A latent class approach to dealing with respondent uncertainty in a stated choice survey for fare simplification on bus journeys

Stephane Hess[1] - Jeremy Shires[2] - Peter Bonsall[3]

Institute for Transport Studies

University of Leeds

Lees LS2 9JT

Abstract

This paper presents the findings of an analysis making use of a stated choice survey looking at bus travellers' preferences for specific fare structures. Specifically, respondents were given the choice between the *current* fare structure, which is largely distance based, a *fixed* fare structure, and a *zonal* fare structure. It was anticipated that there might be two sources of substantial respondent uncertainty, in relation to the current fare as well as the number of zones a specific journey covers. These two types of uncertainty could create a significant level of bias in the results, and in this paper, we put forward a modelling approach based on a latent class structure that allows us to accommodate this respondent uncertainty and avoid the source of bias. Latent class structures are used widely to allow for variations in sensitivities across individual decision makers; here, we show how such a structure is equally well suited to dealing with respondent uncertainty expressed through variations in the *perceived* fares of different journey options.

Keywords: *respondent uncertainty; discrete choice; stated preference; fare simplification; latent class*

1. Introduction

Bus fares in many areas exhibit a high degree of heterogeneity, with fares varying for example by: (1) distance travelled or the number of fare stages passed; (2) by time of day of travel (typically peak and off peak); (3) by passenger characteristics (e.g. eligibility for fare concessions); (4) by journey frequency (inherent in the offer of reduced price season tickets, multi-journey tickets and return journey tickets); and (5) by operator (with different operators offering different fare structures).

Regular bus users become familiar with the fare(s) applicable to their regular journeys and the existence of a range of other fares need not concern them. However, notwithstanding the growing availability of pre-trip information (see e.g. Zito et al., 2010), for non-bus users and for occasional users, or for frequent bus users making unfamiliar journeys, the apparent complexity of the fare structure can be an obstacle and the time and effort required to ascertain the correct fare can be off-putting. Not knowing what the fare should be, the potential passenger may be concerned about it being much higher than they are prepared to pay, about the risk of not having the correct fare

[1] Corresponding author, s.hess@its.leeds.ac.uk, +44 (0)113 34 36611

[2] J.d.shires@its.leeds.ac.uk +44 (0)113 34 35347

[3] P.w.bonsall@its.leeds.ac.uk, +44 (0) 113 34 35335

available in cash or coins or about inadvertently paying more than they need to. This issue is related to recent discussions about passenger perceptions of level of service data (see Tam et al., 2010).

Fare structures can be simplified in various ways - though not without loss of ability to reflect local market conditions. Some of the main types of simplification are outlined below:

- At one level fare simplification might simply mean **rounding the fare** (e.g. to the nearest 50 pence) thus reducing the number of separate fare levels. This simplification is particularly helpful to passengers if the operator has a strict no-change policy.
- **Flat fare schemes** such as those in operation in Brighton and Edinburgh are often quoted as good examples of fare simplification. They offer travel anywhere within the city bus network for a flat fare. The problem is that, to maintain revenue, the introduction of a flat fare scheme involves pricing long journeys lower than the market would bear while pricing some short journeys off the system. This may or may not be acceptable to the operators and their sponsors.
- **Zonal fare schemes** are widely used and offer some of the advantage of flat fares without having to depart so far from the "natural" market fare. However, zonal fares can create boundary problems and rely on the definition of the zones (along with rules on what constitutes a boundary crossing) being clearly understood.
- The introduction of **transferable tickets** covering all services, by all operators which are used to complete a specified journey.
- The introduction of period (day, week, month) **travel cards** removes the need for travellers to know what the fare is for any particular journey – although this advantage is lost if the user finds that the card is not accepted by all operators or for all journeys.
- The use of **stored value cards** similarly removes the immediate need for travellers to know the price of individual journeys (in that they do not have to have the correct cash fare available) but does not help those who wish to know at the journey is going to cost before they decide to make it. London's Oyster card offers the traveller a smarter version of the stored value card by guaranteeing that they will be charged no more than the minimum amount necessary when travelling around the city.

There has been a concerted move in the last decade within the UK transport sector to move towards simplified fares. The Brighton and Hove Bus Company (BHBC) made a much publicised decision, in January 2001, to switch from a graduated fares scale to a widely advertised single flat fare of £1 (Langridge, 2001, Transit 2002). This produced rising revenues and an 8.5% increase in year on year trips. A number of bus companies have since opted to introduce flat fares schemes for particular routes or groups of services but the majority have opted to promote travel cards of varying lengths. Since 2001 BHBC has increased its standard flat fare, which currently stands at £2, and at the same time has also introduced a number of other flat fares for specific short hops journeys/zones reflecting the fact that fare structures need to change over time to reflect underlying changes in bus markets and company circumstances[4].

In London, the advantages of introducing simplified fares are strongly recognised by Transport for London (TfL, 2004 & Brown, 2004) who have continued the trend of fare simplification set by their predecessor, London Transport, and their introduction of the Travelcard and simpler single fares in the 1980s (Grayling & Glaister, 2000). Since then, the introduction of the Oyster card in 2003 has helped to make fares easier to pay and ensures that travellers pay the cheapest fare possible via its 'daily capping' function. There is danger of confusing the impact of 'payment simplification' with those of 'fare simplification'; however both impact strongly upon each other.

