>0.896</td>
<td>4.48</td>
<td>1.0$^*$</td>
<td>5.0$^*$</td>
<td>5.0$^*$</td>
<td>1.0$^*$</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed. $^*$Value is fixed $^*$Standard error using alternate calculation (Rao-Cramer bound rather than the generally more-reliable “sandwich” estimator) is $\infty$; large differences between the two estimates of standard error are indicative that the parameter set is not uniquely-identifiable as estimated.
<table>
<thead>
<tr>
<th>Run #</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-1,344.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-684.1</td>
<td>-684.1</td>
<td>-695.5</td>
<td>-866.1</td>
<td>-866.1</td>
<td>-868.4</td>
</tr>
<tr>
<td>$\rho^2$ / adjusted $\rho^2$</td>
<td>0.49 / 0.48</td>
<td>0.49 / 0.48</td>
<td>0.48 / 0.48</td>
<td>0.36 / 0.35</td>
<td>0.36 / 0.35</td>
<td>0.35 / 0.35</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>0.147</td>
<td>0.147</td>
<td>1.09</td>
<td>0.193</td>
<td>0.193</td>
<td>4.42</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a bicycle</td>
<td>-0.635 (0.10)</td>
<td>-0.635 (0.10)</td>
<td>0*</td>
<td>-0.220 (0.22)</td>
<td>-0.220 (0.22)</td>
<td>0*</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a car</td>
<td>-1.27 (0.16)</td>
<td>-1.27 (0.16)</td>
<td>0*</td>
<td>0.0731 (0.92)</td>
<td>0.0731 (0.92)</td>
<td>0*</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a public transport season ticket</td>
<td>-0.340 (0.09)</td>
<td>-0.340 (0.09)</td>
<td>0*</td>
<td>-0.0634 (0.75)</td>
<td>-0.0634 (0.75)</td>
<td>0*</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Ride a bicycle</td>
<td>0.274 (0.42)</td>
<td>0.205 (0.44)</td>
<td>0*</td>
<td>0.0243 (0.84)</td>
<td>0.0254 (0.84)</td>
<td>0*</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car</td>
<td>-0.233 (0.58)</td>
<td>-0.175 (0.57)</td>
<td>0*</td>
<td>-0.0238 (0.86)</td>
<td>-0.0249 (0.86)</td>
<td>0*</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Shared modes</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Car passenger travel</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Taxi/minicab travel</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Public transport</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Walk</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Fixed holding costs in GBP per month</td>
<td>-0.00538</td>
<td>-0.00538</td>
<td>-0.01*^</td>
<td>-0.00936</td>
<td>-0.00936</td>
<td>-0.01*^</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>-0.116</td>
<td>-0.0869</td>
<td>-0.1*</td>
<td>-0.108</td>
<td>-0.113</td>
<td>-0.1*</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car)</td>
<td>-0.0152 (0.19)</td>
<td>-0.0114 (0.19)</td>
<td>-0.05*</td>
<td>-0.0472</td>
<td>-0.0494</td>
<td>-0.05*</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>-0.0319</td>
<td>-0.0240</td>
<td>-0.05*</td>
<td>-0.0475</td>
<td>-0.0496</td>
<td>-0.05*</td>
</tr>
<tr>
<td>Travel time in minutes (Shared modes)</td>
<td>-0.0763</td>
<td>-0.0573</td>
<td>-0.1*</td>
<td>-0.0978</td>
<td>-0.102</td>
<td>-0.1*</td>
</tr>
<tr>
<td>Travel time in minutes (Car passenger travel)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Taxi/minicab travel)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Public transport)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Walk)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Salience parameter (y) for 'escort' journey purpose</td>
<td>0.297 (0.14)</td>
<td>0.396 (0.16)</td>
<td>0.315 (0.20)</td>
<td>1.13</td>
<td>1.09</td>
<td>1.0*</td>
</tr>
<tr>
<td>y(Leisure)</td>
<td>0.599</td>
<td>0.797</td>
<td>0.745</td>
<td>0.897</td>
<td>0.858</td>
<td>1.0*</td>
</tr>
<tr>
<td>y(Shopping, personal business, and other)</td>
<td>1.00</td>
<td>1.33</td>
<td>1.10</td>
<td>1.00</td>
<td>0.957</td>
<td>1.0*</td>
</tr>
<tr>
<td>y(Social)</td>
<td>1.33</td>
<td>1.77</td>
<td>1.50</td>
<td>0.930</td>
<td>0.890</td>
<td>1.0*</td>
</tr>
<tr>
<td>y(Work and education)</td>
<td>1.34</td>
<td>1.79</td>
<td>1.51</td>
<td>1.00</td>
<td>0.961</td>
<td>1.0*</td>
</tr>
<tr>
<td>Logsum term ($A^{travel}$)</td>
<td>0.751</td>
<td>1.0*</td>
<td>1.0*</td>
<td>5.22</td>
<td>5.0*</td>
<td>5.02</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed. *Value is fixed

Table C.3: Obtained parameter values for runs with simulated datasets B & C
<table>
<thead>
<tr>
<th>Run #</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>E1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-1,344.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-966.2</td>
<td>-966.2</td>
<td>-972.9</td>
<td>-917.9</td>
<td>-1,333.4</td>
</tr>
<tr>
<td>$\rho^2$ / adjusted $\rho^2$</td>
<td>0.28 / 0.27</td>
<td>0.28 / 0.27</td>
<td>0.28 / 0.27</td>
<td>0.32 / 0.31</td>
<td>0.01 / 0.00</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>0.200</td>
<td>0.200</td>
<td>60.58</td>
<td>0.211</td>
<td>0.0626</td>
</tr>
</tbody>
</table>

| (Portfolio Choice level) ASC Own a bicycle | -1.77 | -1.77 | -2.10 | -2.11 | -0.184 (0.16) |
| (Portfolio Choice level) ASC Own a car | 0.144 (0.88) | 0.144 (0.88) | -0.900 | -1.38 (0.17) | -1.43 (0.06) |
| (Portfolio Choice level) ASC Own a public transport season ticket | 1.03 | 1.03 | 0.999 | 1.13 | -0.106 (0.66) |

| (Mode Choice level) ASC Ride a bicycle | 0.223 (0.22) | 0.255 (0.24) | 0* | 0.142 (0.45) | -1.29 (0.33) |
| (Mode Choice level) ASC Drive a car | -0.0647 (0.77) | -0.0740 (0.77) | 0* | 0.319 (0.16) | 0.882 (0.59) |
| (Mode Choice level) ASC Shared modes | 0* | 0* | 0* | 0* | 0* |
| (Mode Choice level) ASC Car passenger travel | -- | -- | -- | -- | -- |
| (Mode Choice level) ASC Taxi/minicab travel | -- | -- | -- | -- | -- |
| (Mode Choice level) ASC Public transport | -- | -- | -- | -- | -- |
| (Mode Choice level) ASC Walk | -- | -- | -- | -- | -- |
| Fixed holding costs in GBP per month | -0.00999 | -0.00999 | -0.01* | -0.0111 | 0.00398 (0.08) |
| Fare costs in GBP per journey | -0.0633 | -0.0725 | -0.1* | -0.0782 (0.12) | 0.125 (0.07) |
| Travel time in minutes (Drive a car) | -0.0339 | -0.0388 | -0.05* | -0.0501 | -0.289 (0.31) |
| Travel time in minutes (Ride a bicycle) | -0.0405 | -0.0464 | -0.05* | -0.0478 | 0.0344 |
| Travel time in minutes (Shared modes) | -0.0760 | -0.0869 | -0.1* | -0.101 | 0.0150 |
| Travel time in minutes (Car passenger travel) | -- | -- | -- | -- | -- |
| Travel time in minutes (Taxi/minicab travel) | -- | -- | -- | -- | -- |
| Travel time in minutes (Public transport) | -- | -- | -- | -- | -- |
| Travel time in minutes (Walk) | -- | -- | -- | -- | -- |
| Salience parameter ($\gamma$) for ‘escort’ journey purpose | 1.08 | 0.941 | 1.0* | 0.778 | 1.66 (0.18) |
| $\gamma$(leisure) | 0.995 | 0.869 | 1.0* | 1.06 | 0.642 (0.20) |
| $\gamma$(shopping, personal business, and other) | 1.00 | 0.874 | 1.0* | 1.00 | 0.587 (0.25) |
| $\gamma$(social) | 0.954 | 0.833 | 1.0* | 0.953 | 1.38e-32 (1.00) |
| $\gamma$(work and education) | 0.940 | 0.822 | 1.0* | 1.08 | 0.0826 (0.75) |
| Logsum term ($\lambda_{\text{travel}}$) | 1.14 | 1.0* | 1.0* | 1.01 | 1.0* |

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed. *Value is fixed

Table C.4: Obtained parameter values for runs with simulated datasets D & E
### Table C.5: Obtained parameter values for runs with simulated datasets F & G

