Public Transp (2011) 3:109–147
DOI 10.1007/s12469-010-0035-4

CASE STUDIES AND APPLICATIONS

Identifying commuter preferences for existing modes and a proposed Metro in Sydney, Australia with special reference to crowding

David A. Hensher · John M. Rose · Andrew T. Collins

Published online: 1 February 2011

Abstract In 2009, the New South Wales government announced that it would be proceeding with a feasibility study to identify the patronage potential of a new Metro rail system for Sydney. As part of this study, a new modal choice study was undertaken to establish the role of traditional attributes such as travel times and costs (and more recently, reliability) but also somewhat neglected influences such as crowding, where the later has a critical role in the calculation of capacity needs at railway stations. This paper focuses on the commuter segment and develops a new stated choice experiment in which travellers are able to compare the proposed new Metro with existing available modal alternatives for access, linehaul and egress trip stages, with a particular emphasis on the incorporation of crowding represented by the availability of a seat vs. standing in existing and new public transport modes. We present the error component choice model together with estimates of mode-specific willingness to pay for travel time components, service frequency and crowding, that latter expressed in terms of the probability of getting a seat and the probability of avoiding standing.

Keywords Mode choice · New mode · Stated choice · Crowding · Frequency · Willingness to pay · Survey design

---

Electronic supplementary material The online version of this article (doi:10.1007/s12469-010-0035-4) contains supplementary material, which is available to authorized users.

D.A. Hensher (✉) · J.M. Rose · A.T. Collins
Institute of Transport and Logistics Studies, The University of Sydney Business School,
The University of Sydney, Sydney, NSW 2006, Australia
e-mail: david.hensher@sydney.edu.au

J.M. Rose
e-mail: john.rose@sydney.edu.au

A.T. Collins
e-mail: andrew.collins@sydney.edu.au

1 Introduction

In 2009, the New South Wales government announced that it would be proceeding with a feasibility study to identify the patronage potential of a new Metro rail system for Sydney. As part of this study, a new modal choice study was undertaken to establish the role of traditional attributes such as travel times and costs, but also the somewhat neglected influences such as crowding, where the latter has a critical role in the calculation of capacity needs at railway stations.

The research involved definition of the modal context in which the Metro is an alternative means of transport, design of a new survey instrument, estimation of choice models to identify the role of candidate influences on modal choice (defined by access, main and egress components of modal alternatives for a given trip), and to establish willingness to pay (WTP) outputs.

This paper presents the approach developed in the context of commuter travel undertaken in 2009 (although non-commuting and employer-business trips were also studied). We set out in some detail all elements of the process beginning with problem definition, phased design of the survey instrument, identifying the geographical catchment area for sampling, piloting the instrument, reviewing the pilot outputs, and determining revisions to a Computer Aided Personal Interview (CAPI). As part of model estimation, we selected an error components choice modelling method. The results from model estimation are presented and discussed together with willingness to pay outputs.

2 Problem definition

The centrepiece of the research is a model of choice amongst relevant existing modal ways of travelling between predefined start and finishing locations and the proposed Metro (the latter defined by the three phases of infrastructure provision, i.e., CBD, Western and North-Western lines).

The main feature of data definition is a stated choice experiment in which, for metropolitan-wide travel in the catchment area, we begin by defining a current or recent trip experience, identify existing modal alternatives that could seriously be considered, and establish the station-to-station pairing for the Metro that would be the sensible Metro trip if the Metro were available today. Importantly, eligible respondents would have been screened as in-scope[1] prior to their participation in the choice experiment (details of this process are given in a later section).

The stated choice (SC) experiment involves a respondent comparing the levels of times, service levels (e.g., frequency, crowding and getting a seat for public transport) and costs (i.e., fares, running costs, tolls and parking for car) of various modes to complete the same current or recent trip. The alternatives that are shown to individual respondents include the modes that are currently available to them plus a metro

[1] In-scope refers to screening eligible trips according to the origin and destination of the trips that the project team agreed would be candidate trips for considering the Metro as a feasible alternative.

option. The respondent has to choose one of these alternatives. The process of choosing amongst the alternatives is repeated a total of six times, in which each choice situation involves varying the levels of times, service levels and costs associated with the different alternatives.

An error components logit model is estimated that accounts for the panel nature of the data (six choice scenarios) for each respondent, taking into account the correlated structure on common-respondent observations. Importantly, the error components logit model allows for degrees of similarity between subsets of modal alternatives (which is also the appeal of nested logit (NL), although NL cannot account for repeated observations from one individual—see Hensher et al. 2005); for example, it might reasonably be hypothesised that public transport modes have a set of unobserved influences that are common within the public modes but different to the car alternative.

The estimated choice models are used to calculate the marginal willingness to pay for various components of travel time, number of transfers, service frequency and crowding.

3 The stated choice approach

A stated choice (SC) experiment is developed in which we systematically vary combinations of levels of each attribute to reveal new opportunities relative to the existing circumstance of time-cost-service level on offer (see Hensher 1994; Louviere et al. 2000; Hensher et al. 2005; Rose et al. 2008). Through the experimental design paradigm, we observe a sample of travellers making choices between the current trip attribute level bundle (or a package of time, cost and service levels), and other alternative attribute level bundles. This approach is the most powerful method capable of separating out the independent contributions of each time, cost and service component between a number of modal options for a specific trip purpose. It is the preferred approach, capable of providing disaggregated willingness to pay estimates for components of trip time (referred to as values of travel time savings), in addition to the weighted average of these time components, the valuation of travel time reliability, and the valuation of other user benefits (in dollars per trip and per hour) such as crowding and getting a seat.

For car and bus, we include the variability of travel time as a combination of a number of separate attributes. The definition of this attribute for road modes (i.e., bus and car), where there is often delays and uncertainty in trip times due to traffic, breakdowns and accidents etc., focuses on the probability of an individual's trip time being a specific level around the average time that a specific trip would take over a given period of time (e.g., a week for commuters). The attribute has three parts—the probability of 'arriving x minutes earlier than expected', 'arriving y minutes later than expected', and 'arriving at the time expected'. A value for the reliability of travel time can be obtained as a weighted average of the three values, where the weights reflect the probability mix. For rail modes, we emphasis a single trip time, on the reasonable assumption that the existing Cityrail system and the proposed new Metro[2] are not delayed by traffic levels. Trip costs are the toll, fuel and parking costs for car, and fares for public transport. Each sampled respondent evaluated six (6) choice scenarios, choosing a preferred mode within each choice scenario from a maximum of three stated choice (SC) alternatives defined by two existing alternatives plus the Metro.

Technically, a Bayesian-efficient design was constructed using D-error as the optimality criteria. Priors were obtained from a pilot survey consisting of 82 respondents conducted in April of 2009. The final design was constructed with 240 choice tasks, allowing for attribute level balance across the design (i.e., each attribute level appears an equal number of times over the experiment, thus minimising any potential bias that may arise from showing one attribute level more than other levels). Given that efficient designs are not orthogonal, the construction of an orthogonal blocking column was not possible. Therefore, a blocking column with 40 levels was generated which minimised the maximum correlation between this column and the other attribute columns. As such, the blocking column was not perfectly orthogonal however the largest correlation was 0.16. Respondents were systematically assigned to each subsequent block during the interview process.

4 The survey[3]

The survey instrument has five major sections:

[2] We were advised by the Metro project that a key objective of the Metro was to 'guarantee' on-time arrival and were instructed to assume this. Furthermore, for existing rail, while there are experienced delays they are quite minimal.

[3] Prior to the main survey, a pilot study was conducted to ensure that the logistics, sampling rules and survey instrument were appropriate for the task at hand. The pilot proved extremely valuable in testing the draft survey instrument and in guiding the revisions of the briefing of interviewers, as well as revisions to the phone and face to face screeners to recruit respondents. The pilot identified a number of features of the instrument that were in need of reconsideration. In particular:

(a) We concluded that it was sensible to limit the number of alternative ways of travelling to a maximum of two current main modes and Metro (previously a person could select all four, although the majority only selected one plus Metro).
(b) It was argued that we should change crucial choice set wording in order to establish eligible alternatives. The new wording is 'would seriously consider'.
(c) The major issue revealed from the pilot was the method of screening to ensure eligibility in respect of the proposed Metro being a feasible alternative. To ensure the appropriate checks are in place in the main survey, we set up a multi-stage process to check that all respondents screened as 'in scope' are indeed eligible before they are interviewed. The potential candidates sourced from the telephone screening are checked by staff at Taverners to ensure that start and finish locations would make Metro a feasible alternative. Interviewers were briefed for six hours in early May to ensure these crucial sampling issues were adhered to at the face to face screening (which complements the telephone screening). The interviewers were also provided with a hard copy numbered map of the metro lines, to be used to ensure at screening that respondents are eligible.
(d) As a result of limiting the number of alternatives in a choice set, plus the simplification of the attitudinal questions, we moved from four to six choice screens, ensuring more robust parameter estimates. The revised survey instrument takes on average 25 minutes compared to over 30 minutes using the pilot instrument.

1. The introduction to the survey task and background on the study.
2. For the metropolitan-wide segments, questions describing a current or recent trip in terms of travel times and costs for available modes for the door-to-door trip in question, including (perceived) times and costs for modes that were not used, up to a maximum of two alternative main modes.
3. The stated choice (SC) experiment (six screens) in which the respondent compares the levels of times, service levels and costs of various modes to complete the same trip described by earlier. The alternatives that are shown to individual respondents include the modes that were available to them as described earlier (for metropolitan-wide segments), plus a metro option. The respondent has to choose one of these alternatives. The process of choosing amongst the alternatives is repeated a total of six times (each choice situation involves varying the levels of times, service levels and costs associated with the different alternatives).
4. A series of attitudinal questions seeking views on the broader set of quality benefits of specific public transport modes; and
5. Some socio-economic questions collected to provide some background of respondents.

The SC experiment offers a total of five different alternatives for metropolitan trips and six for within CBD trips. These alternatives are car, bus, light rail, cityrail and metro. Any one respondent however is limited to choosing amongst a maximum of two existing alternative plus the proposed Metro. The survey itself was conducted using a computer aided personal interview (CAPI) allowing respondents to be asked to provide information, either real or perceived, related to the levels of the relevant alternatives for a recent trip that they undertook. The SC experiment for metropolitan-wide segments then 'pivots' the attribute levels of the various alternatives, where a pivot from the reference trips makes sense. The attributes to pivot are the travel times and costs.[4]

This approach acts to frame the decision context of the choice task within some existing memory schema of the individual respondents, and hence make preference-revelation more meaningful at the level of the individual. Theoretically, the role of reference alternatives in SC tasks is well supported within the literature. For example, prospect theory (Kahnemann and Tversky 1979), which argues that individuals use decision heuristics when making choices, promotes the idea that the very specific context in which a decision is made by each individual is an important determinant of the selection of choice-heuristic, supporting the use of reference alternatives in SC tasks. Framing effects, of which reference dependence is a popular interpretation, provides context support in trading off the desire to make a good choice against the cognitive effort involved in processing the additional information provided in a SC task (Hensher 2010). Case-based decision theory (Gilboa et al. 2002) promotes

[4] A referee suggested that we might have included crowding on existing Cityrail and bus as pivots. Although this is a nice idea, it is one that might be taken on board in further research, since we are unable to redo the study. We should add, however, that the construction of the levels of crowding as a non-pivot attribute is less problematic than time and cost, where the latter two attributes are far more important as pivots as a way of ensuring sensible trip lengths and costs to evaluate. The crowding levels tend to be much more behaviourally meaningful for all trip time and cost situations.

the role of accumulated experience represented by a reference alternative. Starmer (2000, p. 353) in particular argues strongly for the use of reference alternatives (e.g., a current trip) in decision theory:

> "*While some economists might be tempted to think that questions about how reference points [alternatives] are determined sound more like psychological than economic issues, recent research is showing that understanding the role of reference points [alternatives] may be an important step in explaining real economic behaviour in the field.*"

The number and types of alternatives to be shown vary from respondent to respondent. This variation is determined by the responses given by respondents early in the survey in terms of the availability of the various alternatives for the recent trip being examined. For example, if the respondent reports not having a car available for the recent trip, then the car alternative will not be present in the SC experiment.

As well as the number and types of alternatives varying across respondents, several attributes may vary also. For example, access and egress attributes relate to different mode possibilities. As with the main mode alternatives, these attributes only appear if the respondent indicates that they are a valid option for the trip being examined.