The trend of simplification has also been followed by the UK rail industry following long standing criticism about the complexity of fare structures, e.g. two thirds of individual customers said fare

[4] http://www.buses.co.uk/tickets/flatfare.aspx

complexity was a major problem as far back as 2003 (SRA, 2003). Recent moves by the industry and central government have resulted in three ticket types being introduced across the industry with the names chosen to reflect when the ticket can be used, i.e. anytime, off-peak and advanced[5].

In the present paper, we summarise experience from a study that looked at two potential departures from the current fare system which is largely distance based. The first is a *fixed* fare offering travel anywhere within the city bus network for the same specified fare. The second is a *zonal* fare scheme under which the fare payable depends on the number of zones used (a fare-per-zone is specified and multiplied by n, where n is the number of zones used). Preliminary data collection for this project (focus groups, depth interviews and analysis of attitudinal questions in a pilot questionnaire) indicated that many people had an inaccurate and sometimes very approximate idea of the current fare structure and of the likely fare payable for a specified journey and that there was some misunderstanding of the concept of zonal fares.

The question then arises of how to model this kind of uncertainty. Here, a key distinction needs to be made (cf. Bonsall, 2004) between the current case in which the uncertainty is the mind of the decision maker and the more common case when one is seeking to accommodate error and uncertainty in the observation of behaviour, of preferences or of the factors affecting them.

The modelling theory relevant to the treatment of uncertainty in the mind of the decision maker has been reviewed in papers by Bonsall et al. (2007, 2009). These papers draw attention to the considerable body of behavioural theory relevant to travellers' responses to complex or poorly understood price structures and identify the particular relevance of some key theories: The Theory of Rational choice (Von Neumann and Morgenstern, 1944); Expected Utility Theory and the Random Utility Model (Block and Marshack, 1960); Transaction Cost Analysis (Coase, 1937); the Theory of Bounded Rationality (Simon, 1957); Prospect Theory (Kahneman and Tversky, 1979); Ambiguity Avoidance (Ellsberg, 1961); and the Regret Principle (Loomes and Sugden, 1982). The literature includes numerous examples of models which seek to deal with uncertainty. They range from simple representations of individual decision making in which outcomes are based on the decision maker's rational assessment of the expected costs and benefits of the different options available to more psychologically-oriented models which make specific allowance for the extra costs which uncertainty places on the decision maker and for decision makers' asymmetric and non-linear responses to the probabilities of different outcomes.

The use of a modelling approach that does not make provision for misunderstandings and uncertainties in the mind of the traveller is likely to lead to biased results because it cannot capture the relevant asymmetries and non-linearities. The use of expected utility theory in such contexts (e.g. as recently discussed in Liu & Polak, 2007) goes some way towards addressing the issues but does not offer a wholly adequate means of dealing with the fact that decisions are affected by decision makers' attitudes to uncertainty and that these attitudes are generally unobservable and heterogeneously distributed.

The contribution made by this paper is to put forward an alternative framework for modelling heterogeneity arising from misunderstandings and uncertainties in mind of the traveller. The approach is based on a latent class model (see e.g. Greene & Hensher, 2003; Hess et al., 2009). Latent class structures are typically exploited for dealing with heterogeneity in respondent sensitivities, i.e. the marginal utilities that drive a decision maker's choices in different scenarios. In other words, this type of model is typically used to deal with uncertainty from the modeller's perspective, given the inability of the analyst to adequately characterise the taste heterogeneity in a deterministic manner. In the present paper, we exploit the model structure with a different aim,

[5] http://www.nationalrail.co.uk/times_fares/simple_fares.html

namely to account for uncertainty from the respondent's perspective. Our findings not only show significant improvements in model fit compared to models not accommodating this uncertainty but also more reasonable substantive results.

The remainder of this paper is organised as follows. Section 2 talks about survey design and data collection. This is followed in Section 3 by a discussion of modelling methodology, and in Section 4 the presentation of the empirical modelling results. Finally, Section 5 summarises the findings of the paper.

2. Survey design and data collection

Data were collected in October 2008 through computer assisted telephone interviews (CATI). For this, respondents were recruited via face to face interviews (mix of door to door and on street) during which an appointment was agreed for the main interview. Briefing sheets were sent to respondents in advance of the telephone interviews so that they could be referred to during the interview. Strict eligibility constraints were applied to ensure that interviewees were prima-facie likely to be influenced by simplification of fares. The requirement was that interviewees must be local resident but must not be pass-holders, eligible for free bus travel, under 18, or determined non-users of buses – i.e. people were excluded if they would not consider using buses even if they were cheaper, more comfortable and more frequent.