<table>
<thead>
<tr>
<th>Run #</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-1344.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-710.4</td>
<td>-721.6</td>
<td>-723.9</td>
<td>-1032.7</td>
<td>-1026.3</td>
<td>-1026.3</td>
</tr>
<tr>
<td>$\rho^2$ / adjusted $\rho^2$</td>
<td>0.47 / 0.46</td>
<td>0.46 / 0.46</td>
<td>0.46 / 0.46</td>
<td>0.23 / 0.22</td>
<td>0.24 / 0.23</td>
<td>0.24 / 0.23</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>0.169</td>
<td>1.10</td>
<td>2.06</td>
<td>0.206</td>
<td>1.47e-12</td>
<td>1.08e-13</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a bicycle</td>
<td>-0.0906 (0.72)</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
<td>-0.261 (0.31)</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a car</td>
<td>-1.28 (0.21)</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
<td>0.141 (0.87)</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
</tr>
<tr>
<td>(Portfolio Choice level) ASC Own a public transport season ticket</td>
<td>-0.217 (0.27)</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
<td>0.131 (0.53)</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Ride a bicycle</td>
<td>-0.0348 (0.92)</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
<td>-2.24</td>
<td>-1.09</td>
<td>2.75</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car</td>
<td>0.256 (0.43)</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
<td>-1.17</td>
<td>0$^{*}$</td>
<td>3.83</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Shared modes</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
<td>0$^{*}$</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Car passenger travel</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>-24.5 (0.93)</td>
<td>-23.3 (0.56)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Taxi/minicab travel</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1.24</td>
<td>5.07 (1.00)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Public transport</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2.02</td>
<td>5.85</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Walk</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1.95</td>
<td>5.79</td>
</tr>
<tr>
<td>Fixed holding costs in GBP per month</td>
<td>-0.00713</td>
<td>-0.01$^{*}$</td>
<td>-0.01$^{*}$</td>
<td>-0.0106</td>
<td>-0.01$^{*}$</td>
<td>-0.01$^{*}$</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>-0.0853</td>
<td>-0.1$^{*}$</td>
<td>-0.1$^{*}$</td>
<td>-0.143</td>
<td>-0.1$^{*}$</td>
<td>-0.1$^{*}$</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car)</td>
<td>-0.0774</td>
<td>-0.0513</td>
<td>-0.0618</td>
<td>-0.0382</td>
<td>-0.05$^{*}$</td>
<td>-0.05$^{*}$</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>-0.122</td>
<td>-0.130</td>
<td>-0.110</td>
<td>-0.0467</td>
<td>-0.05$^{*}$</td>
<td>-0.05$^{*}$</td>
</tr>
<tr>
<td>Travel time in minutes (Shared modes)</td>
<td>-0.188</td>
<td>--</td>
<td>-0.176</td>
<td>-0.0930</td>
<td>-0.1$^{*}$</td>
<td>-0.1$^{*}$</td>
</tr>
<tr>
<td>Travel time in minutes (Car passenger travel)</td>
<td>--</td>
<td>-0.0989</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Taxi/minicab travel)</td>
<td>--</td>
<td>-0.653 (0.10)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Public transport)</td>
<td>--</td>
<td>-0.220 (0.08)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Walk)</td>
<td>--</td>
<td>-0.210</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Salience parameter ($y$) for ‘escort’ journey purpose</td>
<td>1.78</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
<td>0.794</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
</tr>
<tr>
<td>$y$(leisure)</td>
<td>1.04</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
<td>1.04</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
</tr>
<tr>
<td>$y$(shopping, personal business, and other)</td>
<td>1.00</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
<td>1.0$^{*}$</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
</tr>
<tr>
<td>$y$(social)</td>
<td>1.26</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
<td>0.992</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
</tr>
<tr>
<td>$y$(work and education)</td>
<td>0.620</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
<td>0.898</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
</tr>
<tr>
<td>Logsum term ($\lambda^{\text{travel}}$)</td>
<td>0.824</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
<td>0.713</td>
<td>1$^{*}$</td>
<td>1$^{*}$</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed. *Value is fixed
Appendix D: Further results from joint estimation of parameters, using both the E-NTS and AVATAR datasets

Note that the results of ‘portfolio’ choice run #s 3 and 6 are problematic. Run 3 reports a large negative mode-level ASC for ‘Versatility’, together with an equally-implausible positive marginal effect for time spent travelling in a ‘Versatility vehicle; this does not appear in Run 4 which is identical but for the addition of the logsum term. This issue then re-appears in Run 6, whilst Run 7 terminates successfully but reports that there is a value in the Hessian matrix which is so small as to approach the default criterion (1.0e-4) for declaring the termination to have been unsuccessful (i.e. that at least one parameter is almost completely unidentifiable). These estimation issues were not investigated in detail as they did not arise on the runs taken forward for the forecasting application or further analysis.

<table>
<thead>
<tr>
<th></th>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
<th>Run 4</th>
<th>Run 5</th>
<th>Run 6</th>
<th>Run 7</th>
<th>Run 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td>-2,447.7</td>
<td>-2,447.7</td>
<td>-2,447.7</td>
<td>-2,447.7</td>
<td>-2,447.7</td>
<td>-2,447.7</td>
<td>-2,447.7</td>
<td>-2,447.7</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-2,208.0</td>
<td>-2,134.0</td>
<td>-2,025.8</td>
<td>-1,967.7</td>
<td>-1,957.8</td>
<td>-1,976.2</td>
<td>-2,009.3</td>
<td>-1,957.8</td>
</tr>
<tr>
<td>$p^2$ / adjusted $p^2$</td>
<td>0.10 / 0.10</td>
<td>0.13 / 0.13</td>
<td>0.17 / 0.16</td>
<td>0.196 / 0.188</td>
<td>0.200 / 0.190</td>
<td>0.193 / 0.184</td>
<td>0.18 / 0.17</td>
<td>0.20 / 0.19</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>2.16</td>
<td>0.215</td>
<td>0.0164</td>
<td>0.328</td>
<td>0.058</td>
<td>3.24e-14</td>
<td>0.000327</td>
<td>0.32</td>
</tr>
<tr>
<td>Scale term (θ) for systematic utility (appears only for the E-NTS dataset)</td>
<td>1.24</td>
<td>2.50</td>
<td>0.815</td>
<td>1.16</td>
<td>1.30</td>
<td>1.40</td>
<td>1.35</td>
<td>1.30</td>
</tr>
</tbody>
</table>

(Portfolio Choice level) ASC Own a bicycle

-0.478 (-0.27) | -0.0476 (-0.07) | -0.553 (-0.27) | -0.423 (-0.27) | -0.343 (-0.27) | -0.295 (-0.27) | -0.289 (-0.27) | -0.343 (-0.27) |

(Portfolio Choice level) ASC Own a car

-0.429 | 3.63 | 7.35 | 6.86 | 6.50 | 5.86 | 7.00 | 6.50 |

(Portfolio Choice level) ASC Own a public transport season ticket

-0.331 | 0.877 | 1.45 | 1.14 | 1.01 | 0.915 | 0.536 (0.07) | 1.01 |

(Portfolio Choice level) ASC Subscribe to a car club

-3.22 | -3.18 | -7.85 | -4.35 | -4.82 | -3.73 | -3.12 | -4.82 |

(Portfolio Choice level) ASC Subscribe to ‘Versatility’

0.118 (0.60) | 0.233 (0.32) | 0.430 (0.41) | -0.410 (0.41) | -0.293 (0.51) | 0.369 (0.51) | -0.106 (0.81) | -0.293 (0.51) |

Continued on next page
Continued from previous page

<table>
<thead>
<tr>
<th></th>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
<th>Run 4</th>
<th>Run 5</th>
<th>Run 6</th>
<th>Run 7</th>
<th>Run 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Mode Choice level) ASC Ride a bicycle</td>
<td>--</td>
<td>--</td>
<td>-2.08</td>
<td>-0.0525</td>
<td>-0.431</td>
<td>-1.15</td>
<td>-6.55</td>
<td>-0.121</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.38)</td>
<td>(0.22)</td>
<td></td>
<td></td>
<td>(0.27)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car (E-NTS)</td>
<td>--</td>
<td>--</td>
<td>-0.753</td>
<td>0.0856</td>
<td>0.411</td>
<td>0.501</td>
<td>-3.51</td>
<td>0.115</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.11)</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td>(0.06)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car (AVATAR)</td>
<td>--</td>
<td>--</td>
<td>-1.28</td>
<td>0.146</td>
<td>1.23</td>
<td>1.83</td>
<td>-6.14</td>
<td>0.344</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.05)</td>
<td>(0.21)</td>
<td>(0.14)</td>
<td>(0.23)</td>
<td></td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a car club car</td>
<td>--</td>
<td>--</td>
<td>0.922</td>
<td>2.29</td>
<td>16.3</td>
<td>5.72</td>
<td>40.0</td>
<td>4.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.16)</td>
<td></td>
<td>(0.14)</td>
<td>(0.08)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Mode Choice level) ASC Drive a ‘Versatility’ car</td>
<td>--</td>
<td>--</td>
<td>-68.4</td>
<td>-0.251</td>
<td>-1.66</td>
<td>-28.1</td>
<td>-12.6</td>
<td>-0.467</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.14)</td>
<td>(0.07)</td>
<td>(0.41)</td>
<td>(0.10)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>(Mode Choice level) ASC Shared modes</td>
<td>--</td>
<td>--</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
<td>0*</td>
</tr>
<tr>
<td>Fixed holding costs in GBP per month</td>
<td>-- -0.0115</td>
<td>-- -0.0310</td>
<td>-0.0270</td>
<td>-0.0253</td>
<td>-0.0233</td>
<td>-0.0250</td>
<td>-0.0253</td>
<td></td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>--</td>
<td>--</td>
<td>-0.0716</td>
<td>-0.175</td>
<td>-1.13</td>
<td>-0.766</td>
<td>-7.16</td>
<td>-0.317</td>
</tr>
<tr>
<td>Travel time in minutes (Generic)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>-0.0236</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car, E-NTS)</td>
<td>--</td>
<td>--</td>
<td>-0.0208</td>
<td>-0.00470</td>
<td>-0.0251</td>
<td>--</td>
<td>-0.156</td>
<td>-0.00703</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car, AVATAR)</td>
<td>--</td>
<td>--</td>
<td>-0.0353</td>
<td>0.00473</td>
<td>0.0368</td>
<td>--</td>
<td>0.410</td>
<td>0.0103</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.07)</td>
<td>(0.30)</td>
<td>(0.27)</td>
<td></td>
<td></td>
<td>(0.28)</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>--</td>
<td>--</td>
<td>-0.0374</td>
<td>-0.0108</td>
<td>-0.0567</td>
<td>--</td>
<td>-0.281</td>
<td>-0.0199</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car club car)</td>
<td>--</td>
<td>--</td>
<td>-0.0586</td>
<td>-0.0687</td>
<td>-0.454</td>
<td>--</td>
<td>-1.42</td>
<td>-0.127</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a ‘Versatility’ car)</td>
<td>--</td>
<td>--</td>
<td>0.525</td>
<td>0.0245</td>
<td>0.169</td>
<td>--</td>
<td>1.06</td>
<td>0.0475</td>
</tr>
<tr>
<td>Travel time in minutes (Shared modes)</td>
<td>--</td>
<td>--</td>
<td>-0.0237</td>
<td>-0.00508</td>
<td>-0.0282</td>
<td>--</td>
<td>-0.193</td>
<td>-0.00792</td>
</tr>
<tr>
<td>Scale term (V) for all y terms in the AVATAR dataset only</td>
<td>--</td>
<td>--</td>
<td>3.10</td>
<td>1.96</td>
<td>1.67</td>
<td>1.55</td>
<td>1.53</td>
<td>1.28</td>
</tr>
<tr>
<td>y for ‘escort’ journey purpose</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.101</td>
<td>0.117</td>
<td>1.53</td>
<td>0.468</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.11)</td>
<td>(0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y(leisure)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.127</td>
<td>0.140</td>
<td>0.554</td>
<td>0.588</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.11)</td>
<td>(0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y(shopping, personal business, and other)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.170</td>
<td>0.179</td>
<td>0.521</td>
<td>0.787</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.11)</td>
<td>(0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y(social)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.100</td>
<td>0.106</td>
<td>0.747</td>
<td>0.466</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.11)</td>
<td>(0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y(work and education)</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>1*</td>
<td>0.216</td>
<td>0.229</td>
<td>1*</td>
<td>1*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.11)</td>
<td>(0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsum term (A^{travel})</td>
<td>--</td>
<td>--</td>
<td>1*</td>
<td>5.67</td>
<td>1.0*</td>
<td>1.0*</td>
<td>0.199</td>
<td>3.57</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed.