Exactly how analysts distribute the levels of the design attributes over the course of an experiment (which typically is via the underlying experimental design), may play a big part in whether or not an independent assessment of each attribute's contribution to the choices observed to have been made by sampled respondents can be determined. Conceptually, an experimental design may be viewed as nothing more than a matrix of values that is used to determine what goes where in a SC survey. The values that populate the matrix represent the attribute levels that will be used in the SC survey, whereas the columns and rows of the matrix represent the choice situations, attributes and alternatives of the experiment.

The experimental design consists of 240 different choice tasks. Individual respondents were asked to review 6 of the 240 choice tasks and to indicate which alternative they prefer. The exact levels shown to each respondent vary depending on the times and costs that each respondent reports for a recently undertaken trip. The trip times and costs in the choice scenarios associated with each existing mode are pivoted around current perceived levels; however for the other attributes such as frequency, crowding, getting a seat and number of transfers, the range was fixed and unlinked from the current trip experience (see Table 1). The mode specific attributes and the percentage levels to be used in the survey are given in Tables 2 to 5. The majority of the attributes have five levels (exception being crowding, number of transfers and service frequency), which were found to deliver a sufficiently wide range as well as levels within the range that would be meaningful to respondents, and in the case of the proposed Metro, give us enough variability in the range that the metro is likely to be represented when constructed.

The attributes and attribute levels for the study were derived in conjunction with the Metro project authority. The literature was also consulted to determine relevant pivot levels. The pilot study conducted in April 2009 was used to test the logistical aspects to be implemented for the main field phase of the project, as well as to test the operational capabilities of the Computer Aided Personal Interview (CAPI) software that will be used. Furthermore, results from the pilot were used to provide priors for constructing the design used in the main field phase. The median time per survey from the pilot study was 38 minutes, the largest proportion of which occurs from the non-SC questions of the survey (time per SC question is approximately 20–30 seconds per choice task). The median 38 minutes compares to 20–25 minutes originally planned for. For this reason, the survey was shortened in terms of additional questions asked such as attitudinal questions. Further, partly to help limit the amount of survey time, and partly to do with the complexity of the SC screens, it was felt that limiting respondents to four choice tasks was best in the final field phase. This was also confirmed in interviews held with the pilot sample after they had undertaken the survey.

Example choice scenario screens are shown in Figs. 1(a) and (b) for metropolitan wide trips.

Crowding required careful consideration since it is a focal issue in this study. We are interested in both the number of seats occupied and the number of people standing in the train, metro and bus/light rail. The approach adopted to represent crowding on a public mode involves both a written description and schematic diagrams of the seating configuration for each mode showing the people seated and standing. The combinations defining the crowding attribute are summarised in Table 1. For example, level 10 is a situation where every seat is occupied and there are either 0, 15 or 16 standing depending on which mode the level is assigned to in the experimental design. The graphical schema depicting crowding for each mode is based on the 16 levels[5] with an example in Fig. 2 of a 100 percent seated train carriage and 10 people standing, based on ideas developed by Whelan and Crockett (2009). The selection of the attribute levels was determined in consultation with the Sydney Metro Authority.

5 The main survey field activity

620 individuals resident in Sydney were selected as the total commuter sample size. The sampling plan reflects travel across the catchment area that not only includes the CBD, but also areas within the metropolitan area that extend as far west as Westmead (in terms of the rail network), and the north west of Sydney (out to Rouse Hill).

The proposed Metro lines (Fig. 3) define the catchment area. Considerable thought was given to defining candidate trip origins and destinations, with daily checks of potential in-scope respondents prior to interview. These checks were undertaken by the managers of the survey team. Having a trip commence and finish in the catchment area does not mean in-scope; rather some of these trips (e.g., Castle Hill to Parramatta) are deemed out of scope since the Metro is hardly a meaningful mode for such a trip either as an access, main or egress mode. Furthermore, some in-scope trips can begin in the catchment area and finish outside (e.g., Gladesville to Kensington).

The sample was designed to capture trips in both the proposed Metro's geographical catchment area and the rest of the metropolitan area to the extent that the trips can

[5] The full set of 16 levels are available on supplementary material from the first author.

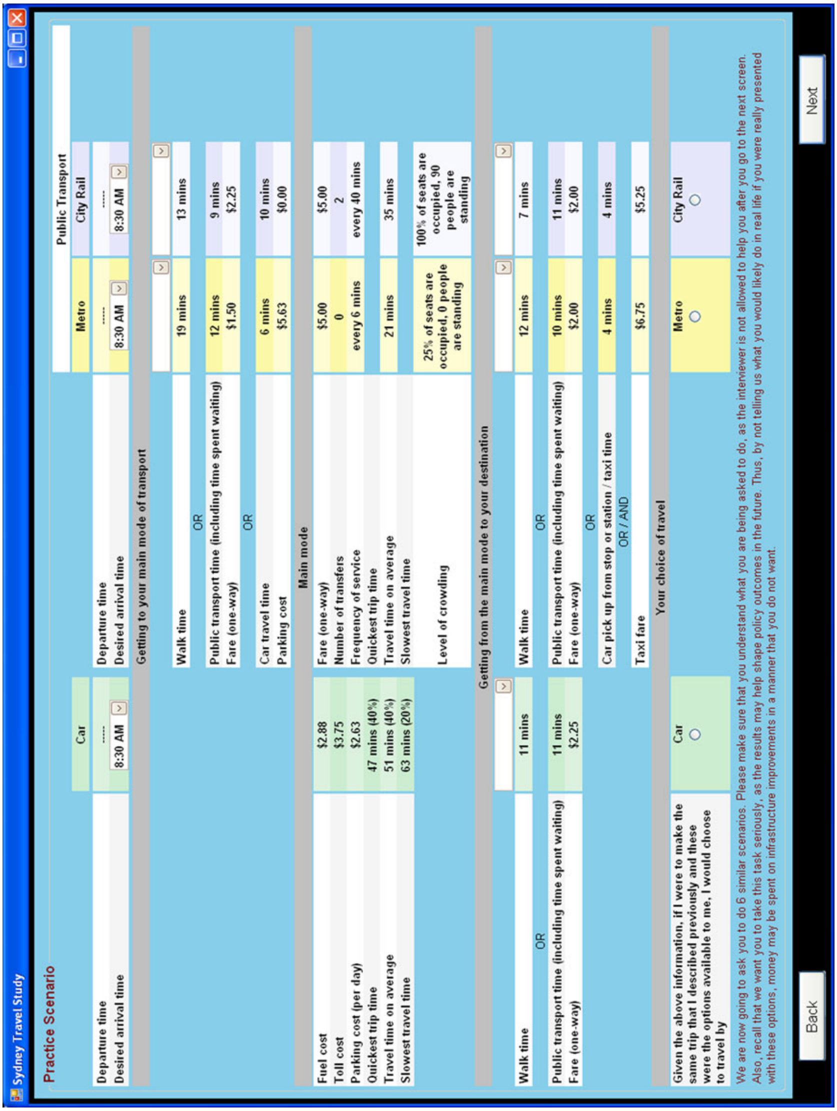


Fig. 1a Example Metro-wide choice scenario screen

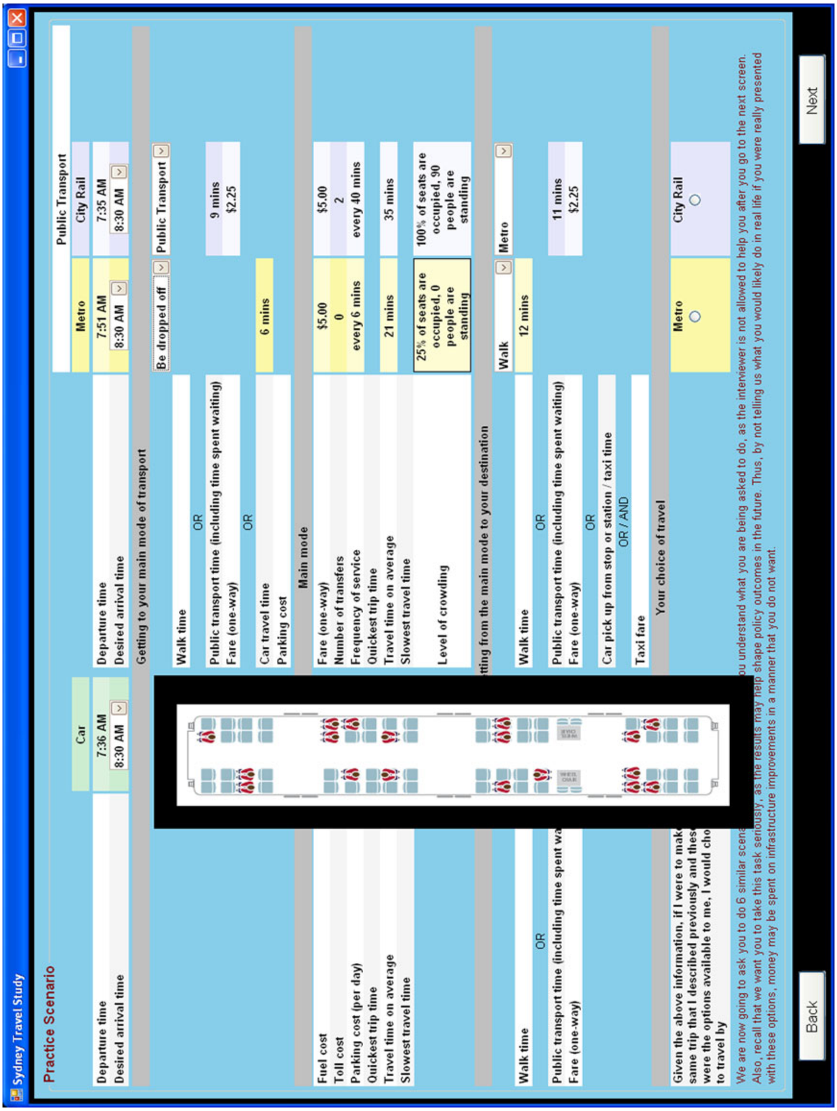


Fig. 1b Example Metro-wide choice scenario screen

Table 1 Crowding attribute levels by mode

Level	Seated	Bus Standing	Train Standing	Metro Standing
1	25%	0	0	0
2	50%	0	0	0
3	60%	0	0	0
4	70%	0	0	0
5	80%	0	0	0
6	80%	5	5	5
7	90%	0	0	0
8	90%	5	5	5
9	100%	0	15	16
10	100%	3	30	32
11	100%	7	45	47
12	100%	11	60	63
13	100%	15	75	78
14	100%	19	90	94
15	100%	23	105	109
16	100%	27	120	125

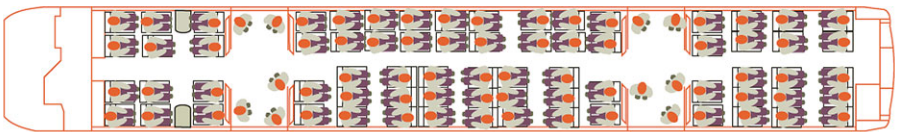


Fig. 2 Example of crowding within a public mode

finish anywhere in the metropolitan area, provided that they commence in the catchment area. The final sample profile is summarised in Table 6. The difference with the required quotas in large measure is attributable to the following considerations that emerged as the fieldwork progressed:

1. Commuter off-peak trips are difficult to find, and so we have removed the peak and off-peak distinction.
2. The postcode allocation was modified to be more in line with the percent catchment demand of approximately 50 percent/30 percent/20 percent for East/Middle/West trips.
3. The candidate access and egress modes associated with each existing main mode are not all easy to identify as actual trips as per the quotas. Examples are public transport access to train for commuters and car access to bus or light rail. We have relaxed the quotas after 50 percent of the data was collected, to allow more flexibility in the selection of the access and/or egress modes.

The valid sample cell in Table 6 counts only the quota or how much of the quota is full, not oversubscribed interviews. The main field work commenced in early May and was completed on Sunday May 31, eight days ahead of schedule.

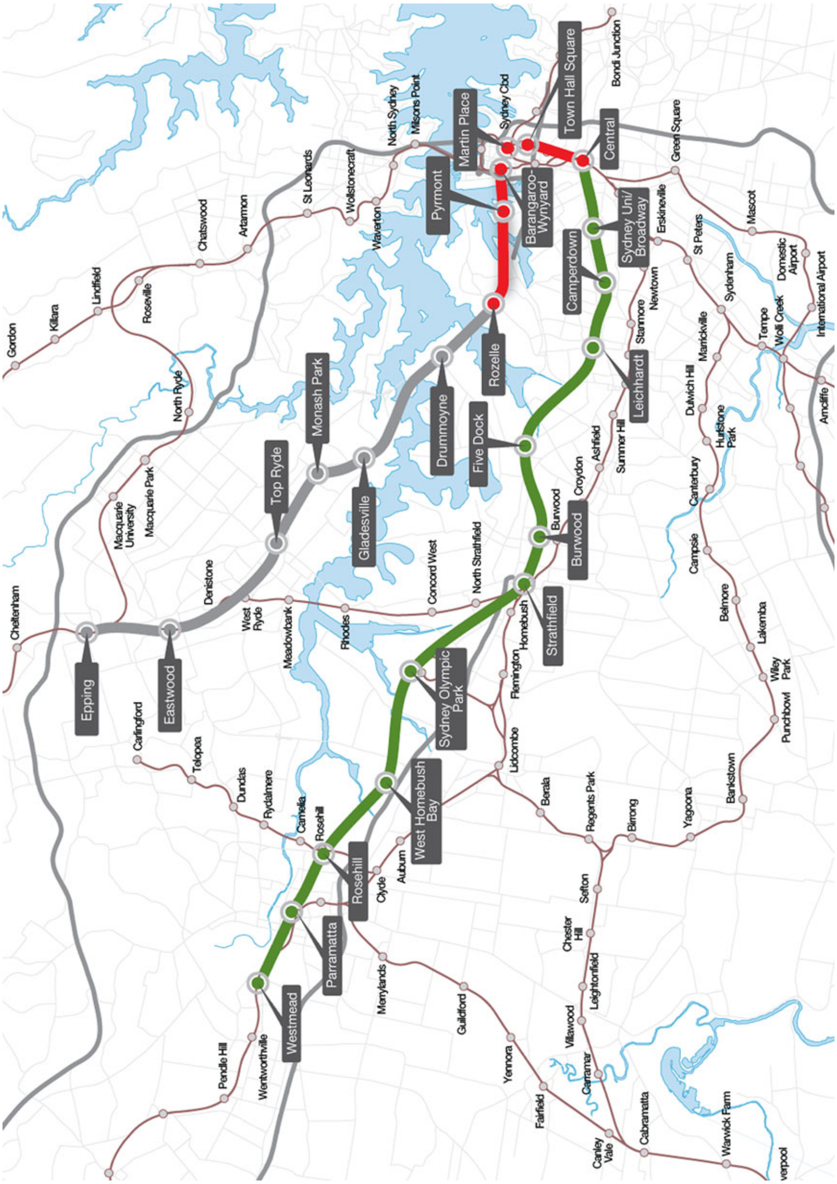


Fig. 3 Proposed Metro lines

Table 2 Car attributes and attribute levels

Attribute	# levels	Levels	Pivot/rule
Main mode			
Fuel cost ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Base = (average speed/60× mins)/(100 × 8) × $1[a]
Toll cost (per day) ($)	5	−25%, −12.5%, 0%, 12.5%, 25% or $0, $2, $4, $6, $8	Pivot from reported toll, or if non-toll road, use levels $0, $2, $4, $6, $8
Parking cost (per day) ($)	5	−25%, −12.5%, 0%, 12.5%, 25% or $0, $5, $10, $15, $20	Pivot from reported parking cost, or if no cost reported, use levels $0, $5, $10, $15, $20
Quickest trip time (mins)	5	−25%, −20%, −15%, −10%, −5%	Value less than average pivot time[b]
Travel time on average (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported trip time
Slowest trip time (mins)	5	10%, 15%, 20%, 25%, 30%	Value more than average pivot time
Quickest trip time %	5	10%, 15%, 20%, 25%, 30%	
Travel time on average %	5	40%, 45%, 50%, 55%, 60%	
Slowest trip time %		–	100% − (Travel time on average % + Quickest trip time %)
Egress mode			
Walk time (mins) to final destination	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivotted off respondent provided level if eligible alternative is indicated by respondent
Public transport time (including time spent waiting) (mins)	5	−25% , −12.5%, 0%, 12.5%, 25%	Pivotted off respondent provided level if eligible alternative is indicated by respondent
Fare (one-way) ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivotted off respondent provided level if eligible alternative is indicated by respondent

[a] Assumes average fuel efficiency of 8 litres for every 100 km travelled, and fuel of $1 per litre

[b] If the reported car trip time is 30 minutes, and the average trip time in a choice set is −25% (hence = 22.5 minutes), then the quickest trip time will be this % less than 22.5 minutes

Table 3 Bus and light rail attributes and attribute levels

Attribute	# levels	Levels	Pivot/rule
Access mode			
Walk time (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported walk time
Car Travel time (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported drive time
Park and Ride Cost ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported drive to station and park cost
Main mode			
Fare (one-way) ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivotted off respondent provided level
Quickest trip time (mins)	5	−25%, −20%, −15%, −10%, −5%	Value less than average pivot time[a]
Travel time on average (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported trip time
Slowest trip time (mins)	5	10%, 15%, 20%, 25%, 30%	Value more than average pivot time
Slowest trip time %	5	10%, 15%, 20%, 25%, 30%	
Quickest trip time %	5	40%, 45%, 50%, 55%, 60%	
Travel time on average %	5	–	100% − (Travel time on average % + Quickest trip time %)
Number of transfers	3	0, 1, 2	
Frequency of service (mins)	6	Every 10, 20, 30, 40, 50, 60 mins	
Crowding	16	1–16	
Egress mode			
Walk time (mins) to final destination (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported time
Car pick up from stop or station (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported time
Taxi Fare ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported fare

[a]If the reported car trip time is 30 minutes, and the average trip time in a choice set is −25% (hence = 22.5 minutes), then the quickest trip time will be this % less than 22.5 minutes

Table 4 Train attributes and attribute levels

Attribute	# levels	Levels	Pivot/rule
Access mode			
Walk time (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported walk time
Public transport time (including time spent waiting) (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported PT time
Fare (one-way) ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported
Car Travel time (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported drive time
Park and Ride Cost ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported drive to station and park cost
Main mode			
Fare (one-way) ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported fare
Travel time (mins)	5	−25%, −20%, −15%, −10%, −5%	Pivot from reported time
Number of transfers	3	0, 1, 2	
Frequency of service (mins)	6	Every 10, 15, 20, 25, 30, 35 mins	
Crowding	16	1–16	
Egress mode			
Walk time (mins) to final destination	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported time
Public transport time (including time spent waiting) (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported time
Fare (one-way) ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported fare
Metro time (including time spent waiting) (mins)	5	2, 4, 6, 8, 10 mins	
Metro Fare (one-way) ($)	5	$2.60, $2.80, $3.00, $3.20, $3.60	
Car pick up from stop or station (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported time
Taxi Fare ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported fare

Table 5 Metro attributes and attribute levels

Attribute	# levels	Levels	Pivot/rule
Access mode			
Walk time (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported walk time
Public transport time (including time spent waiting) (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported PT time
Fare (one-way) (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported
Car Travel time (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported drive time
Park and Ride Cost ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported drive to station and park cost
Main mode			
Fare (one-way) ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from fare matrix
Travel time (mins)	5	−25%, −20%, −15%, −10%, −5%	Pivot from matrix
Number of transfers	3	0, 1	
Frequency of service (mins)	6	Every 2, 4, 6, 8, 10, 12 mins	
Crowding	16	1–16	
Egress mode			
Walk time (mins) to final destination	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported time
Public transport time (including time spent waiting) (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported time
Fare (one-way) ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported fare
Train time (including time spent waiting) (mins)	5	2, 4, 6, 8, 10 mins	
Train Fare (one-way) ($)	5	$2.60, $2.80, $3.00, $3.20, $3.60	
Car pick up from stop or station (mins)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported time
Taxi Fare ($)	5	−25%, −12.5%, 0%, 12.5%, 25%	Pivot from reported fare

Table 6 Final sample sizes

Commuter

Main mode	Car		Train						Bus_light rail				Total
Time	Peak	Off peak	Peak			Off peak			Peak		Off peak		
Access mode			Walk	PT	Car	Walk	PT	Car	Walk	Car	Walk	Car	
Sample size	80	80	40	40	40	40	40	40	100	40	40	40	**620**
Done	**137**	**34**	**161**	**48**	**87**	**29**	**2**	**3**	**132**	**17**	**20**	0	670
Required	−57	46	−121	−8	−47	11	38	37	−32	23	20	40	−50

In addition to the modal quotas, we introduced (post-pilot) postcode grouping quotas[6] for commuter trips to ensure geographical coverage and to avoid (as occurred in the pilot) a disproportionately higher incidence of trips from the outer west, and especially the north west (Hills District). These are summarised in Fig. 4. The interview locations are shown in Fig. 5. Each week, interviews were undertaken in three or four centres. These interviews are a mixture of telephone screened respondents and individuals screened on-site.

6 Descriptive overview of data

This section provides an overview of the modal profile actually obtained from the effective sample (Table 7), as well as a socio-economic overview of the data (Table 8), with details of the trip characteristics in terms of times, costs, frequency, crowding etc. presented and discussed in Sect. 9. Table 7 is useful in identifying what modes are used for access to the main modes.

An overview of the socioeconomic data for the sample is given in Table 8, with comparisons with the Sydney Household Travel Survey (HTS)[7] in Table 9(a). We do not present the trip attributes in this table. They are given in the next section where we present the final models for each trip segment.

In addition to the socioeconomic descriptors, we also summarise (in Table 9(b)) the extent of data screening prior to model estimation. The commuting segment had an initial maximum sample of 670 respondents, which is a maximum of 4,020 observations for model estimation. We removed data that concerned us. This was necessary in order to remove observations that were subsequently found to be either a serious outlier problem (in terms of trip times in particular where the total trip time was less than 10 minutes and greater than 90 minutes), and/or the open ended comments suggested that respondent participation was motivated by reasons

[6] A question was added into the final CAPI to gather the starting postcode for trips. This enables counts for each postcode at the end of each day to be gathered. A code of −999 is allowed if a respondent really does not know an answer, but these should be minimal.

[7] The HTS data refers to the catchment area and not to the in-scope origin-destination trips reflected in the sample used herein. This survey is undertaken annually by the NSW government and collects data on over 3,000 households.