* Value is fixed

b Parameter is output from the estimation process but not used in forecasting applications. See Section 7.5.2.

Table D.1: Results from parameter estimation of ‘portfolio’ choice using the combined E-NTS and AVATAR survey datasets
<table>
<thead>
<tr>
<th></th>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
<th>Run 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null log-likelihood</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-24,447.2</td>
<td>-18,753.4</td>
<td>-17,103.8</td>
<td>-17,356.6</td>
</tr>
<tr>
<td>$p^2$ / adjusted $p^2$</td>
<td>0.23 / 0.23</td>
<td>0.30 / 0.30</td>
<td>0.29 / 0.29</td>
<td>0.25 / 0.25</td>
</tr>
<tr>
<td>Smallest singular value in Hessian matrix</td>
<td>2.95</td>
<td>0.19</td>
<td>0.218</td>
<td>1.18</td>
</tr>
<tr>
<td>Scale term ($\eta$) for systematic utility (appears only for the E-NTS dataset)</td>
<td>2.21</td>
<td>2.56</td>
<td>2.26</td>
<td>0.851</td>
</tr>
<tr>
<td>ASC Drive a car</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>ASC Ride a bicycle</td>
<td>-1.37</td>
<td>-1.37</td>
<td>-1.26</td>
<td>-2.31</td>
</tr>
<tr>
<td>ASC Drive a car club car</td>
<td>-1.33</td>
<td>-0.593 (0.35)</td>
<td>-0.734 (0.10)</td>
<td>-5.88</td>
</tr>
<tr>
<td>ASC Drive a ‘Versatility’ car</td>
<td>-0.774</td>
<td>-0.734</td>
<td>-0.500</td>
<td>-3.91</td>
</tr>
<tr>
<td>ASC Car passenger travel</td>
<td>-0.895</td>
<td>-0.828</td>
<td>-0.859</td>
<td>-0.684</td>
</tr>
<tr>
<td>ASC Taxi/Minicab travel</td>
<td>-2.16</td>
<td>-1.64</td>
<td>-1.87</td>
<td>-3.29</td>
</tr>
<tr>
<td>ASC Public transport</td>
<td>-1.11</td>
<td>-0.849</td>
<td>-0.696</td>
<td>-2.73</td>
</tr>
<tr>
<td>ASC Walk</td>
<td>-1.13</td>
<td>-0.537</td>
<td>-0.551</td>
<td>0.0106 (0.92)</td>
</tr>
<tr>
<td>Fare costs in GBP per journey</td>
<td>--</td>
<td>-0.0617</td>
<td>-0.0342</td>
<td>-0.105</td>
</tr>
<tr>
<td>Travel time in minutes (Generic)</td>
<td>--</td>
<td>--</td>
<td>-0.0119</td>
<td>--</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car, E-NTS)</td>
<td>--</td>
<td>-0.0179</td>
<td>--</td>
<td>0.00944</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car, AVATAR)</td>
<td>--</td>
<td>0.00382$^b$</td>
<td>--</td>
<td>-0.150</td>
</tr>
<tr>
<td>Travel time in minutes (Ride a bicycle)</td>
<td>--</td>
<td>-0.0166</td>
<td>--</td>
<td>-0.0187</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a car club SDCS car)</td>
<td>--</td>
<td>-0.00626 (0.70)</td>
<td>--</td>
<td>0.0705</td>
</tr>
<tr>
<td>Travel time in minutes (Drive a ‘Versatility’ car)</td>
<td>--</td>
<td>0.00172 (0.86)</td>
<td>--</td>
<td>0.105</td>
</tr>
<tr>
<td>Travel time in minutes (Car passenger travel)</td>
<td>--</td>
<td>-0.0155</td>
<td>--</td>
<td>0.0117</td>
</tr>
<tr>
<td>Travel time in minutes (Taxi/Minicab travel)</td>
<td>--</td>
<td>-0.0121</td>
<td>--</td>
<td>0.0209</td>
</tr>
<tr>
<td>Travel time in minutes (Public transport)</td>
<td>--</td>
<td>-0.00919</td>
<td>--</td>
<td>0.0152</td>
</tr>
<tr>
<td>Travel time in minutes (Walk)</td>
<td>--</td>
<td>-0.0132</td>
<td>--</td>
<td>-0.0314</td>
</tr>
</tbody>
</table>

NB: Values in parentheses are p-values; values smaller than 0.05 are suppressed.

$^a$ Value is fixed

$^b$ Parameter is output from the estimation process but not used in forecasting applications. See Section 7.5.2.

Table D.2: Results from parameter estimation of mode choice using the combined E-NTS and AVATAR survey datasets
Appendix E: Sample gaming-simulation survey instrument package

(Rest of this page intentionally left blank)
Good ____________.

First let me thank you for agreeing to participate in this study of how and why Londoners’ travel. My name is ________________, and I’ll be your survey coordinator. Your participation will be in several parts:

- First, this initial meeting
- Then, your activity-travel diaries, which we will go through in a minute.
- And finally, a meeting at your home, scheduled a few days after I collect the diaries.

First, I’d like to ask about your postcode, but not the whole postcode. Please just say the first part of the postcode and the first number of the second part.

**Postcode**

Now, we’ll talk about your family members living in this home, starting with you.

**Person 1 (An adult)**

<table>
<thead>
<tr>
<th>Name:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Year Born &amp; Gender:</td>
<td></td>
</tr>
<tr>
<td>Relationship:</td>
<td>SELF</td>
</tr>
<tr>
<td>If a couple, how long have you been a couple?</td>
<td></td>
</tr>
<tr>
<td>Employment:</td>
<td></td>
</tr>
<tr>
<td>Address of Employer or School:</td>
<td></td>
</tr>
<tr>
<td>Thinking about your neighbourhood, how would you describe it?</td>
<td></td>
</tr>
<tr>
<td>What is it like to park a car in your neighbourhood? Is there on-street parking? Do many homes have driveways or garages? Do you need a residential parking permit?</td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Answer</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>What kinds of services are there in the neighbourhood? (convenience store, bank branch, high street shops, etc.)</td>
<td></td>
</tr>
<tr>
<td>Is there a nearby tube station? Bus stop? National Rail station?</td>
<td></td>
</tr>
<tr>
<td>Driving Licence? When did you first receive a driving license?</td>
<td></td>
</tr>
<tr>
<td>Do you have an Oyster card, Freedom Card or other public transport pass?</td>
<td></td>
</tr>
<tr>
<td>Highest Level of Education:</td>
<td></td>
</tr>
<tr>
<td>How long have you lived at this place of residence? Have you lived in other places or cities? When you lived there, how did you get around?</td>
<td></td>
</tr>
<tr>
<td>Have you owned a car in the past? (If yes, what were the circumstances)</td>
<td></td>
</tr>
<tr>
<td>Have you had transit passes in the past? (If yes, what were the circumstances)</td>
<td></td>
</tr>
</tbody>
</table>