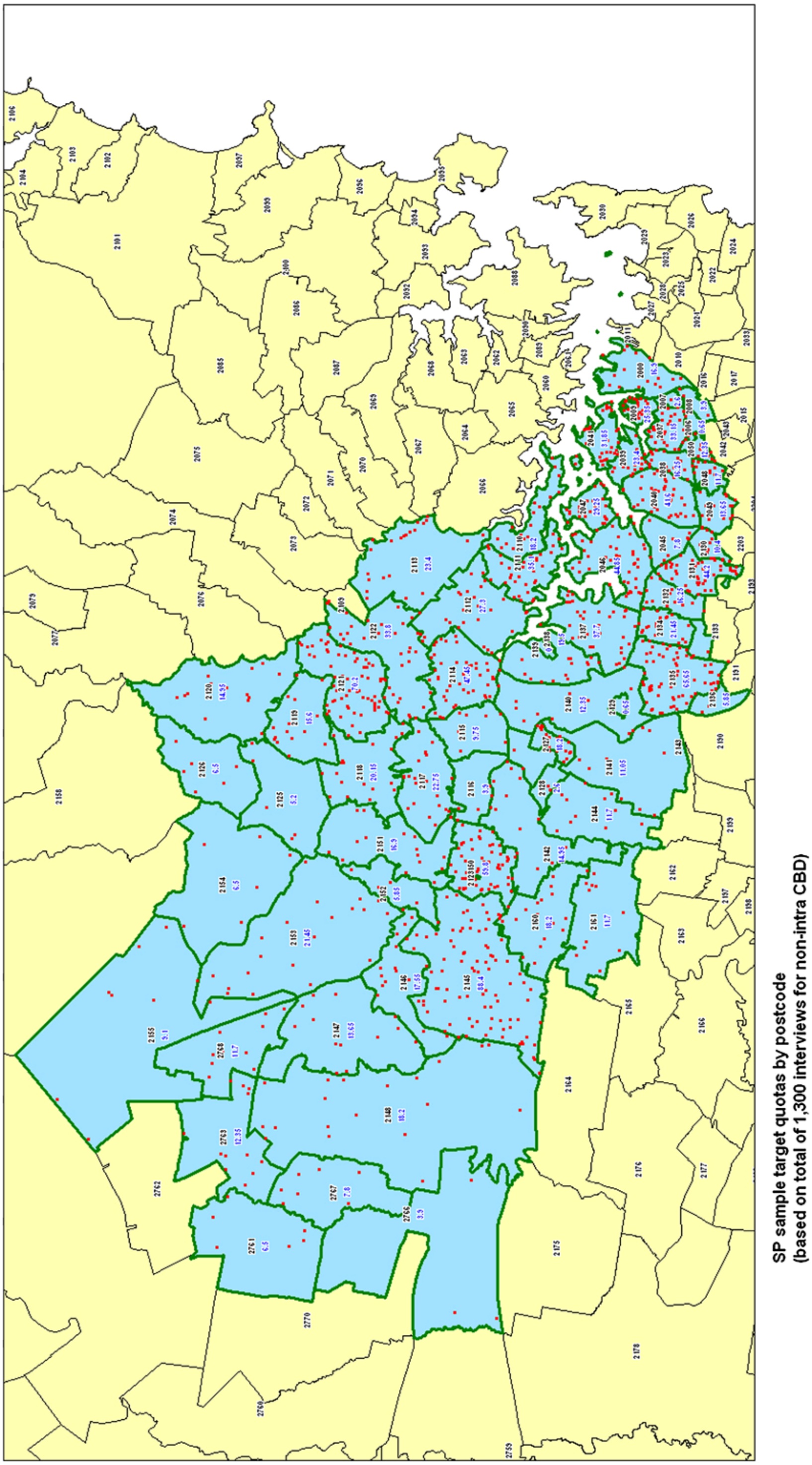


Fig. 4 Postcode quota guidelines for commuting trips

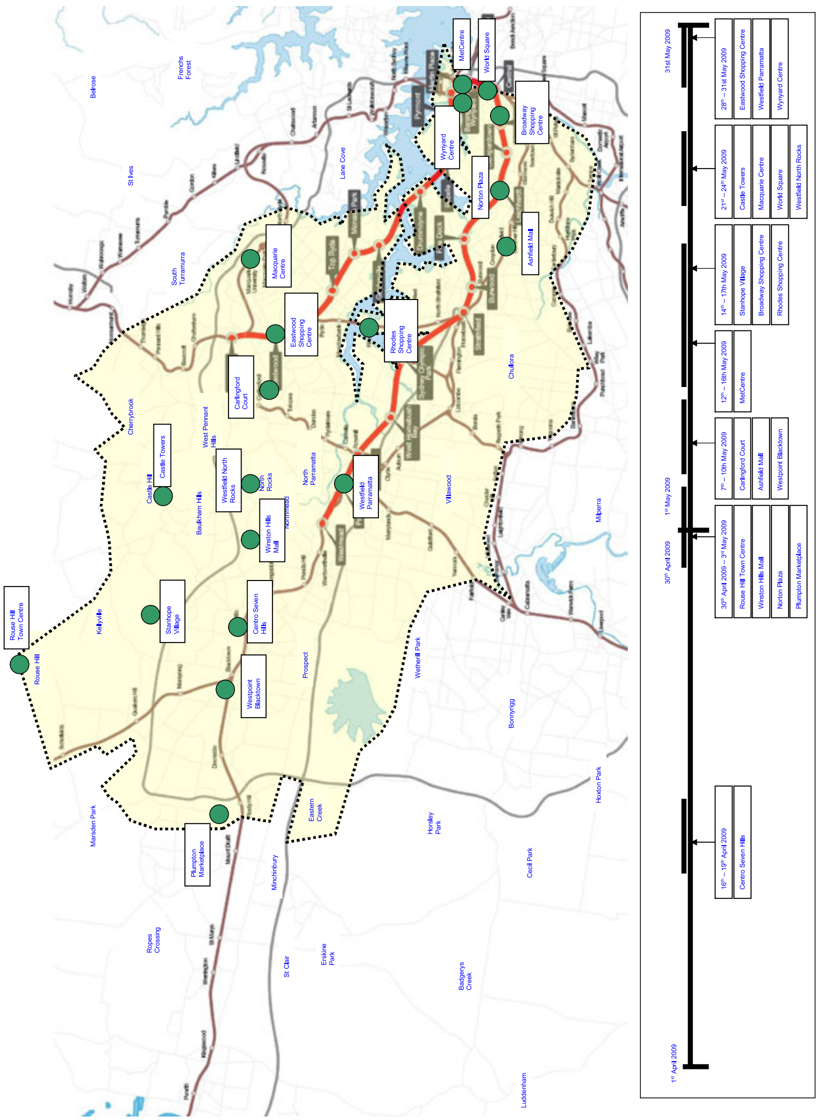


Fig. 5 Interview locations

Table 7 Overview of the modal profile (access, main and egress modes) of the sample

Main purpose	Main mode	Time period	Access mode	Count
Commuter	Bus/LR	Off-peak	Car	0
Commuter	Bus/LR	Off-peak	Walk	20
Commuter	Bus/LR	Peak	Car	17
Commuter	Bus/LR	Peak	Walk	132
Commuter	Car	Off-peak		34
Commuter	Car	Peak		137
Commuter	Train	Off-peak	Car	3
Commuter	Train	Off-peak	PT	2
Commuter	Train	Off-peak	Walk	29
Commuter	Train	Peak	Car	87
Commuter	Train	Peak	PT	48
Commuter	Train	Peak	Walk	161

that suggest a strategic bias in responses (e.g., totally opposed to the government spending money on a metro). The short travel context caused problems with some parameter estimates, given little variation in trip times for such trips. Trips in excess of 90 minutes are very infrequent and these were deemed to be outliers. The total sample for model estimation was reduced to 524 respondents or 3,144 observations.

7 The modelling approach

The model results are obtained from an error components (EC) model. Let U_{nsj} denote the utility of alternative j obtained by respondent n in choice situation s. The utility U_{nsj} is partitioned into three components, an observed component of utility, V_{nsj}, an unobserved component of utility, η_{nsj}, and an unobserved component, ε_{nsj} where the latter is identically and independently extreme value type 1 (EV1) distributed Type I. Under this representation, utility may be written as:

$$U_{nsj} = V_{nsj} + \eta_{nsj} + \varepsilon_{nsj}. \tag{1}$$

As is common practice, the observed component of utility is assumed to be described by a linear relationship of observed attribute levels of each alternative, x, and their corresponding weights (parameters), β. In the *multinomial logit* (MNL) model, the parameter weights for each attribute are invariant over respondents, such that the observed component of utility can be represented as:

$$V_{nsj} = \sum_{k=1}^{K} \beta_{jk} x_{nsjk}. \tag{2}$$

The second component of overall utility, which we have designated η_{nsj}, consists of a set of parameters, θ_{ji} and variables z_{nsji}. Unlike the observed component of utility

Table 8 Socioeconomic profile of data

Age	Commuter
24 or under	72
25 to 34	180
35 to 44	162
45 to 54	146
55 to 64	90
65 and over	19
Refused	0
Total	669
Income	**Commuter**
Under $10K	38
$10K–$15K	21
$15K–$20K	22
$20K–$30K	39
$30K–$40K	73
$40K–$50K	98
$50K–$60K	66
$60K–$80K	116
$80K–$100K	89
$100K–$120K	44
$120K–$150K	24
$150K–$200K	11
Over $200K	11
Refused	18
Total	670
Gender	**Commuter**
Male	314
Female	356
Total	670

however, z_{nsji} is either known, partially known or completely unknown to the analyst whilst θ_{ji} is simply not modelled at all. Given that θ_{ji} and z_{nsji} are either unobserved or not modelled by the analyst, η_{nsj} is typically assumed to be randomly distributed over the population, and although not necessary, to be generic across subsets of alternatives. This component of utility may be written as:

$$\eta_{nsj} = \sum_{i=1}^{i} \theta_{nji} z_{nsji}. \tag{3}$$

Table 9 (a) Comparison of HTS segments to sample segments by proportion. (b) Summary of data removed in preparation for model estimation

(a)

	Commuter	
	HTS	Sample
Personal Gross Income		
Less than 10K	0.078	0.057
10–20K	0.123	0.064
20–30K	0.224	0.058
30–60K	0.295	0.355
60K +	0.281	0.466
Gender		
Male	0.564	0.467
Female	0.436	0.533
Number of Vehicles in Household		
0	0.065	0.106
1	0.327	0.427
2	0.348	0.348
3+	0.260	0.119
Mode Usage		
Car	0.748	0.255
Train	0.165	0.493
Bus	0.086	0.252

(b)

Sample	Collected	Removed due to				Total
		Trip $<$ 10 mins	Trip $>$ 90 mins	Sub total	Other reasons[a]	
Commuter	670 (4,020)	11 (66)	110 (660)	121 (726)	25 (150)	146 (876)
Total	2108 (11,034)	32 (192)	328 (1,968)	360 (2,160)	70 (420)	430 (2,580)

[a]Respondents removed due to providing unrealistic data (36); the sampled trip was not deemed an eligible trip (21); the respondent made a political remark about the Metro (15 respondents); or the surveyed trip involved transporting work equipment (8). Numbers in brackets refer to number of cases (i.e., respondents by choice six scenarios)

Assuming the analyst knows *a priori* all the attributes in z_{nsji}, it is possible for z_{nsji} to enter into (1) such that $x_{nsjk} = z_{nsji}$. In such an instance, (1) will collapse to (4)

$$U_{nsj} = V_{nsj} + \varepsilon_{nsj}. \tag{4}$$

Equation (5) represents the most common form of utility representation within the literature. In some instances, the analyst may have no knowledge of z_{nsji}. In such instances, it is possible without loss of generality, to assume that $z_{nsji} = 1$, in which case (1) becomes:

$$U_{nsj} = V_{nsj} + \varpi_{nsj}, \tag{5}$$

where $\varpi_{nsj} = (\theta_{nsj}, \varepsilon_{nsj})$ and θ_{nsj} has some distribution over sampled individuals and alternatives. Equation (5) represents the EC model. Assuming that the ε_{nsj}'s are EV1 IID and that preferences are homogenous within the sampled population, the probability, P_{nsj}, that respondent n chooses alternative j in choice situation s is given by the MNL model, (6)

$$P_{nsj} = \frac{\exp(V_{nsj})}{\sum_{i \in J_{ns}} \exp(V_{nsi})}. \tag{6}$$

We do not report the results for MNL models, but rather for models estimated assuming the panel formulation of the EC including error components. This model, whilst more complex than the MNL model, is equivalent to the Nested Logit (NL) model in that it accounts for differences in error variances for the different alternatives, but has the added benefit of taking into account the pseudo panel nature of SC data (something that the NL model does not do). As such, the model outputs may be used in exactly the same way as those obtained from a NL model; however the parameters obtained from the model are less likely to be biased as a result of respondents completing more than one choice task. In the formulation of the EC model, the marginal probability of observing a sequence of choices, s_n, is what is modelled, as opposed to the within choice set probabilities themselves. This sequence of probabilities made by respondent n in s choice situations is given as:

$$P(s_n) = \int_{\eta_{sp}} \prod_{s=1}^{S_n} \frac{\exp(\beta_i + \beta x_{si} + \eta_{sp,i})}{\sum_j (\beta_j + \beta x_{sj} + \eta_{sp,j})} \varphi(\eta_{sp}|\sigma^2) d\eta_{sp}, \tag{7}$$

where $\eta_{sp,j} = 0$ for at least one alternative j.

The integrals in (7) require, in order for the probabilities to be calculated, simulation over the random parameters and error terms. This simulation requires that draws be taken over the random parameter (and/or error component) distributional space. For convenience, and to ensure coverage of the entire space of the parameter distributions, typically researchers rely on quasi-random Monte Carlo methods in order to take the simulated draws (see e.g., Bhat 2001, 2003; Hess et al. 2005; Train 2003). The expected probabilities derived from (7) are then used to calculate the log-likelihood function for the model.[8]

In the models reported below, we used 500 Halton draws in a simulated log-likelihood function. The error components that capture influences that are related to alternatives in contrast to attributes are included by constructing a set of independent individual terms, E_{im}, $m = 1, \ldots, M \sim N[0, 1]$ that can be added to the utility functions. This device allows us to create what amounts to a random effects model and, in addition, a very general type of nesting of alternatives. Let θ_m be the scale parameter

[8] All models were estimated using Nlogit4.

(standard deviation) associated with these effects. Then, each utility function can be constructed as:

$$U_{ijt} = \alpha_{ji} + \beta_i' x_{jit} + (\text{any of } \theta_1 E_{i1}, \theta_2 E_{i2}, \ldots, \theta_M E_{iM}). \tag{8}$$

Consider, for example, a four outcome structure

$$\begin{aligned} U_{i1t} &= V_{i1t} + \theta_1 E_{i1} + \theta_2 E_{i2}, \\ U_{i2t} &= V_{i2t} + \theta_2 E_{i2}, \\ U_{i3t} &= V_{i3t} + \theta_1 E_{i1} + \theta_3 E_{i3}, \\ U_{i4t} &= V_{i4t} + \theta_4 E_{i4}. \end{aligned}$$

Thus, U_{i4t} has its own uncorrelated effect, but there is a correlation between U_{i1t} and U_{i2t} and between U_{i1t} and U_{i3t}. This example is fully populated, so the covariance matrix is block diagonal with the first three freely correlated. The model might usefully be restricted in a specific application.

8 The modelling results

In this section we present the final error component logit model, arrived at after extensive investigation of alternative specifications.[9]

An important point to be aware of when evaluating the findings is that the catchment area was defined to ensure that the metro was able to be evaluated as a physically available alternative, with a time-cost profile that was deemed sensible in satisfying in-scope eligibility. Hence the sample has selectivity bias in that *the choices are conditional on metro being deemed a considered alternative in the choice set*. In contrast, all other means of transport (apart from the mode actually chosen in the recently sampled trip for an in-scope interview), may or may not be in the choice set, depending of the responses of respondents. Consequently it will come as no surprise that the current modal share is not a reproduction of the wider modal share of the Sydney metropolitan area or of those who live and travel from a location within the catchment area.

The pre-defined catchment was assumed to be the locality where eligible trips would commence; however many trips in the catchment area would not have a destination that would allow the Metro to be a candidate alternative; hence the use of population data such as that provided by the NSW Transport Data Centre for exogenous weighting turned out to be of little value since the data assumes that everyone who is a ‘Residents of occupied private dwellings in the CATCHMENT’ (defined by

[9] A an aside, after finalising the models we ran models in which we included a dummy variable to distinguish between observations recruited via the phone screener and those recruited via the shopping centres. In all cases this dummy variable was not statistically significant, and indeed t-ratios were significantly less than 1.0. Generic Public transport recruitment type dummy and alternative-specific public transport recruitment type dummy variables were evaluated. Importantly it should be noted that screening for eligibility at shopping centres was based on the exact same screener as used on the phone.

a list of postcodes in Fig. 4) was a candidate potential users of Metro.[10] Given that it is unlikely that one would be able to extract the relevant sub-sample from the HTS survey (given the physical catchment area), we have chosen to focus on the unweighted models, which provide parameter estimates that are a true reflection of the preferences of the sampled respondents for whom the Metro is a genuinely relevant alternative. A descriptive profile of the data used in model estimation is given in Table 10.

The final model is presented in Table 11. This was selected from a large number of models. The overall goodness-of-fit on accepted criteria is impressive (rho^2 of 0.6075). Some parameter estimates are alternative-specific (e.g., car costs and times), and some are generic to a subset of modes (e.g., train and metro or all public transport modes). There are no fully generic parameter estimates (i.e., across all modes), suggesting that there are mode-specific sensitivities on many influencing attributes. For example, the mean parameter estimates (or marginal disutilities) for public transport fares are −0.4345 for bus, and −0.3994 for train and metro, and all are statistically significant (lowest t-value is −5.21).

Taking a closer look at each explanatory variable within class, we see that for access, park or kiss and ride time is the most statistically significant influence on main mode choice for train and metro, followed by the use of a public transport mode (in terms of both time and cost). Note that for the main public mode, it is feasible to access it by either bus, train or the proposed metro. Walk time to a public transport mode was marginally significant. The access public mode cost is also statistically significant, so overall the access trip does have a strong role to play in establishing the preference for public transport. There is clear evidence of a high degree of sensitivity to travel time to drive to, and park near or be dropped off (kiss n ride), a train or proposed metro station, as well as using public transport to access the train or metro. Thus station spacing can have a significant influence on the role of this attribute, given the marginal disutility weight.

The crowding attributes associated with public transport are statistically significant. The number standing is defined as a quadratic interacted with in-vehicle travel time for the main mode, and the proportion seated is defined by the natural logarithm transformation interacted with the main mode in-vehicle travel time. The nonlinearity is intuitively plausible and is very useful in establishing the willingness to pay to get a seat and to stand under various loading scenarios by time of day (or time slice during a particular period such as the peak period) and trip length. The willingness to pay to obtain a seat or to avoid standing is set out in the next section for a range of scenarios. The (dis)utility function for the number standing and the proportion seated are respectively:

$$\text{Utility (\# standing)} = -0.00000147 \times \text{\#standing}^2 \times \text{in-vehicle time},$$

[10] Within the catchment area, the population weights are based on estimated resident population (ERP) as at 30 June 2006. The HTS data is pooled data from 5 waves—02/03 to 06/07 HTS weighted to 30 June 2006 population. We did however investigate weighting with the unlinked trip data supplied by TDC from the Household travel survey (HTS) to identify population modal shares on existing modes and we re-weighted on this criterion. Unlinked trip data was selected in contrast to linked trip data, given that the latter has a priority mode assignment that is not appropriate for the exogenous weighting of relevance herein. The estimated models (available on supplementary material) had a significantly worse overall goodness of fit on log-likelihood as well as producing WTP measures with confidence limits that were unacceptable.

Table 10 Descriptive profile of attributes and choices made: commuter segment. (Mean and standard deviation except for choice shares.) 3,144 observations from 524 respondents, all data is based on sample (given in [ ] except where full sample) who had that attribute in their trip choice set profile

Attribute	Car	Bus	Train	Metro
Choice shares	10.66	8.08	18.0	63.3
#observations (#respondents*#choice scenarios) (i.e., actual number of choice scenarios where this alternative is present)	1,068	972	1,830	3,144
Access walk time (mins)	–	5.42 (2.94) [911]	9.19 (5.15) [1121]	7.325 (4.63) [1188]
Crowding defined as (number standing)^2	–	141.76 (227.9)	3008 (4468)	9003 (6813)
Crowding defined as (LN[Proportion seated])	–	−0.215 (0.352)	−0.211 (0.355)	−0.038 (0.10)
Egress walk time (mins)	7.57 (4.35) [1044]	6.34 (4.04) [954]	7.29 (4.09) [1648]	7.57 (4.35) [2530]
Frequency of service (mins)	–	35.12 (16.73)	34.85 (27.76)	8.41 (4.58)
Fare ($)	–	3.50 (1.88)	4.24 (1.84)	3.77 (1.55)
Main mode travel time (mins)	38.05 (16.21)	34.95 (15.52)	33.07 (13.6)	16.23 (8.42)
Number of transfers (0,1,2)	–	0.994 (0.827)	1.03 (0.82)	0.493 (0.50)
Egress time by public transport (mins)	–	–	13.74 (5.68) [168]	15.71 (8.23) [591]
Egress fare ($)	–	–	2.90 (3.21) [168]	3.15 (1.77) [591]
Running cost ($)	2.02 (1.16)	–	–	–
Toll cost ($)	4.26 (2.89)	–	–	–
Parking cost ($)	12.37 (8.97)	–	–	–
Access mode—park n ride (mins)	–	9.57 (6.73) [47]	9.31 (5.34) [324]	11.45 (6.57) [570]
Access mode—kiss n ride (mins)	–	8.50 (3.08) [14]	7.81 (5.87) [128]	8.18 (4.15) [206]
Access mode—park n ride cost ($)	–	4.42 (2.73) [47]	3.72 (2.87) [324]	4.87 (4.61) [570]
Access mode—public transport time (mins)	–	–	13.61 (7.19) [248]	13.14 (7.73) [933]
Access mode—public transport fare ($)	–	–	2.13 (0.89) [248]	2.51 (1.18) [933]
Proportion of sample whose chosen SC mode is their reference trip mode	0.736	0.840	0.843	–

Note. Choice shares refers to percentage choosing each mode across all choice sets

Table 11 Commuter model

Attribute	Modes	Par	(*t*-ratio)
Access			
Walk (mins)	All PT	−0.04287	(−1.80)
Car time (Park and ride and Kiss and ride) (mins)	All PT	−0.00152	(−5.10)
PT time (metro, bus and train) (mins)	Train, Metro	−0.03501	(−1.87)
PT cost (metro, bus and train) ($)	Train, Metro	−0.00145	(−7.84)
Public transport Main Mode Attributes			
Mode-specific constant	Bus only	−4.7661	(−3.06)
Mode-specific constant	Train only	−4.2748	(−2.91)
Mode-specific constant	Metro only	−3.3051	(−2.34)
Fare ($)	Bus only	−0.4345	(−5.21)
Fare ($)	Train and Metro	−0.3994	(−7.60)
Expected travel time (mins)[a]	All PT	−0.0537	(−7.50)
Crowding ([number standing]$^2 \times$ travel time)	All PT	−0.00000147	(−3.84)
Crowding (LN[Proportion seated] × travel time)	All PT	−0.01978	(−3.96)
[Frequency of service]^2(every number of mins)	All PT	−0.00031	(−6.47)
Number of Transfers	Bus only	−0.2931	(−2.32)
Number of Transfers	Train only	−0.1615	(−1.80)
Number of Transfers	Metro only	−0.2111	(−1.90)
Car Main Mode			
Petrol and Toll Cost ($)	Car only	−0.1671	(−4.45)
Parking Cost ($)	Car only	−0.1518	(−8.43)
Travel Time (mins)	Car only	−0.06508	(−4.48)
Egress			
Walk (mins)	All modes	−0.0755	(−3.19)
PT cost (metro, bus and train) ($)	All modes	−0.00145	(−7.84)
Socio-demographic variables			
Current Bus User (1, 0)	Bus only	1.1163	(2.07)
Current Train User (1, 0)	Train only	0.2952	(0.84)
Curent Car User (1, 0)	Car only	4.7780	(4.46)
Error Components			
Car, Bus		1.164	(3.44)
Bus, Train		1.663	(12.84)
Bus, Train, Metro		2.417	(7.02)
Model Fits			
LL(β)		−1710.90	
LL(0)		−4358.509	
LL(ASC)		−2278.538	
Rho^2 (relative to zero parameters)		0.6075	
Rho^2 (relative to ASCs)		0.2491	
Number of respondents		524	
Number of observations		3,144	

[a]Defined as the probability weighted travel time for each of the three times in the choice scenarios for each alternative

$$\text{Marginal Utility (\# standing)} = 2 \times -0.00000147 \times \text{\#standing} \times \text{in-vehicle time},$$

$$\text{Utility (proportion seated)} = -0.01978 \times \text{Ln(proportion seated)} \times \text{in-vehicle time},$$

$$\text{Marginal Utility (probability of getting a seat)}$$
$$= (-0.01978 \times \text{in-vehicle time})/(\text{proportion seated}).$$

Of particular note is that the mode-specific constants for all public transport modes are negative (relative to the car-specific constant set to 0.0) and statistically significant. What this suggests is that, after accounting for the differential and correlated variances associated with the unobserved influences associated with each alternative, as well as the rich specification of factors that really do matter to travellers who are in-scope, that there remain significant unobserved influencing effects on average. The metro has the least negative mode-specific constant (with bus the most negative) indicating its preference over train and bus after accounting for service and cost levels. In application one can construct a modified metro-specific constant from all those attributes in the model that are statistically significant that the analysts may choose not to incorporate in their application model. One would have to define a specific (fixed) level for each of these attributes, and use that to obtain the additive metro-specific constant.

We have introduced dummy variables to represent the mode currently being used. These variables are measures of inertia and suggest, all other factors remaining unchanged, that an individual is more likely to choose the alternative they currently use. Since many trips are essentially habitual (in contrast to variety seeking), this is an important conditioning effect. The dummy variables for the revealed preference (RP) modes car and bus are statistically significant, but not for train. This might suggest that the proposed metro will appeal mostly to existing train users. The strongest RP inertia effect is for car, with a parameter estimate of 4.7780, in comparison to the bus estimate of 1.1163.