**Person #2**

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name:</td>
<td></td>
</tr>
<tr>
<td>Age &amp; Gender:</td>
<td></td>
</tr>
<tr>
<td>Relationship:</td>
<td></td>
</tr>
<tr>
<td>Employment:</td>
<td></td>
</tr>
<tr>
<td>Address of Employer or School:</td>
<td></td>
</tr>
<tr>
<td>Driving Licence?</td>
<td></td>
</tr>
<tr>
<td>Do you have an Oyster card, Freedom Card or other public transport pass?</td>
<td></td>
</tr>
<tr>
<td>Highest Level of Education:</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>How long have you lived at this place of residence? Have you lived in other places or cities? When you lived there, how did you get around?</td>
<td></td>
</tr>
<tr>
<td>Have you owned a car in the past? (If yes, what were the circumstances)</td>
<td></td>
</tr>
<tr>
<td>Have you had transit passes in the past? (If yes, what were the circumstances)</td>
<td></td>
</tr>
</tbody>
</table>

### Person #3

<table>
<thead>
<tr>
<th>Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &amp; Gender:</td>
</tr>
<tr>
<td>Relationship:</td>
</tr>
<tr>
<td>Employment:</td>
</tr>
<tr>
<td>Address of Employer or School:</td>
</tr>
<tr>
<td>Driving Licence?</td>
</tr>
<tr>
<td>Do you have an Oyster card, Freedom Card or other public transport pass?</td>
</tr>
<tr>
<td>Highest Level of Education:</td>
</tr>
<tr>
<td>How long have you lived at this place of residence? Have you lived in other places or cities? When you lived there, how did you get around?</td>
</tr>
<tr>
<td>Have you owned a car in the past? (If yes, what were the circumstances)</td>
</tr>
<tr>
<td>Have you had transit passes in the past? (If yes, what were the circumstances)</td>
</tr>
</tbody>
</table>
### Person #4

<table>
<thead>
<tr>
<th>Name:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &amp; Gender:</td>
<td></td>
</tr>
<tr>
<td>Relationship:</td>
<td></td>
</tr>
<tr>
<td>Employment:</td>
<td></td>
</tr>
<tr>
<td>Address of Employer or School:</td>
<td></td>
</tr>
<tr>
<td>Driving Licence?</td>
<td></td>
</tr>
<tr>
<td>Do you have an Oyster card, Freedom Card or other public transport pass?</td>
<td></td>
</tr>
<tr>
<td>Highest Level of Education:</td>
<td></td>
</tr>
<tr>
<td>How long have you lived at this place of residence? Have you lived in other places or cities? When you lived there, how did you get around?</td>
<td></td>
</tr>
<tr>
<td>Have you owned a car in the past? (If yes, what were the circumstances)</td>
<td></td>
</tr>
<tr>
<td>Have you had transit passes in the past? (If yes, what were the circumstances)</td>
<td></td>
</tr>
</tbody>
</table>

### Person #5

<table>
<thead>
<tr>
<th>Name:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &amp; Gender:</td>
<td></td>
</tr>
<tr>
<td>Relationship:</td>
<td></td>
</tr>
<tr>
<td>Employment:</td>
<td></td>
</tr>
<tr>
<td>Address of Employer or School:</td>
<td></td>
</tr>
<tr>
<td>Driving Licence?</td>
<td></td>
</tr>
<tr>
<td>Do you have an Oyster card, Freedom Card or other public transport pass?</td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Answer</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>How long have you lived at this place of residence?</td>
<td></td>
</tr>
<tr>
<td>Have you lived in other places or cities? When you lived there, how did you get around?</td>
<td></td>
</tr>
<tr>
<td>Have you owned a car in the past? (If yes, what were the circumstances)</td>
<td></td>
</tr>
<tr>
<td>Have you had transit passes in the past? (If yes, what were the circumstances)</td>
<td></td>
</tr>
</tbody>
</table>

**Person #6**

<table>
<thead>
<tr>
<th>Name:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &amp; Gender:</td>
<td></td>
</tr>
<tr>
<td>Relationship:</td>
<td></td>
</tr>
<tr>
<td>Employment:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Address of Employer or School:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Driving Licence?</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Do you have an Oyster card, Freedom Card or other public transport pass?</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Highest Level of Education:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>How long have you lived at this place of residence?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have you lived in other places or cities? When you lived there, how did you get around?</td>
</tr>
<tr>
<td>Have you owned a car in the past? (If yes, what were the circumstances)</td>
</tr>
<tr>
<td>Have you had transit passes in the past? (If yes, what were the circumstances)</td>
</tr>
</tbody>
</table>
Do you own or rent your home?  

Is your home a:  (house)  (flat or maisonette)  (room or rooms)  

Do you own any cars or trucks? ________ (IF YES) Now I'll ask you some questions about them.

**Car or truck #1**

<table>
<thead>
<tr>
<th>Model Year:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Make and Model:</td>
<td></td>
</tr>
<tr>
<td>Who drives this vehicle the most?</td>
<td></td>
</tr>
<tr>
<td>Does anyone else drive it regularly?</td>
<td></td>
</tr>
<tr>
<td>Where is it parked?</td>
<td>Day:</td>
</tr>
</tbody>
</table>

**Car or truck #2**

<table>
<thead>
<tr>
<th>Model Year:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Make and Model:</td>
<td></td>
</tr>
<tr>
<td>Who drives this vehicle the most?</td>
<td></td>
</tr>
<tr>
<td>Does anyone else drive it regularly?</td>
<td></td>
</tr>
<tr>
<td>Where is it parked?</td>
<td>Day:</td>
</tr>
</tbody>
</table>

**Car or truck #3**

<table>
<thead>
<tr>
<th>Model Year:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Make and Model:</td>
<td></td>
</tr>
<tr>
<td>Who drives this vehicle the most?</td>
<td></td>
</tr>
<tr>
<td>Does anyone else drive it regularly?</td>
<td></td>
</tr>
<tr>
<td>Where is it parked?</td>
<td>Day:</td>
</tr>
</tbody>
</table>
Now, I’d like to ask you to complete a questionnaire on travel and lifestyle questions. Please hand it back when you are finished.

OK, now we’ll go through the activity-travel diaries. First, I’ll explain how to complete them, and then we’ll do a practice example. They’re meant to help you record the activities you do and the trips you make during your survey week. There’s one sheet for each day.

First, I’ll draw your attention to the day and date shown at the top. In your diaries, they’re all filled in and in order. We put an extra sheet at the back of the diary if you make a mistake.

Now we’ll go over the section on the left-hand side, which will record the activities you do during the day. There are letter codes for different kinds of activities just to the right. You’ll start by writing the code of where your day started, normally at your home. Then you’ll draw a horizontal line when you leave home to do your first activity of the day. Always put a horizontal line when you begin a journey to a new activity. Each time that you leave to go somewhere else during the day, just put the code of the new activity and a horizontal line.

Now, if you drop off or pick up someone else, please use the code letter ‘B’. and write it in at the appropriate time. Also, if none of these codes makes sense, code the letter ‘O’ and make a note here <<point to location>>.

Now, on the right hand side you’ll record your day’s travel. Let’s first go over when to record a journey. We’ll classify a journey as a one-way trip. So, your commute to work would be a journey. If you went to the store to pick up some milk and then returned home, that would be two separate journeys.

If you drop someone off, please record two separate journeys because the number of people travelling will change. It’s important to remember this.

The first thing to record is where the journey begins. If it’s your home, workplace, or school, just write ‘home’, ‘work’, or ‘school’. For any other places, note the address and postcode, or if you don’t know the address just write in the street name and place name.

Next, you’ll write the time the journey begins and ends. It doesn’t have to be exact, as close as you can recall is fine. The you’ll write in if anyone else travelled with you.

The method of travel is next. You’ll see at the bottom listings of different ways to travel. The first, second, and third spaces are for if you use more than one way to make a journey, maybe something like a bus then a tube. If you drive a car or a passenger in a car, write in the model if it’s one of your family cars, or if it’s someone else’s car just write in whose car it is. So, for instance, if you rode in a friend’s car, just write in ‘Psgr, friend’s car’

Then the last thing to note for each journey is how long you had to walk at the beginning or end. So if you walked to a tube station, or to your parked car, write in about how many minutes you walked at each end of the journey.
Now, when you move on to record the next journey, it will usually begin at the same place as your last journey ended. So in that case, just put an arrow <<draw in on sample>> and you don’t need to write the address again. Then you’ll proceed with the rest of the journey information, starting with where it ended.

Now, I’ll hand you a sample diary page and we’ll go through an example together. Would anyone like to volunteer to record their day’s activities and travel?

<<Work through example. If no volunteer, use coordinator’s day as sample>>

<<If any young children>> Please help <<names>> to complete their diaries.

OK, that’s all for setting you up with the activity-travel diaries. Please start recording tomorrow. Feel free to fold up your diary if you’d like to keep it in your purse or briefcase. Do you have a good understanding of how to complete your diaries?

My contact information is on the front page if you have any questions at all whilst the week progresses. I will call you on <<date on front of diaries>> to check on how you are faring and to make the arrangement for collecting them. When I come to collect them, I’ll have a quick follow-up questionnaire which should take you no more than 5 minutes. We’ll hold the last meeting, which will be about an hour and a half, on <<arranged date & time>>. It is important that all family members are available for the follow-up discussion.
These questions ask about how frequently you do various activities. Please tick the box indicating how often you do each of them. Ask your survey coordinator for help if you have any questions.