We see a very statistically significant public transport frequency attribute (an impressive t-ratio of -6.47). Frequency is defined as a quadratic of minutes between services, allowing for the marginal disutility to vary by headway. The negative parameter estimate indicates all other influences remaining constant, that an increased frequency associated with reduced time between services, will reduce the marginal disutility of public transport, and increase the probability of choosing public transport. The willingness to pay for increased frequency is statistically significant and non-marginal.

Finally, the presence of a transfer within public transport is negative as expected, and statistically significant for bus, and marginally significant for train and metro. Thus relative to metro, the transfer penalty is lower for train and higher for bus.

9 Willingness to pay outputs

The key behavioural outputs are the estimates of various willingness to pay (WTP) measures such as the value of travel time savings (incorporating the distribution of travel time reliability), value of getting a seat and avoiding standing, and the value placed on increased service frequency. A willingness to pay indicator is a measure of how much an individual is willing to pay to avoid or receive a unit of a specific attribute, holding income and tastes constant. We provide a full set of WTP outputs, and where appropriate, contrast the evidence with benchmark estimates.

We express the WTP in dollars per minute or per trip. The values of travel time savings (VTTS) can be converted to minutes (or hours) per dollar (i.e., WTP in time units); the values attached to seating and standing refer to a trip and so to convert from $/trip to minutes per trip we would need to know the travel time of the trip.

The Krinsky and Robb (1986, 1990) method is used to obtain mean and median WTP along with 95 percent confidence intervals. The procedure involves a number of steps summarised below:

1. Estimate the WTP model of interest.
2. Obtain the vector of parameter estimates β' and the variance-covariance (VCV) matrix $V(\beta')$ (these are saved at model estimation).
3. Calculate the Cholesky decomposition, C, of the VCV matrix $CC' = V(\beta')$.
4. Randomly draw from the standard normal distribution, a vector x with k independent elements.
5. Calculate a new vector of parameter estimates Z such that $Z = \beta' + C'x$.
6. Use the new vector Z to calculate the WTP measures of interest.
7. Repeat steps 4, 5, and 6 $N(>= 5000)$ times to obtain an empirical distribution of WTP.
8. Sort the N values of the WTP function in ascending order.
9. Obtain a 95% confidence interval around the mean/median by dropping the top and bottom 2.5% of the observations.

The Commuter WTP estimates are presented in Table 12. We have included the output from the Krinsky and Robb procedure. The mean values of travel time savings (VTTS) for the main modes are $7.68 per hour for bus, $8.22 for train and metro, and $24.78 for car. The weighted average VTTS is $9.95 person hour.

Importantly the VTTS for road modes (i.e., car and bus) include an allowance for travel time uncertainty on repeated trips over a week, and so the VTTS is based on a recognition that the trip time in a main mode can vary over repeated trips for car and bus according to the probability profile provided for quickest trip time, travel time on average and slowest trip time. So we might expect the median estimates of VTTS to be higher than previous studies, since we have removed the confoundment between trip specific travel time and reliability of time.

Polydoropoulou and Ben-Akiva (2001) investigated *the probability of getting a seat* (out of 10 times) for bus and mass transit, using a nested logit specification on joint RP and SP data (essentially ignoring the panel nature of the data) for all trip purposes. The variable was treated as a linear effect. The authors found it to be a statistically significant effect, but did not report a WTP. They provide two cost variables (fare for those with no cars, and fare for those who own one or more cars). The WTP calculated from the data (in Israeli shekel (NIS) 1999) is 36.39 NIS and 46.61 NIS respectively for 0 and 1-plus cars owned. This suggests that if a traveller can always get a seat (10 out of 10 times), then they are willing to pay up to 36.39 NIS to 46.61 NIS for this situation, which translates into (given the linear definition),

Table 12 Commuter willingness to pay ($/min except transfers)

					95% confidence interval	
		Average	Median	Std Dev.	Lower	Upper
Access Walk Time	Bus	$0.102	$0.098	$0.061	$−0.007	$0.236
	Train, Metro	$0.112	$0.107	$0.068	$−0.007	$0.259
Access Public Transport Time	Train, Metro	$0.089	$0.090	$0.049	$−0.006	$0.180
Access Drive Time	Bus	$0.004	$0.003	$0.001	$0.002	$0.006
	Train, Metro	$0.004	$0.004	$0.001	$0.002	$0.006
In-vehicle Time	Bus	$0.128	$0.123	$0.031	$0.082	$0.205
	Train, Metro	$0.137	$0.134	$0.027	$0.094	$0.201
	Car	$0.413	$0.391	$0.143	$0.202	$0.730
Number of Transfers ($/transfer)	Bus	$0.702	$0.678	$0.336	$0.134	$1.413
	Train	$0.411	$0.413	$0.240	$−0.052	$0.901
	Metro	$0.539	$0.524	$0.294	$0.000	$1.157
Egress Time Walk	Bus	$0.182	$0.175	$0.071	$0.066	$0.342
	Train, Metro	$0.193	$0.189	$0.065	$0.077	$0.336
	Car	$0.486	$0.459	$0.216	$0.161	$1.004

3.639 NIS to 4.661 NIS per trip. If there is a 50 percent chance of getting a seat, then the values per trip are 1.819 NIS and 2.331 NIS.

The direct comparison with our evidence is tricky, since 100 percent seating does not imply a seat 100 percent of the time. We have to look at a scenario such as 'no one standing and plenty of seats available'. Within the range, we have investigated what the best estimates might be when there is a 45 per cent chance that you get a seat, which has a median WTP per trip averaging 30 minutes of 2009$AUD2.76 (or 2009$US2.15) for metro. If we recognise that the study by Polydoropoulou and Ben-Akiva (2001) was undertaken in Tel Aviv 10 years ago, the evidence is not directly comparable. The figure in 1999 for Australia, given the Sydney metro estimate in 2009, would be 1999$Aus2.06[11] (or 1999US1.34) given an exchange rate of $AUD1 = $US0.64518. This is higher than the Tel Aviv evidence in 1999, in the range $US0.439 and $US0.563 (one US dollar was equal to 4.14 NIS in 1999). Another way of comparing the evidence is to express these WTP estimates as a percentage of the prevailing average wage rate. For the Israel study this is 6.4 to 8.2 percent of the average wage rate and for the Australian evidence it is 7.10 percent of the average wage rate, the latter being in the range of the Tel Aviv evidence, which is encouraging.

Pepper et al. (2003) investigated crowding in trains. The numbers and lengths of New Jersey Transit (NJT) commuter trains accessing New York's Penn Station are currently at the limits of available capacity during peak periods, as evidenced by the

[11] Conversion using http://www.rba.gov.au/calculator/calc.go.

significant number of standees at these times. A study was conducted to determine the design of the multilevel coaches, so that they will provide the extra capacity needed and also reflect customers' preferences. The study focused on interior issues, including the seating configuration and seat design, that relate directly to the amount of seated (and standee) capacity on the new coaches. A detailed, computer-based survey was administered to customers to quantify their preferences among key elements of the multilevel concepts, and to estimate their WTP for those elements. The study found that additional seating capacity in the configuration preferred by customers provides a substantial net benefit to NJT passengers, equivalent to *about* \$2.20 *fare value per trip*. The benefits are higher for this application because of the crowded conditions on existing trains. The study also suggests that multilevel coaches and improved interior design have benefits well beyond increased capacity. This mean estimate cannot be compared to our study without knowing the baseline mix of seats and standing capacity, and the actual loadings in terms of those sitting and standing.

Tables 13 and 14 provide detailed scenarios of the WTP associated with the number of standees (Table 13) and the percentage of seats occupied (or the converse as the probability of getting a seat) (Table 14). This format of presentation enables analysts to select a specific standing and seating mix under a range of scenarios for each public transport mode.

A number of WTP measures are non-linear. The two crowding attributes (i.e., number standing and the proportion seated) are respectively specified as a quadratic form and in natural logarithmic form, as well as being interacted with travel time. The frequency attribute is treated as a quadratic. The WTP formulae for these attributes are:

$$\text{WTP (\# standing)} = (2 \times \beta_{\#standing} \times \text{\#standing} \times \text{in-vehicle time})/\beta_{cost},$$

$$\text{WTP (proportion seated)} = (\beta_{seat} \times \text{in-vehicle time})/(\beta_{cost} \times \text{proportion seat}),$$

$$\text{WTP (frequency)} = (2 \times \beta_{frequency} \times \text{frequency})/\beta_{cost}.$$

The WTP associated with the proportion seated (or 1-probability of getting a seat) is a logarithmic functional form, and hence the WTP is based on the equation below (where tt = travel time and x_k is the proportion seated):

$$\text{WTP} = \frac{\frac{d}{dx_k}\beta_k \ln(x_k) tt}{\frac{d}{dx_C}\beta_C x_c} = \frac{\beta_k \frac{1}{x_k} tt}{\beta_C} = \frac{\beta_k tt}{\beta_C x_k}. \tag{9}$$

The WTP associated with getting a seat is [(−0.01978 × in-vehicle time)/(−0.4345 × proportion seated)] for train and metro and [−0.01978 × in-vehicle time/(−0.3994 × proportion seated)] for bus. For example, in Table 13, for metro, if we assume a probability of obtaining a seat per trip of 0.5 per carriage for a 30 minute in-vehicle trip time for the main mode, the mean WTP is \$2.97; hence a commuter is willing to pay on average up to \$2.97 to ensure that a seat is available 50 percent of the time. To ensure a seat is available 75 percent of the time, they are willing to pay on average up to \$6.08. Likewise, the WTP associated with a trip environment in which there are 75 standees at the time of boarding over the 30 minute in-vehicle trip is a very small

Table 13 Commuter Public transport standing WTP

(a) Bus

	95% confidence interval						95% confidence interval				
	Average	Median	Std Dev.	Lower	Upper		Average	Median	Std Dev.	Lower	Upper
# standing	10 mins					# standing	20 mins				
5	$0.00035	$0.00034	$0.00011	$0.00017	$0.00061	5	$ 0.00070	$0.00067	$0.00023	$0.00033	$0.00123
10	$0.00070	$0.00067	$0.00023	$0.00033	$0.00123	10	$ 0.00140	$0.00135	$0.00045	$0.00066	$0.00246
15	$0.00105	$0.00101	$0.00034	$0.00050	$0.00184	15	$ 0.00210	$0.00202	$0.00068	$0.00100	$0.00369
20	$0.00140	$0.00135	$0.00045	$0.00066	$0.00246	20	$ 0.00280	$0.00269	$0.00091	$0.00133	$0.00491
25	$0.00175	$0.00168	$0.00057	$0.00083	$0.00307	25	$ 0.00350	$0.00337	$0.00113	$0.00166	$0.00614
30	$0.00210	$0.00202	$0.00068	$0.00100	$0.00369	30	$ 0.00421	$0.00404	$0.00136	$0.00199	$0.00737
# standing	30 mins					# standing	40 mins				
5	$0.00105	$0.00101	$0.00034	$0.00050	$0.00184	5	$ 0.00140	$0.00135	$0.00045	$0.00066	$0.00246
10	$0.00210	$0.00202	$0.00068	$0.00100	$0.00369	10	$ 0.00280	$0.00269	$0.00091	$0.00133	$0.00491
15	$0.00315	$0.00303	$0.00102	$0.00149	$0.00553	15	$ 0.00421	$0.00404	$0.00136	$0.00199	$0.00737
20	$0.00421	$0.00404	$0.00136	$0.00199	$0.00737	20	$ 0.00561	$0.00539	$0.00181	$0.00266	$0.00983
25	$0.00526	$0.00505	$0.00170	$0.00249	$0.00921	25	$ 0.00701	$0.00674	$0.00227	$0.00332	$0.01229
30	$0.00631	$0.00606	$0.00204	$0.00299	$0.01106	30	$ 0.00841	$0.00808	$0.00272	$0.00398	$0.01474
# standing	50 mins					# standing	60 mins				
5	$0.00175	$0.00168	$0.00057	$0.00083	$0.00307	5	$ 0.00210	$0.00202	$0.00068	$0.00100	$0.00369
10	$0.00350	$0.00337	$0.00113	$0.00166	$0.00614	10	$ 0.00421	$0.00404	$0.00136	$0.00199	$0.00737
15	$0.00526	$0.00505	$0.00170	$0.00249	$0.00921	15	$ 0.00631	$0.00606	$0.00204	$0.00299	$0.01106
20	$0.00701	$0.00674	$0.00227	$0.00332	$0.01229	20	$ 0.00841	$0.00808	$0.00272	$0.00398	$0.01474
25	$0.00876	$0.00842	$0.00283	$0.00415	$0.01536	25	$ 0.01051	$0.01011	$0.00340	$0.00498	$0.01843
30	$0.01051	$0.01011	$0.00340	$0.00498	$0.01843	30	$ 0.01262	$0.01213	$0.00408	$0.00598	$0.02211