1. **How you get around**

<table>
<thead>
<tr>
<th>In general, how frequently do you...</th>
<th>Never</th>
<th>Almost Never</th>
<th>Several Times a Year</th>
<th>About once a Month</th>
<th>About once a Week</th>
<th>Almost Every Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk more than 5 minutes to get somewhere</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drive your own car</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drive another household member’s car</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrow a car from a friend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Give someone a ride</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ride as a passenger in a car</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hire a rental car</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use a black cab or minicab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Take a bus or coach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ride on the tube</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Take an overground train</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fly commercially</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## 2. Personal Activities

**In general, how frequently do you...**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Never</th>
<th>Almost Never</th>
<th>Several Times a Year</th>
<th>About once a Month</th>
<th>About once a Week</th>
<th>Almost Every Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eat lunch out</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Eat dinner out</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Regular household cleaning or chores</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>DIY</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Watch broadcast television</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Watch videos at your home</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Shop for groceries</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Shop for other things</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Have friends round to visit</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Meet friends out of your home</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Visit relatives</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Go to the cinema, theatre, live sport, or music concerts</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Go to the bank</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Go to a government office</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Visit a doctor (for yourself or to escort someone)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Do volunteer work</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Actively exercise (in a gym or outdoors)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Travel as part of your job, not commuting (within Greater London)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Work from home</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
### 3. Electronic Communications

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Almost Never</th>
<th>About once a Month</th>
<th>About once a Week</th>
<th>Almost Every Day</th>
<th>Several Times a Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talk on a mobile phone</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Send text messages</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Use email</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Seek information online (such as clubs, products, services, or other web searching)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Get news online</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Online banking or purchase something online</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Online social networking</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

### 4. Activities Outside Greater London

**Do you own a second home outside of London?**

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>(circle one)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Almost Never</th>
<th>Several Times a Year</th>
<th>About once a Month</th>
<th>More than once a month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Travel outside London for work</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Make day trips outside London for leisure</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Take overnight holidays (including weekends) in the UK</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Take holidays outside the UK</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Travel outside London on personal business, or for other reasons (neither business or leisure)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

**Thank you**

Please hand back to your survey coordinator
This sheet will record the activities you do and the trips you make. Please use a separate sheet for each day of your survey week.

### Your Day’s Activities

What did you do through the course of the day?  
(Please select amongst the categories at right. Include travel to each activity within each block of time. Or, write in a brief description of the activity.)

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>A) Any activity at home</td>
</tr>
<tr>
<td>0100</td>
<td>B) Dropping off or picking up someone</td>
</tr>
<tr>
<td>0200</td>
<td>C) At work</td>
</tr>
<tr>
<td>0300</td>
<td>D) Working (off-site location)</td>
</tr>
<tr>
<td>0400</td>
<td>E) At school</td>
</tr>
<tr>
<td>0500</td>
<td>F) Grocery shopping</td>
</tr>
<tr>
<td>0600</td>
<td>G) Shopping (not groceries)</td>
</tr>
<tr>
<td>0700</td>
<td>H) Personal business (medical, etc.)</td>
</tr>
<tr>
<td>0800</td>
<td>I) Eating out</td>
</tr>
<tr>
<td>0900</td>
<td>J) Visiting friends at their home</td>
</tr>
<tr>
<td>1000</td>
<td>K) Visiting relatives at their home</td>
</tr>
<tr>
<td>1100</td>
<td>L) Entertainment or public activities</td>
</tr>
<tr>
<td>1200</td>
<td>M) Play sport or exercise</td>
</tr>
<tr>
<td>1300</td>
<td>N) At hotel (on holiday or business)</td>
</tr>
<tr>
<td>1400</td>
<td>O) Some other activity, not at home (make notes below)</td>
</tr>
</tbody>
</table>

### Your Day’s Travel

<table>
<thead>
<tr>
<th>Trip #</th>
<th>Where did the journey start?</th>
<th>Where did the journey end?</th>
<th>When did you leave and arrive?</th>
<th>Who travelled with you?</th>
<th>Which method of travel?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>Left: ______________________</td>
<td>2nd: _____________________</td>
<td>Walk at start of trip (minutes):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>Left: ______________________</td>
<td>2nd: _____________________</td>
<td>Walk at end of trip (minutes):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd</td>
<td>Left: ______________________</td>
<td>2nd: _____________________</td>
<td>Start: ______________________</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th</td>
<td>Left: ______________________</td>
<td>2nd: _____________________</td>
<td>End: _______________________</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5th</td>
<td>Left: ______________________</td>
<td>2nd: _____________________</td>
<td>Start: ______________________</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6th</td>
<td>Left: ______________________</td>
<td>2nd: _____________________</td>
<td>End: _______________________</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Example

- **Drive Car:** (Indicate model if a household car, e.g. Drive, Renault. Otherwise indicate the owner, e.g. friend, relative, rental car, etc.)
- **Passenger in Car:** (Indicate model if a household car, e.g. PSGR, Renault. Otherwise indicate the owner, e.g. friend, relative, rental car, etc.)
- **Black Cab or Minicab**
- **Bus or Coach**
- **Tube**
- **Overground Train**
- **Fly**
- **Any Other Way (Write in)**
Scenario ‘A’

Inspections have found a number of cracked rails throughout the National Rail network.

Until the system can be fully checked and repaired, strict speed restrictions are in place. Travel times are roughly doubled.

This is expected to last at least 3 months. In order to pay for the emergency work, all National Rail tickets now carry a £10 surcharge. (£20 for return tickets)
Scenario ‘B’

As previously, the National rail network is under strict speed restrictions.

However, the economy has deteriorated. XXXX’s firm has cancelled the company car programme as an austerity measure. Effective immediately, pool cars will be available for business use from a garage in Central London.

All staff are to be provided £35 per week (taxable) as partial compensation for the change.
Scenario ‘C’

As previously, the National rail network is under strict speed restrictions, and XXXX’s firm has switched to a pool car system.

On Monday afternoon, XXXXX receives a phone to discuss an attractive job offer, from a firm under pressure to improve its environmental record. They would like you to come for an interview the next day (Tuesday afternoon). If selected, they will pay half-a-year's salary for a six-to-8-week specialised placement, to start as soon as practical. They are located in a part of XXXXX with poor public transport connections
Car Club

A car club provides hire cars in your neighbourhood that you can use without paperwork or waiting. They are parked in spaces within a short walk of your home. You need to reserve a car ahead of time. You can use your mobile phone or the internet to make a reservation, at any time before you want to use it.

It costs about £5 per hour, and you are charged from the time you pick the car up until you bring it back. The car club covers MOT, petrol, insurance, and all other costs. You are billed each month for your use, and pay £50 per year for membership.

You unlock the car doors by touching a smartcard to the windscreen. The smartcard also starts the car. Each time you use it, you must bring the car back to where you started.

Other car club members use the car at different times, so there is a 1 in 20 chance that it isn’t available to reserve when you want.
Scenario ‘D’

As previously, the National rail network is under strict speed restrictions, XXXXX’s firm has switched to a pool car system, and XXXXX has a job interview at a site in Surrey.

The ‘South-Car’ car club has opened in the Southfields area.

One of the service points is located in your neighbourhood within a short walk.
Scenario ‘A’

Inspections have found a number of cracked rails throughout the National Rail network.

Until the system can be fully checked and repaired, strict speed restrictions are in place. Travel times are roughly doubled.

This is expected to last at least 3 months. In order to pay for the emergency work, all National Rail tickets now carry a £10 surcharge. (£20 for return tickets)

On a scale from zero to ten, how would your week have been in this set of circumstances?  
(Please mark on the scale below)

0 1 2 3 4 5 6 7 8 9 10
Substantially worse About the same as my week was Substantially better
Scenario ‘B’

As previously, the National rail network is under strict speed restrictions.

However, the economy has deteriorated. XXXX’s firm has cancelled the company car programme as an austerity measure. Effective immediately, pool cars will be available for business use from a garage in Central London.

All staff are to be provided £35 per week (taxable) as partial compensation for the change.

On a scale from zero to ten, how would your week have been in this set of circumstances?  (Please mark on the scale below)

0    1    2    3    4    5    6    7    8    9    10
Substantially worse  About the same as my week was  Substantially better
Scenario ‘C’

As previously, the National rail network is under strict speed restrictions, and XXXXX’s firm has switched to a pool car system.

On Monday afternoon, XXXXXX receives a phone to discuss an attractive job offer, from a firm under pressure to improve its environmental record. They would like you to come for an interview the next day (Tuesday afternoon). If selected, they will pay half-a-year's salary for a six-to-8-week specialised placement, to start as soon as practical. They are located in a part of XXXXXX with poor public transport connections.

On a scale from zero to ten, how would your week have been in this set of circumstances? (Please mark on the scale below)

0 Substantially worse
1 About the same as my week was
2 Substantially better

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substantially worse</td>
<td>About the same as my week was</td>
<td>Substantially better</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Scenario ‘D’

As previously, the National rail network is under strict speed restrictions, XXXXX’s firm has switched to a pool car system, and XXXXX has a job interview at a site in XXXXX.

The ‘South-Car’ car club has opened in the Southfields area.

One of the service points is located in your neighbourhood within a short walk.

**On a scale from zero to ten, how would your week have been in this set of circumstances?**

(Please mark on the scale below)

<table>
<thead>
<tr>
<th>Scale</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Substantially worse</td>
</tr>
<tr>
<td>1</td>
<td>About the same as my week was</td>
</tr>
<tr>
<td>2</td>
<td>Substantially better</td>
</tr>
</tbody>
</table>
Appendix F: Sample AVATAR survey instrument package

How should Jane get around? How should Joe get around?

Let's start by asking a bit about you

Next
OK, before we begin let me introduce you to Jane

Jane lives with her partner, son (age 7) and an adult family member. They are moving to London soon, to a place in Outer London

She is in her 40s, and works in an office

Her flat search has narrowed to four neighbourhoods in Outer London

She has asked for your advice:

**She wants to know how you would get around from each place that she is considering**
We'll look at some of the things she will do in a typical week, and how she could get to them.

This first one is just practice -- I'll show you how each bit of the screen works, one section at a time.

Just ask if you have any questions. I'll tell you when we're about to start for real.

Next
Jane lives with her partner, son (age 7) and an adult family member. They are moving to London soon, to a place in Outer London.

She is in her 40s, and works in an office.

Some things Jane will do in a typical week...

A car must be taxed, insured and maintained.

Jane does not expect to have a parking space, so she would park her car on the street at night.

Motoring organisations say that owning a mid-sized car in London costs around £4,000 per year, excluding petrol and parking.

What it means...

Travel & spending:
- by public transport
- by taxi
- walking

Weekly travel & spending:
- £30
- £10
- £50

Time walking:
- 50 mins
- 20 mins

I would do this if I were Jane
### Should Jane...

**A car club provides by-the-hour hire cars in Jane’s neighbourhood that can be used without paperwork or waiting.**

Jane would need to reserve a car ahead of time.

Jane can use her mobile or the internet to make a reservation.

---

#### Weekly travel & spending:

<table>
<thead>
<tr>
<th>Mode</th>
<th>Cost</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public transport fare</td>
<td>£10</td>
<td>50 mins</td>
</tr>
<tr>
<td>CBD fare</td>
<td>£50</td>
<td></td>
</tr>
<tr>
<td>Time on public transport</td>
<td>90 mins</td>
<td></td>
</tr>
<tr>
<td>Time walking</td>
<td>50 mins</td>
<td></td>
</tr>
</tbody>
</table>

**I would do this if I were Jane**

---

**What it means...**

It costs £5 per hour, and Jane would be charged from the time she picks the car up until she brings it back.

Jane would be billed each month for her use, and pay £50 per year for membership.

Jane would unlock the car doors by touching a smart card to the windscreen. Each time she uses it, she must bring the car back to where she started.

---

**I would do this if I were Jane**
### Should Jane...

**Some things Jane will do in a typical week...**

<table>
<thead>
<tr>
<th>Take a Taxi</th>
<th>Drive her own car</th>
<th>Drive a car club</th>
<th>Walk</th>
<th>Take public transport</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
</tr>
</tbody>
</table>

**What it means...**

**Petrol** is included in the £5 hourly fee, but not parking.

Other members use the cars at different times, so there is a 1 in 20 chance that a car isn’t available when Jane wants.

Jane is responsible to check for damage before taking a car club car.

**Click here to return**

### Her weekly travel & spending:

**Public transport fares:** £30

**Taxi fares:** £10

Time on public transport: 9h 10m

Time walking: 50 mins

Time as taxi passenger: 20 mins

**I would do this if I were Jane**

### Should Jane...

**Some things Jane will do in a typical week...**

<table>
<thead>
<tr>
<th>Take a Taxi</th>
<th>Drive her own car</th>
<th>Drive a car club</th>
<th>Walk</th>
<th>Take public transport</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
<td>10 mins</td>
</tr>
</tbody>
</table>

**What it means...**

A public transport season ticket lets Jane use public transport as much as she likes without paying per-journey fares.

A season ticket would cost Jane £1.50 per month.

**Click here to return**

**A public transport season ticket lets Jane use public transport as much as she likes without paying per-journey fares. A season ticket would cost Jane £1.50 per month.**

**Click here to return**
S Le Vine thesis

Some things Jane will do in a typical week...

<table>
<thead>
<tr>
<th>Should Jane...</th>
<th>TAKE A TAXI</th>
<th>DRIVE HER | MIN</th>
<th>CAR CLUB</th>
<th>WALK</th>
<th>TAKE PUBLIC</th>
<th>CYCLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some things Jane will do in a typical week...</td>
<td>20 min</td>
<td>15 min</td>
<td>3 min</td>
<td>4 min</td>
<td>3 min</td>
<td>15 min</td>
</tr>
</tbody>
</table>

Jane will need accessories along with a bicycle, such as bike lights and a helmet.

She expects to have a place off-street to store the bicycle at night.

What it means...

Her weekly travel & spending:

- Public transport fares: £70
- Taxi fares: £10
- Time on public transport: 9h 10m
- Time walking: 30 mins
- Time as taxi passenger: 20 mins

I would do this if I were Jane

Jane can get free advice from cycling organisations on safe cycling and recommended routes for cycling in London.

Cycling organisations say that it costs about £150 per year to own and maintain a bicycle, including a helmet, accessories, and maintenance.

What it means...

Her weekly travel & spending:

- Public transport fares: £80
- Taxi fares: £20
- Time on public transport: 9h 10m
- Time walking: 30 mins
- Time as taxi passenger: 20 mins

I would do this if I were Jane
That was practice

Now, are you ready to get started?

Just ask if you have any questions

No, show me again  Yes

Let's sort out how Jane could get around in the FIRST neighbourhood she's considering

Jane

Next
Some things Jane will do in a typical week...

- **Buying a car**: 80p per week
  - 15 mins each way
  - 80p each way
- **Taking a public transport ticket**: £1.50 per week
  - 30 mins each way
  - 70p each way
- **Walking**: 60 mins each way
  - Free

**Her weekly travel & spending**
- Total weekly cost: £1.60
- Travel time: 70 mins
- Travel distance: 10 km

I would do this if I were Jane.

Now, let’s sort out how Jane could get around in the SECOND neighbourhood she’s considering.
Some things Jane will do in a typical week...

Some things Jane will do in a typical week...

<table>
<thead>
<tr>
<th>Should Jane...</th>
<th>What it means...</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TAKE A TAXI</strong></td>
<td>She will: join a car club. <strong>FULL</strong> year.</td>
</tr>
<tr>
<td><strong>DRIVE A CAR</strong></td>
<td>10 min. walk. £6.25 in total.</td>
</tr>
<tr>
<td><strong>WALK</strong></td>
<td>10 min. walk. £0.00 in total.</td>
</tr>
<tr>
<td><strong>TAKE PUBLIC TRANSPORT</strong></td>
<td>10 min. walk. £0.00 in total.</td>
</tr>
<tr>
<td><strong>CYCLE</strong></td>
<td>10 min. walk. £0.00 in total.</td>
</tr>
</tbody>
</table>

**Her weekly travel & spending:**

- **Period/packaging costs:** £5
- **Public transport fares:** £3
- **Car club use fees:** £10
- **Taxi fares:** £15
- **Time walking:** 1hr 30min
- **Time on public transport:** 1hr 20min
- **Time walking:** 5hr 20min

**I would do this if I were Jane.**

Now, let’s sort out how Jane could get around in the **THIRD** neighbourhood she’s considering.

![Diagram of buildings and a person named Jane]

Next
But now things are a bit different...

A large car hire company has recently introduced a service they call ‘VERSATILITY’

VERSATILITY cars are parked on street corners around London.

You walk up to a car, with or without a reservation, and then drop it off at ANY ON-STREET PARKING SPACE IN LONDON at the end of your journey.

You only pay until you drop it off at the end of your journey — 20p A MINUTE (PARKING IS FREE)
VERSATILITY isn't perfect...

It costs £10 a month to subscribe.

At busy times it can be difficult to find a car nearby, and you may have to wait.

Every so often you might find a car you would like to use in an unclean state or low on fuel.

The cars can't be taken outside London.

Now, let's sort out how Jane could get around in the THIRD neighbourhood she's considering.
Now, let's sort out how Jane could get around in the **FOURTH** neighbourhood she's considering.

'Versatility' is still available to use

XX 4

Next
Jane says: good-bye and thanks you -- that's what she wanted to ask you about the neighbourhoods she is considering.

Now I'd like to ask several general questions about you.
### Do you agree or disagree with these statements:

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Neutral</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I frequently pay more for products or services that are environmentally friendly</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am generally willing to try out new ways of doing things, even if there is a chance that it might not work out</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I tend to plan the things I do several days in advance</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td></td>
<td>Next</td>
</tr>
<tr>
<td>I frequently buy organic foods</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I wouldn't mind sharing a table at a busy restaurant if it meant not waiting</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I prepare and live within a detailed personal budget</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td></td>
<td>Next</td>
</tr>
</tbody>
</table>
Thank you -- that's the main part of the survey

Before we wrap up, can I ask you about the advice you gave Jane?

Would you say Jane's routine of activities is similar or different to yours?

Very different  Neutral  Very similar

How similar or different was this game to how you think about getting around?

Very different  Neutral  Very similar

You just gave some advice to Jane. As you thought through Jane's choices, how close was your thinking to how you make choices for yourself?

Very different  Neutral  Very similar

What was different? (Please type in boxes below)

What was similar? (Please type in boxes below)
I feel like I have a good understanding of how a car club works

Strongly disagree
Neutral
Strongly agree

I would need to know more about car clubs before I would consider using one

Strongly disagree
Neutral
Strongly agree

Thank you very much for taking part in this study

Would you care to leave any comments, suggestions or thoughts?

<<Type your comments, suggestions or thoughts here>>

If you have any further questions or comments, or wish to receive updates on this study, please contact:

Scott Le Vine
Centre for Transport Studies, Imperial College
slevine@imperial.ac.uk
020 7594 6105

Alternatively, you may leave an e-mail address:
(We will only contact you about this study)
TRAINING AGENDA

I)    Introductions
II)   Sample interview with an interviewer as a respondent
III)  De-brief sheet & Interviewer Synthesis
IV)   Frequent points of confusion & questions from respondents
V)    Starting / closing survey in MS Access (make all pages suitable so can have context menu)
VI)   Transmitting data to Imperial (fax or email)
VII)  Recruitment
VIII) Format for incentives
IX)   Remaining Issues
NOTES FOR INTERVIEWERS

TO SUPPORT YOU

1) CALL SCOTT WITH ANY QUESTIONS, DAY OR EVENING: (0) 79 6456 8035. My top priority for the next few weeks is having this go well— I’m available to support you as little or much as needed.