Table 13 (*Continued*)

(b) Train/Metro

	95% confidence interval						95% confidence interval				
	Average	Median	Std Dev.	Lower	Upper		Average	Median	Std Dev.	Lower	Upper
# standing	10 mins					# standing	20 mins				
5	$0.00038	$0.00037	$0.00011	$0.00018	$0.00063	5	$ 0.00075	$0.00073	$0.00023	$0.00035	$0.00126
15	$0.00113	$0.00110	$0.00034	$0.00053	$0.00188	15	$ 0.00226	$0.00220	$0.00068	$0.00105	$0.00377
25	$0.00188	$0.00184	$0.00057	$0.00088	$0.00314	25	$ 0.00377	$0.00367	$0.00113	$0.00175	$0.00628
50	$0.00377	$0.00367	$0.00113	$0.00175	$0.00628	50	$ 0.00754	$0.00735	$0.00226	$0.00351	$0.01255
75	$0.00565	$0.00551	$0.00170	$0.00263	$0.00941	75	$ 0.01130	$0.01102	$0.00340	$0.00526	$0.01883
100	$0.00754	$0.00735	$0.00226	$0.00351	$0.01255	100	$0.01507	$0.01470	$0.00453	$0.00701	$0.02510
120	$0.00904	$0.00882	$0.00272	$0.00421	$0.01506	120	$0.01809	$0.01764	$0.00543	$0.00842	$0.03012
# standing	30 mins					# standing	40 mins				
5	$0.00113	$0.00110	$0.00034	$0.00053	$0.00188	5	$ 0.00151	$0.00147	$0.00045	$0.00070	$0.00251
15	$0.00339	$0.00331	$0.00102	$0.00158	$0.00565	15	$ 0.00452	$0.00441	$0.00136	$0.00210	$0.00753
25	$0.00565	$0.00551	$0.00170	$0.00263	$0.00941	25	$ 0.00754	$0.00735	$0.00226	$0.00351	$0.01255
50	$0.01130	$0.01102	$0.00340	$0.00526	$0.01883	50	$ 0.01507	$0.01470	$0.00453	$0.00701	$0.02510
75	$0.01696	$0.01654	$0.00509	$0.00789	$0.02824	75	$ 0.02261	$0.02205	$0.00679	$0.01052	$0.03765
100	$0.02261	$0.02205	$0.00679	$0.01052	$0.03765	100	$0.03015	$0.02940	$0.00906	$0.01403	$0.05020
120	$0.02713	$0.02646	$0.00815	$0.01263	$0.04518	120	$0.03617	$0.03528	$0.01087	$0.01684	$0.06025
# standing	50 mins					# standing	60 mins				
5	$0.00188	$0.00184	$0.00057	$0.00088	$0.00314	5	$ 0.00226	$0.00220	$0.00068	$0.00105	$0.00377
15	$0.00565	$0.00551	$0.00170	$0.00263	$0.00941	15	$ 0.00678	$0.00661	$0.00204	$0.00316	$0.01130
25	$0.00942	$0.00919	$0.00283	$0.00438	$0.01569	25	$ 0.01130	$0.01102	$0.00340	$0.00526	$0.01883
50	$0.01884	$0.01837	$0.00566	$0.00877	$0.03138	50	$ 0.02261	$0.02205	$0.00679	$0.01052	$0.03765
75	$0.02826	$0.02756	$0.00849	$0.01315	$0.04707	75	$ 0.03391	$0.03307	$0.01019	$0.01578	$0.05648
100	$0.03768	$0.03675	$0.01132	$0.01754	$0.06276	100	$0.04522	$0.04410	$0.01359	$0.02104	$0.07531
120	$0.04522	$0.04410	$0.01359	$0.02104	$0.07531	120	$0.05426	$0.05292	$0.01630	$0.02525	$0.09037

Table 14 Commuter Public transport seating WTP

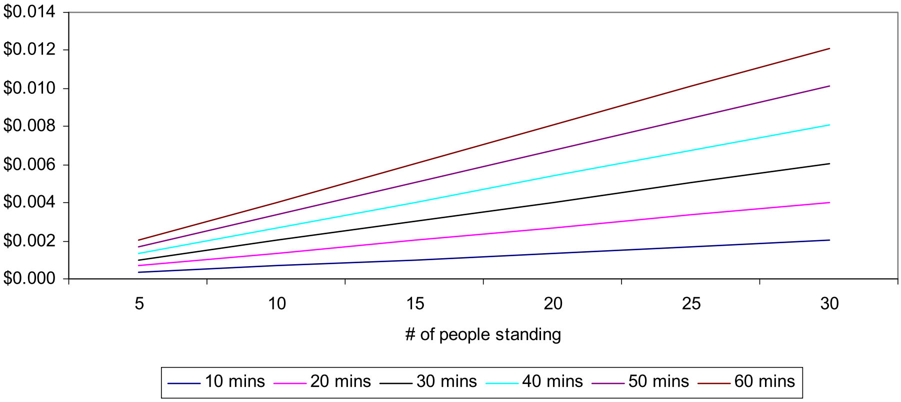
(a) Bus

	95% confidence interval						95% confidence interval				
	Average	Median	Std Dev.	Lower	Upper		Average	Median	Std Dev.	Lower	Upper
Prop. Seated	10 mins					Prop. Seated	20 mins				
0.25	$1.90511	$1.79998	$0.65873	$0.91345	$3.44117	0.25	$3.81023	$3.59997	$1.31745	$1.82689	$6.88234
0.4	$1.19070	$1.12499	$0.41170	$0.57090	$2.15073	0.4	$2.38139	$2.24998	$0.82341	$1.14181	$4.30146
0.55	$0.86596	$0.81817	$0.29942	$0.41520	$1.56417	0.55	$1.73192	$1.63635	$0.59884	$0.83041	$3.12834
0.7	$0.68040	$0.64285	$0.23526	$0.32623	$1.22899	0.7	$1.36080	$1.28570	$0.47052	$0.65246	$2.45798
0.85	$0.56033	$0.52941	$0.19374	$0.26866	$1.01211	0.85	$1.12066	$1.05881	$0.38749	$0.53732	$2.02422
1	$0.47628	$0.45000	$0.16468	$0.22836	$0.86029	1	$ 0.95256	$0.89999	$0.32936	$0.45672	$1.72058
Prop. Seated	30 mins					Prop. Seated	40 mins				
0.25	$5.71534	$5.39995	$1.97618	$2.74034	$10.32351	0.25	$7.62046	$7.19994	$2.63491	$3.65378	$13.76468
0.4	$3.57209	$3.37497	$1.23511	$1.71271	$6.45219	0.4	$4.76279	$4.49996	$1.64682	$2.28362	$8.60292
0.55	$2.59788	$2.45452	$0.89826	$1.24561	$4.69250	0.55	$3.46384	$3.27270	$1.19769	$1.66081	$6.25667
0.7	$2.04119	$1.92855	$0.70578	$0.97869	$3.68697	0.7	$2.72159	$2.57141	$0.94104	$1.30492	$4.91596
0.85	$1.68098	$1.58822	$0.58123	$0.80598	$3.03633	0.85	$2.24131	$2.11763	$0.77497	$1.07464	$4.04843
1	$1.42884	$1.34999	$0.49405	$0.68508	$2.58088	1	$ 1.90511	$1.79998	$0.65873	$0.91345	$3.44117
Prop. Seated	50 mins					Prop. Seated	60 mins				
0.25	$9.52557	$8.99992	$3.29364	$4.56723	$17.20584	0.25	$11.43068	$10.79990	$3.95236	$5.48068	$20.64701
0.4	$5.95348	$5.62495	$2.05852	$2.85452	$10.75365	0.4	$7.14418	$6.74994	$2.47023	$3.42542	$12.90438
0.55	$4.32980	$4.09087	$1.49711	$2.07601	$7.82084	0.55	$5.19577	$4.90905	$1.79653	$2.49122	$9.38501
0.7	$3.40199	$3.21426	$1.17630	$1.63115	$6.14494	0.7	$4.08239	$3.85711	$1.41156	$1.95738	$7.37393
0.85	$2.80164	$2.64704	$0.96872	$1.34330	$5.06054	0.85	$3.36197	$3.17644	$1.16246	$1.61196	$6.07265
1	$2.38139	$2.24998	$0.82341	$1.14181	$4.30146	1	$ 2.85767	$2.69998	$0.98809	$1.37017	$5.16175

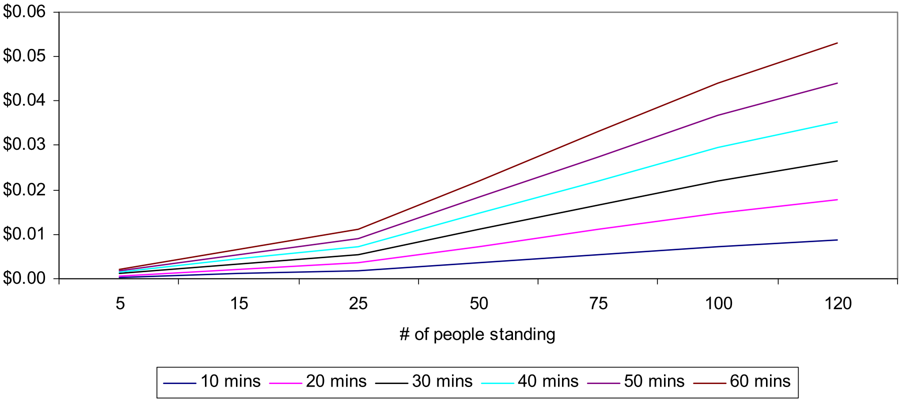
Table 14 (*Continued*)

(b) Train/Metro

	95% confidence interval						95% confidence interval				
	Average	Median	Std Dev.	Lower	Upper		Average	Median	Std Dev.	Lower	Upper
Prop. Seated	10 mins					Prop. Seated	20 mins				
0.25	$2.02593	$1.99416	$0.58157	$1.00387	$3.33226	0.25	$4.05186	$3.98832	$1.16315	$2.00774	$6.66452
0.4	$1.26621	$1.24635	$0.36348	$0.62742	$2.08266	0.4	$2.53241	$2.49270	$0.72697	$1.25484	$4.16532
0.55	$0.92088	$0.90644	$0.26435	$0.45630	$1.51466	0.55	$1.84176	$1.81287	$0.52870	$0.91261	$3.02933
0.7	$0.72355	$0.71220	$0.20770	$0.35853	$1.19009	0.7	$1.44709	$1.42440	$0.41541	$0.71705	$2.38019
0.85	$0.59586	$0.58652	$0.17105	$0.29526	$0.98008	0.85	$1.19172	$1.17304	$0.34210	$0.59051	$1.96015
1	$0.50648	$0.49854	$0.14539	$0.25097	$0.83306	1	$ 1.01297	$0.99708	$0.29079	$0.50194	$1.66613
Prop. Seated	30 mins					Prop. Seated	40 mins				
0.25	$6.07780	$5.98248	$1.74472	$3.01161	$9.99678	0.25	$8.10373	$7.97664	$2.32629	$4.01548	$13.32904
0.4	$3.79862	$3.73905	$1.09045	$1.88226	$6.24799	0.4	$5.06483	$4.98540	$1.45393	$2.50968	$8.33065
0.55	$2.76263	$2.71931	$0.79305	$1.36891	$4.54399	0.55	$3.68351	$3.62575	$1.05741	$1.82522	$6.05865
0.7	$2.17064	$2.13660	$0.62311	$1.07558	$3.57028	0.7	$2.89419	$2.84880	$0.83082	$1.43410	$4.76037
0.85	$1.78759	$1.75955	$0.51315	$0.88577	$2.94023	0.85	$2.38345	$2.34607	$0.68420	$1.18102	$3.92031
1	$1.51945	$1.49562	$0.43618	$0.75290	$2.49919	1	$ 2.02593	$1.99416	$0.58157	$1.00387	$3.33226
Prop. Seated	50 mins					Prop. Seated	60 mins				
0.25	$10.12966	$9.97080	$2.90786	$5.01935	$16.66130	0.25	$12.15559	$11.96496	$3.48944	$6.02322	$19.99356
0.4	$6.33104	$6.23175	$1.81741	$3.13710	$10.41331	0.4	$7.59724	$7.47810	$2.18090	$3.76451	$12.49597
0.55	$4.60439	$4.53218	$1.32176	$2.28152	$7.57332	0.55	$5.52527	$5.43862	$1.58611	$2.73783	$9.08798
0.7	$3.61774	$3.56100	$1.03852	$1.79263	$5.95046	0.7	$4.34128	$4.27320	$1.24623	$2.15115	$7.14056
0.85	$2.97931	$2.93259	$0.85525	$1.47628	$4.90038	0.85	$3.57517	$3.51911	$1.02630	$1.77154	$5.88046
1	$2.53241	$2.49270	$0.72697	$1.25484	$4.16532	1	$ 3.03890	$2.99124	$0.87236	$1.50581	$4.99839



(a)



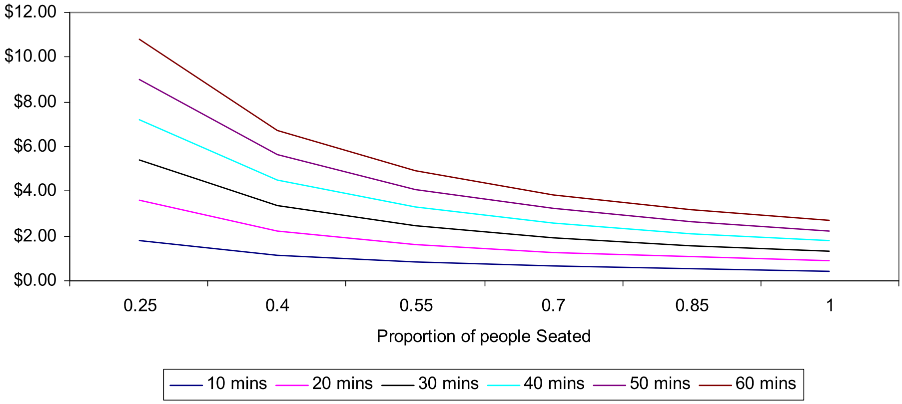
(b)

Fig. 6 (**a**) Commuter Median Willingness to Pay to avoid standing in a bus. (**b**) Commuter Median Willingness to Pay to avoid standing in metro/train (*note that the kink is a result of the non-linear increase in the number of standees showing on the X axis). (**c**) Commuter Median Willingness to Pay for proportion seated on a bus. (**d**) Commuter Median Willingness to Pay for proportion seated on a metro/train

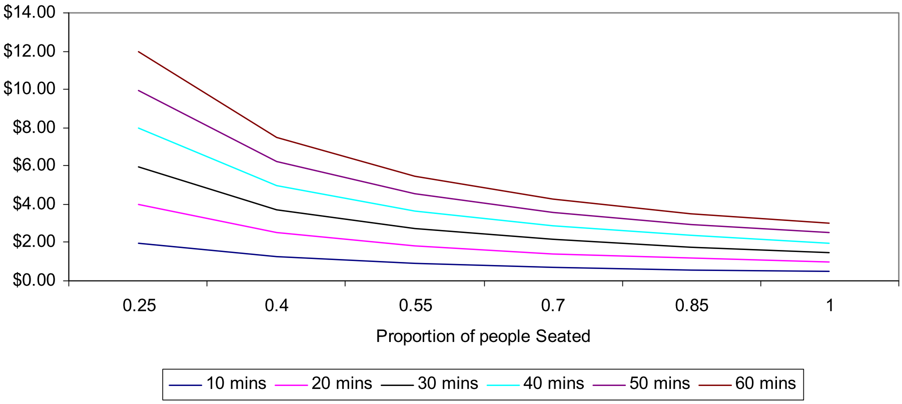
amount, namely 1.7 cents, suggesting that all the focus should be on the WTP to get a seat.

In order to implement the 'crowding' WTP, one would need to identify the number of commuters using the metro in any time slice, and then establish the number of standees and those sitting, given the configuration of the carriages. For example, suppose the demand estimates suggest that, given 70 seats per metro carriage, there are 140 passengers per carriage. Then the probability of getting a seat is 0.5. The WTP associated with this scenario for a 30 minute trip is $2.97 per passenger per trip. Based on this logic, we can graph (Figs. 6(c), (d)) the WTP associated with getting a seat as well as a WTP associated with the number standing (Figs. 6(a), (b)).

Importantly, the perception of crowding is built into this valuation (given the pictures shown to respondents). Suppose that the average travel time in the metro is 30 minutes, then the 'crowding penalty' in cents per minute in the example above



(c)



(d)

Fig. 6 (*Continued*)

for a probability of 0.5 of getting a seat is 9.9 cents. Given the exchange rate of \$AUD1 = \$0.48 pounds, this is equivalent to 4.76 pence per minute, which is in the range of the UK London evidence. As a mark up on the VTTS per person hour, this adds 0.736 on the metro VTTS.

The graphs in Figs. 6(a)–(d) show the shape of the function for the relationship between WTP and (i) number standing and in-vehicle time; and (ii) the proportion seated (= 1-probability of getting a seat) and in-vehicle time. The service frequency distribution by headway is summarised in Table 15.

10 Conclusions

The study has focussed on obtaining new empirical evidence on the factors influencing the choice of mode of transport for commuting trips in Sydney, and especially the contribution to the relatively small literature on valuation of crowding. The primary

Table 15 Commuter frequency of service WTP values by mode

Frequency	Average	Median	Std Dev.	95% confidence interval Lower	Upper
Bus					
10	$0.01482	$0.01436	$0.00369	$0.00925	$0.02341
15	$0.02224	$0.02154	$0.00554	$0.01388	$0.03511
20	$0.02965	$0.02871	$0.00738	$0.01850	$0.04681
25	$0.03706	$0.03589	$0.00923	$0.02313	$0.05851
30	$0.04447	$0.04307	$0.01107	$0.02775	$0.07022
35	$0.05188	$0.05025	$0.01292	$0.03238	$0.08192
40	$0.05930	$0.05743	$0.01476	$0.03700	$0.09362
45	$0.06671	$0.06461	$0.01661	$0.04163	$0.10532
50	$0.07412	$0.07178	$0.01845	$0.04625	$0.11703
55	$0.08153	$0.07896	$0.02030	$0.05088	$0.12873
60	$0.08894	$0.08614	$0.02214	$0.05550	$0.14043
Train					
10	$0.00791	$0.00780	$0.00150	$0.00532	$0.01106
15	$0.01582	$0.01560	$0.00301	$0.01065	$0.02212
20	$0.02373	$0.02340	$0.00451	$0.01597	$0.03317
25	$0.03164	$0.03120	$0.00602	$0.02130	$0.04423
30	$0.03955	$0.03900	$0.00752	$0.02662	$0.05529
35	$0.04746	$0.04680	$0.00903	$0.03195	$0.06635
40	$0.05537	$0.05460	$0.01053	$0.03727	$0.07741
45	$0.06328	$0.06240	$0.01204	$0.04259	$0.08847
50	$0.07119	$0.07020	$0.01354	$0.04792	$0.09952
55	$0.07910	$0.07800	$0.01505	$0.05324	$0.11058
60	$0.08701	$0.08580	$0.01655	$0.05857	$0.12164
Metro					
2	$0.00316	$0.00312	$0.00060	$0.00213	$0.00442
4	$0.00633	$0.00624	$0.00120	$0.00426	$0.00885
6	$0.00949	$0.00936	$0.00181	$0.00639	$0.01327
8	$0.01266	$0.01248	$0.00241	$0.00852	$0.01769
10	$0.01582	$0.01560	$0.00301	$0.01065	$0.02212
12	$0.01898	$0.01872	$0.00361	$0.01278	$0.02654

focus is on establishing the role of a number of trip attributes in the definition of the modal preference expression for the proposed metro in the presence of existing modes of transport for access, main and egress stages of door-to-door travel activity.

The evidence on the relative marginal disutility (or utility) associated with the influencing attributes, essentially components of service level (time, frequency, transfers), costs and crowding as well as the metro-specific constant that captures unobserved influences associated with the metro option, enable us to derive estimates of willingness to pay for each component.

The application modelling system can import the various WTP estimates and use them to build up the empirical expressions of relative utility associated with the competing modes of transport. This can be achieved in a generalised time or cost specification.

A number of comments are appropriate as a way of reinforcing some important assumptions that underlie (or condition) the behavioural research herein:

1. The estimated model has been developed in a context where the metro is deemed to be a 'feasible' alternative. Essentially, the catchment origin-destination set has been premised on the metro being included in each respondent's choice set as if they were asked, in contrast to imposing it on them.
2. The application system within which the model can be applied must recognise that the parameter estimates and associated WTP estimates are applicable to a specific context in Sydney, and as such, their incorporation in a Sydney-wide modelling system must recognise this, and be prepared to assume that the parameters are transferable to the wider context.
3. On the reasonable assumption that the key outputs are portable to the application model system, the only constant that can be ported is the metro-specific constant. The numerical estimates for all segments are intuitively plausible relative to the car, train and bus and so we can be somewhat confident that we have identified the relative magnitude of the mean unobserved influences associated with the metro.
4. All modelling was undertaken as if the entire metro network will be in place. Consequently any applications that focus on a staged introduction must be mindful of the context in which all parameter estimates are derived.

Acknowledgements Special thanks to members of Parsons Brinkenhoff and Sydney Metro, in particular Rory Garland, William Davidson, Dan Jones, Gillian Akers and Tim Raimond for their participation and advice. David Ashley of SKM, is thanked for his careful review of our research. The support of Connie Jansen and Patrick Meza of Taverner Research assisted greatly in the collection of the data. Permission to publish our research was given by Transport NSW. Discussions on crowding and advice on the diagrams used with Gerard Whelan are appreciated. We also thank two referees for their comments.

References

Bhat CR (2001) Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transp Res, Part B, Methodol 35:677–693

Bhat CR (2003) Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transp Res, Part B, Methodol 37:837–855

Gilboa I, Schmeidler D, Wakker P (2002) Utility in case-based decision theory. J Econ Theory 105:483–502

Hensher DA (1994) Stated preference analysis of travel choices: the state of practice. Transportation 21(2):107–134

Hensher DA (2010) Attribute processing, heuristics and preference construction in choice analysis. In: Hess S, Daly A (eds) State-of art and state-of practice in choice modelling. Emerald, Bradford, pp 35–70

Hensher DA, Rose JM, Greene WH (2005) Applied choice analysis: a primer. Cambridge University Press, Cambridge

Hess S, Train KE, Polak JW (2005) On the use of a modified Latin hypercube sampling (MLHS) approach in the estimation of a mixed logit model for vehicle choice. Transp Res, Part B, Methodol 40(2):147–163

Kahnemann D, Tversky A (1979) Prospect theory: an analysis of decisions under risk. Econometrica 47(2):263–291

Krinsky I, Robb AL (1986) On approximating the statistical properties of elasticities. Rev Econ Stat 68:715–719

Krinsky I, Robb AL (1990) On approximating the statistical properties of elasticities: a correction. Rev Econ Stat 72(1):189–190

Louviere JJ, Hensher DA, Swait JF (2000) Stated choice methods and analysis. Cambridge University Press, Cambridge

Pepper J, Spitz G, Adler T (2003) Customer perspectives on multilevel coaches for increasing rail system capacity. Transp Res Rec 1838:19–29

Polydoropoulou A, Ben-Akiva ME (2001) Combined revealed and stated preference nested logit access and mode choice model for multiple mass transit technologies. Transp Res Rec 1771:38–45

Rose JM, Bliemer MC, Hensher DA, Collins AT (2008) Designing efficient stated choice experiments in the presence of reference alternatives. Transp Res, Part B, Methodol 42(4):395–406

Starmer C (2000) Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. J Econ Lit XXXVIII:332–382

Train K (2003) Discrete choice methods with simulation. Cambridge University Press, Cambridge

Whelan G, Crockett J (2009) A stated strength-of-preference investigation of the willingness to pay to reduce rail overcrowding. Paper presented at the 1st international conference on choice modelling conference, Harrogate, UK