2) If a respondent wishes to hold an interview at Imperial – call Scott to reserve a room (not a problem, even on short notice)

3) Main sections of the survey:
   CASI
   a. Gathering basic demographic information
   b. Introduction to choice experiment & practice – the most common point of confusion
   c. Choice experiment (four replications)
   d. Post-game questions (in software)
   MANUAL
   e. De-brief sheet (open-ended questions & informal discussion)
   f. Interviewer Synthesis, after leaving from the interview

4) Your role in the CASI (stages a through d) is to support and observe the respondent, noting any ‘alerts’ at the end of your De-Brief Sheet.

5) In stages e and f, after the software portion has ended, your role is to ask the questions on the de-brief sheet, and lead an informal discussion from which to complete a brief Interviewer Synthesis afterwards.

START-UP

1) Script for introduction:

   Thanks for coming along. As a reminder, we’ve asked you to take part in this survey about ways of getting around in urban areas.

   You’ll go through the main part of the survey on the computer, and then I’ll ask you some questions about it afterwards.

   It takes between a half-hour and 45 minutes. I’ll be here to guide you through the survey – do you have any questions before we begin? When you’re ready to get started, go ahead & click [Jane/Joe].

HELPING RESPONDENTS

1) Most common point of confusion – introduction to choice experiment. Be aware of possible need for guidance. Gently direct respondent to read the blue text & practice making choices in the left-hand column and in the centre. Ask them to read back to you their choices before clicking and moving on.

2) Common questions from respondents:
   a. “I’m not sure if I live in Inner or Outer London.” See A4 sheet with list of Inner/Outer London boroughs.
b. MISFIT RESPONSES SUCH AS: “I’d take the taxi back from the food store but would [walk/take the bus, etc.] to get there.” Tell them to select “taxi” & make a handwritten alert with round/activity #s.

c. “Is Jane/Joe dropping the son off or staying with him?” Staying with him.

d. “What’s this all about?” We’re interested in how people get around in urban neighbourhoods.

e. “Does it matter how many hours a week I work?” 10+ hours/week = “employed”

f. “Does [Jane/Joe] have a driving licence?” Yes

g. “The journey times and costs look implausible to me.” They’re what [Jane/Joe] has been told about the neighbourhood by [his/her] estate agent

USING THE COMPUTER

1) Open survey instrument by:
   a. Opening the survey file (Survey vXXX.accdb) which is on the Desktop.
   b. Then double-clicking on the form (a red and white icon) labelled “Page1” on the left-hand side (one line below “Background Page”)

2) If it becomes necessary to re-start survey instrument:
   a. Press CTRL-ALT-DEL & select “Task Manager”
   b. Right-click on MS Access & choose “end task”
   c. Re-start survey software as above. The respondent will have to re-start the software from the beginning.
   d. If the software fails a second time – end the interview.

3) Occasionally the software may pause for up to about a minute – re-assure the respondent & only re-start the survey after more than a minute or two of the software not responding.

4) The software will save the data from all of the interviews completed on that computer cumulatively in the survey file.

5) Sending data from completed surveys to Imperial (at the end of each day during which at least one interview is completed)
   a. Fax hard copies of de-brief sheets to (0) 207 594 6102 or scan & email to slevine@imperial.ac.uk
   b. Zip the survey file (Survey vXXX.accdb) & send to Scott at above address, either by email as an attachment, or via Imperial’s file exchange system: http://fileexchange.imperial.ac.uk/
   c. I’ll be checking the data as it comes in & will be in touch if I have any suggestions.

6) There is a backup copy of the survey file (Survey vXXX.accdb) in the folder C:\Survey.

7) Laptop login information:
   a. Machine #1 Login name: cvlocal Password: cvlocal
   b. Machine #2 Login name: projectlaptop Password: projectlaptop

FINISHING THE INTERVIEW

1) Please complete your Interviewer’s Synthesis as soon as possible after leaving the interview
REVISED NOTES FOR INTERVIEWERS

(For interviews from 25th February onwards)

TO SUPPORT YOU

1) CALL SCOTT WITH ANY QUESTIONS, DAY OR EVENING: (0) 79 6456 8035. My top priority for the next few weeks is having this go well—I’m available to support you as little or much as needed.
2) If a respondent wishes to hold an interview at Imperial—call Scott to reserve a room (not a problem, even on short notice)
3) Main sections of the survey:
   CAPI
   a. Gathering basic demographic information
   b. Introduction to choice experiment & practice – you will guide them in this portion
   c. Choice experiment (four replications)
   d. Post-game questions (in software)
   MANUAL
   e. De-brief sheet (open-ended questions & informal discussion)
   f. Interviewer Synthesis, after leaving from the interview
4) Your role in the CAPI (stages a through d) is to support and observe the respondent, noting any ‘alerts’ at the end of your De-Brief Sheet. The exception to this is stage b, where you guide them through the introduction to the choice experiment
5) In stages e and f, after the software portion has ended, your role is to ask the questions on the de-brief sheet, and lead an informal discussion from which to complete a brief Interviewer Synthesis afterwards.

START-UP

1) Script for introduction:

   Thanks for coming along. As a reminder, we’ve asked you to take part in this survey about ways of getting around in urban areas.

   You’ll go through the main part of the survey on the computer, and then I’ll ask you some questions about it afterwards.

   It takes between a half-hour and 45 minutes. I’ll be here to guide you through the survey – do you have any questions before we begin? When you’re ready to get started, go ahead & click [Jane/Joe].

HELPING RESPONDENTS

1) The most common point of confusion – introduction to choice experiment. Follow the oral script and ensure that the respondent has understood each function before moving on to the next function. Ask them to read back to you their choices before clicking and moving on.
2) Common questions from respondents:
a. “I’m not sure if I live in Inner or Outer London.” See A4 sheet with list of Inner/Outer London boroughs.

b. MISFIT RESPONSES SUCH AS: “I’d take the taxi back from the food store but would [walk/ take the bus, etc.] to get there.” Tell them to select “taxi” & make a handwritten alert with round/activity #s.

c. “Is Jane/Joe dropping the son off or staying with him?” Staying with him.

d. “What’s this all about?” We’re interested in how people get around in urban neighbourhoods.

e. “Does it matter how many hours a week I work?” 10+ hours/week = “employed”

f. “Does [Jane/Joe] have a driving licence?” Yes

g. “The journey times and costs look implausible to me.” They’re what [Jane/Joe] has been told about the neighbourhood by [his/her] estate agent.

USING THE COMPUTER

1) Open survey instrument by:
   a. Opening the survey file (Survey vXXX.accdb) which is on the Desktop.
   b. Then double-clicking on the form (a red and white icon) labelled “Page1” on the left-hand side (one line below “Background Page”)

2) If it becomes necessary to re-start survey instrument:
   a. Press CTRL-ALT-DEL & select “Task Manager”
   b. Right-click on MS Access & choose “end task”
   c. Re-start survey software as above. The respondent will have to re-start the software from the beginning.
   d. If the software fails a second time – end the interview.

3) Occasionally the software may pause for up to about a minute – re-assure the respondent & only re-start the survey after more than a minute or two of the software not responding.

4) The software will save the data from all of the interviews completed on that computer cumulatively in the survey file.

5) Sending data from completed surveys to Imperial (at the end of each day during which at least one interview is completed)
   a. Fax hard copies of de-brief sheets to (0) 207 594 6102 or scan & email to slevine@imperial.ac.uk
   b. Zip the survey file (Survey vXXX.accdb) & send to Scott at above address, either by email as an attachment, or via Imperial’s file exchange system: http://fileexchange.imperial.ac.uk/
   c. I’ll be checking the data as it comes in & will be in touch if I have any suggestions.

6) There is a backup copy of the survey file (Survey vXXX.accdb) in the folder C:\Survey.

7) Laptop login information:
   a. Machine #1 Login name: cvlocal Password: cvlocal
   b. Machine #2 Login name: projectlaptop Password: projectlaptop

FINISHING THE INTERVIEW

1) Please complete your Interviewer’s Synthesis as soon as possible after leaving the interview
Interviewer initials & interview # (e.g. SL-5): ___________ Respondent first name: ________________
Date/time of interview: __________________ Location: __________________________________________
London borough: __________________________ Means of contact: ____________________________

REVISED SCRIPT FOR PRACTICE ROUND

(For interviews from 23rd February onwards)

Check after reviewing each item:

INTRODUCTION:

I’m going to take you through the practice portion of the interview. Remember, I’ll let you know before we begin for real.

JOE/JANE’S ACTIVITIES:

We know a number of the things that Jane/Jo does in a typical week, two of them are shown here [POINT TO ACTIVITY PICS]. The type of activities that Jane/Jo does is the same in each of the four neighbourhoods. You can click on each activity to find out a bit more about each of them, and how often Jane/Jo does each of them. Go ahead and click on his/her first activity [POINT TO FIRSTACTIVITY]. You can click again to bring back the picture if you like.

INFO ABOUT JANE/JOE:

You’ll see a picture of Jane/Jo on the left side of the screen. In any neighbourhood, clicking on the round blue button next to this picture will bring up some information about Jane/Jo’s life circumstances. Go ahead and click on the round blue button now. [ THEN MAKE SURE THEY CLICK TO RETURN TO THE MAIN PRACTICE SCREEN]

SWITCH TO 2ND PRACTICE SCREEN

Now, go ahead and click on the button that says ‘Next’ in the bottom right and I’ll show you more bits of the practice screen.
WAYS OF GETTING AROUND:

Jane/Joe is asking you for advice on how he/she should get around. If you look at the very top of the screen, you’ll see the different ways that he/she could travel. For each thing he/she does, you’ll need to select how he/she should travel.

‘GETTING AROUND’ BUTTONS:

You’ll see buttons to the right of each activity, which offer different ways to get there and back.

MOUSE-OVERS:

These buttons in each row only say how much time each option would take and how much it would cost. If you leave the mouse on top of any of these buttons for a moment, you’ll see more about what it would be like to use that option to get to that activity. Go ahead and move the mouse over one of the buttons to give it a try.

MAKING CHOICES ABOUT GETTING AROUND:

After you’ve thought a bit about the different options for Jane/Joe to get to an activity, you can select one of them by clicking on the button. The text will change colour to show that you’ve selected it. Take a few moments now and think about the options for getting to Jane’s first activity. When you’re ready, go ahead and select one of the options to get to Jane’s first activity.

BIG DECISIONS:

You’ll notice that the taxi and walking options are black, but the others are grey. To use any of the other options for getting around, Jane/Joe must first make a ‘big decision.’

INFORMATION BUTTONS:

To understand these big decisions better, information is available by clicking the round blue buttons near the top of the screen. Go ahead and click each of the blue buttons one at a time, I’ll wait for you.

BUY BUTTONS:

If you click on any of the ‘buy’ buttons directly below the information buttons, you’ll see that Jane/Joe can now use that way of getting around. The travel options directly below it will turn from grey to black.

BIG DECISIONS AREN’T FREE:

These big decisions aren’t free and Joe/Jane will have to pay if he/she wishes to ‘buy’ them. The top right of the screen keeps track of the purchases. Go ahead and select each of them to see how much they would cost Jane/Joe.

CANCEL BUTTONS:

Any ‘big decision’ can be reversed by clicking the ‘cancel’ button, and then the buttons below become grey again. The only exception is, public transport can be used either Pay-as-you-go or by prepaying unlimited journeys with a season ticket.
GET JANE/JOE TO EACH ACTIVITY:

Now, in each neighbourhood we’ll ask you to sort out how Jane/Joe should get to each one of his/her activities. If you wish, you may select as many or few of the big decisions as you like, and none if you want.

EXPERIMENTING:

You can experiment as much as you like with the ‘big decisions’ and the different options for getting around.

When you’ve got each one of Jane’s/ Joe’s activities sorted out, you’ll be able to click the button in the bottom right that says ‘I would do this if I were Jane/ Joe’. There’s a lot to consider, so feel free to explore and take your time. Remember that I’m here to answer any questions you’ve got as you go through, so just take the time you need to reach your decisions. As you consider the advice you’ll give to Jane/ Joe, keep in mind that there are no right or wrong answers, and that the choices in each neighbourhood are judgment calls.

Go ahead now and sort out getting to these two activities, and pause after you’ve made a choice in each of the two rows.

HOW MUCH TRAVEL/EXPENSE THIS MEANS:

In the bottom right side of the screen, you’ll see some information on how much time and expense Jane/ Joe will spend travelling to get to and from his/her activities.

MOVING ON TO THE NEXT NEIGHBOURHOOD:

When have made one travel choice for each one of Jane’s/ Joe’s activities, and are happy with your advice, you can click the button on the bottom right that says ‘I would do this if I were Jane/ Joe’.

EACH NEIGHBOURHOOD IS A NEW SET OF OPPORTUNITIES:

Before you do that, keep in mind that each neighbourhood is a completely different place, and you’ll be starting fresh. You’ll be making all the decisions again in each new neighbourhood, they don’t carry through.

Go ahead and click the button on the bottom right that says ‘I would do this if I were Jane/ Joe’.

RETURN TO PRACTICE IF DESIRED:

Now, if you’d like we can go through the practice again, or if you’re ready to get started then go ahead. I’ll introduce you to the first neighbourhood, and then it’s your turn to give Jane/ Joe advice.

AS SOON AS THEY ENTER THE FIRST NEIGHBOURHOOD:

This is the first neighbourhood, and you’ll see that there are now more activities for Jane/ Joe to get to. Now, I’ll ask you to take a moment to learn more about each of Jane’s/ Joe’s activities.

Go ahead and click on each of them, and consider what they might be like. Remember that you can click again to bring back the picture if you like, I’ll wait for you to go through them.

[ONLY AFTER THEY’VE GONE THROUGH EACH ACTIVITY]

OK, now go ahead and sort out how Jane/ Joe should get around from this neighbourhood. I’ll leave you to proceed from here, but will be right here if you have any questions.
Interviewer initials & interview # (e.g. SL-5): ___________ Respondent first name: ___________

Date/time of interview: ________________ Location: __________________________________________

London borough: ______________________ Means of contact: ___________________________________

ALERTS – INDICATE NEIGHBOURHOOD # (E.G. N1)

____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
Interviewer initials & interview # (e.g. SL-5): ___________ Respondent first name: ______________

QUESTIONS TO ASK RESPONDENT:

SOME QUESTIONS ABOUT THE GAME YOU JUST PLAYED

Thinking back, how happy are you with your advice to [Jane/Joe] – if you were doing it again, would you change any of your advice?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

What do you think about the five examples of activities that we picked for [Jane/Joe] to:

1) Give [him/her] advice about getting around day-to-day **AND**
2) What about the big decisions?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

SOME OTHER QUESTIONS ABOUT YOU

Two or three years from now, in which ways – if any – do you expect your life to be different? [“Anything else?” – probe gently for additional responses]

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

If you were moving to another neighbourhood, how would you find out about ways to get around? [“Anything else?” - probe gently for additional responses]

[Note to interviewers: the number of distinct options that they come up with will be important in the analysis]

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
Interviewer initials & interview # (e.g. SL-5): ___________ Respondent first name: ___________

IF RESPONDENT HAS A CAR: As we discussed, you generally have a car available for your own use. 
What would your life be like without one?

________________________________________________________________________________________

Can you imagine circumstances where you could do away with it?
________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

IF RESPONDENT DOES NOT HAVE A CAR: As we discussed, you don’t own a car. If you had one, 
what would you do with it?
________________________________________________________________________________________

________________________________________________________________________________________

Can you imagine a future point in your life where it would be tough to make do without owning a 
car? [Note to interviewers: probe for life circumstances, not a particular time/place need for a car]
________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

________________________________________________________________________________________

IF CAR CLUB MEMBER – HOW LONG?  Years: _____ Months_______ OR Year joined:__________

• If the respondent has not said very much about how they experienced the choice game, feel free to ask about it to help you later with your synthesis

• End of formal interview – pay attention to noteworthy comments they may make as they relax and you put away the computer
INTERVIEWER’S SYNTHESIS: WHAT WENT ON? YOUR OBSERVATIONS MUST INCLUDE AT LEAST THE ITEMS BELOW:

1. Did they find any parts of the CASI difficult to use or understand? Which?
2. Briefly describe how they went about the choice exercise. (e.g. systematic v. chaotic, etc.)
3. What seemed to you to grab their attention the most as they made their choices on-screen?
4. Did they make any noteworthy comments? (e.g. reason for survey)

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
References

Adamou, O. (2011) *A Need-based Analysis Framework to characterise Induced and Suppressed Travel associated with Subscription Drive-it-yourself Car Services*. Undergraduate thesis, Imperial College London.


Carplus (2007), ‘Car club accreditation criteria (v4)’. Leeds, UK.


City Carshare, (2005). Bringing carsharing to your community. San Francisco, US.


Diem, W. (2008); *In Germany, Cars by the Hour*; New York Times; 26 October.


Doherty S., Lee-Gosselin M., Burns K., Andrey J. (2002); *Household Activity Rescheduling in Response to Automobile Reduction Scenarios*; Transportation Research Record #1807, 174-182.


Firnkorn, J. & Müller, M. What will be the environmental effects of new free-floating car-sharing systems? The case of car2go in Ulm. Ecological Economics, 2011


Fishman, L. and Wabe, S. J. (1967), 'Restructuring the form of car ownership: A proposed solution to the problem of the motor car in the United Kingdom', Transportation Research 3(4), 429-442.


Harmer, C and Cairns S (2011) *Carplus annual survey of car clubs 2010/11.* Transport Research Laboratory, Crowthorne


Hilkevitch, J. (2008), *'CTA, car-sharing group announce plans for new transit card',* Chicago Tribune (11 October).


Ibanez, J.N., Toner, J. and Daly, A. (2007) Optimality and efficiency requirements for the design of stated choice experiments, European Transport Conference.


268


269

Morenc, C., Trepanier, M. and Martin, B. (2007), 'Object-oriented analysis of a car-sharing system', Transportation Research Record. 2063 (105-112)


Nassauer, S. (2008), 'If your Zipcar is costing more, the taxman may be to blame', The Wall Street Journal (19 June).


Pas, E. I. Weekly Travel-Activity Behavior. Transportation, 15, 89-109


Robert, B. (2005c), Recueil de textes sur communauto, Master's thesis, Université Laval (Québec

Roberts, A. (2008), Personal correspondence via e-mail, 31 October.


Turrentine, T. S., Lee-Gosselin, M., Kurani, K. and Sperling, D. (1992), *A study of adaptive and optimizing behavior for electric vehicles based on interactive simulation games and revealed behavior of electric vehicle owners*, in 'World Conference on Transport Research; Lyon, France'.