Respondents received a compensation of £10 for taking part in the survey. The sample for the main survey was specified as 300 individuals split equally between three areas, namely Warwick, Manchester and Leeds. These areas were chosen to represent three complex fare structures and to include a substantial sample from a non-metropolitan area to contrast with that from metropolitan areas. Table 1 shows the socio-demographic characteristics of the three areas. Clearly, Warwick is different from both Leeds and Manchester, in terms of higher gross disposable income, higher car ownership and a smaller population. Leeds would also appear to be different to Manchester but the picture is less clear since in effect Manchester here does not include Greater Manchester but inner Manchester which tends to be a more deprived area than the former. Additionally, both Leeds and Manchester have much larger populations than Warwick and are considerably less car dependent.

The questionnaire included several sections. The first sought background and attitudinal data including the respondent's estimate of the current fare between two specified points (and their confidence in that estimate), the second section included the Stated Choice (SC) exercise with which the current paper is concerned, the third section included stated intention questions and the final section included more background questions.

The specific reason for using a SC survey in the present study is to look at the preferences for different hypothetical fare structures under changing attribute values. The SC component of the survey presented each respondent with six binary choice situations. These choices were all in the context of a specific journey. Respondents were shown a map and were asked to "imagine that you have to make a single journey, by bus, from [X] to [Y] at about 11 o'clock on a cloudy but dry morning" (where [X] was a prominent location in the relevant city centre, and [Y] was a specified location around 2 miles away which had already been identified as one which the respondent knew but to which they had not travelled to by bus in the last year). As an example, Figure 1 shows the map used in one of our survey areas; respondents were told they were travelling from a specified point the city centre [in zone 1] to a specified point in zone 2.

The set of six choices was split into two choices involving the *as now* and *zonal* fare options, two choices involving the *as now* and *fixed* fare options, and two choices involving the *fixed* and *zonal* fare options. The order of presentation of the six choice sets, as well as the ordering of alternatives within each choice set, were randomised. The only exception to this was that the first choice set always involved a choice between the *as now* and *zonal* fare options. Finally, each respondent was also given a trial choice set to start with, involving the *as now* and *fixed* fare options.

The attributes used in the presentation of the alternatives included journey time as well as the fare structure and level. The reason for including journey time was to move away from a simple scenario in which respondents traded only on fare levels and fare structures. As a result, it should also be noted that travel time became a design attribute, hence varying across alternatives and choice tasks, just as was the case for fare. This may seem counterintuitive, as it allows for two journeys on the same route (but using different fare structures) to have different travel times. However, such variation is required in order to enable the formation of appropriate trade-offs. The approach was further justified by pilot work which showed no *realism* concerns from a respondent perspective; one possible perceived reason for time differences across different fare structures would for example be the use of different operators, with different equipment, or a difference in the number of stops.

Two important points need to be discussed. Firstly, for the *as now* options, the actual fare level was not shown, and secondly, for the *zonal* option, although the fare-per-zone was shown, respondents were expected to use this information in conjunction with the map to work out the correct fare for the specified journey. The omission of the actual fare level for the *as now* option was deliberate - we wanted to replicate the real life situation where some respondents do not know the existing fare (a large part of the appeal of a simplified fare structure might be that it is easier to estimate than an un-simplified fare and, had we revealed the true value of the as-now fare, this effect would have been lost). Similarly, the omission of the actual fare level for the *zonal* option was deliberate - we wanted to replicate the real life situation where some respondents might misunderstand a zonal fare and we did not want to hide this effect by revealing the actual fare payable.

Separate designs were generated for each of the three locations (Leeds, Manchester and Warwick) in which we intended to conduct our survey work. The design process started with the generation of three D-efficient designs (cf. Bliemer & Rose, 2009) in each of the three locations, where each design looked at one of the three possible binary combinations, and contained 18 rows. Special care was taken to avoid scenarios that involved combinations that gave one of the two alternatives an excessively large advantage (this being judged using the true values of the fares and a-priori values for the implied coefficients). Each respondent was then assigned two rows from each of the three designs, where the blocking was performed in such a way as to be almost orthogonal, i.e. minimising the correlations between blocks and the attributes of the alternatives.

Four levels were used for the time attribute, with the appropriate values being 22 minutes, 25 minutes, 27 minutes and 28 minutes for Leeds, 30 minutes, 35 minutes, 39 minutes and 40 minutes for Warwick, and 25 minutes, 30 minutes, 34 minutes and 35 minutes for Manchester. For fare, the costs for the as now option were £1.60 for Leeds, £3.60 for Warwick, and £2.60 for Manchester. For the *fixed* and *zonal* alternatives, three fare levels were used in each case, where the actual levels were the same for both alternatives (for the whole journey, not per zone), namely £1.60, £1.80, and £2.00 in Leeds, £3.60, £4.00, and £4.40 in Warwick, and £2.60, £3.00, and £3.40 in Manchester. In each case, fare simplification thus implies a slightly higher fare, allowing us to test the willingness of respondents to pay for fare simplification. An example of a choice scenario is shown in Figure 2.

3. Modelling methodology

In a discrete choice framework[6], the choice between a finite number of mutually exclusive alternatives is represented by making use of a mathematical construct of the attractiveness of each

[6] See Train (2003) for a thorough introduction.

alternative, its utility, and a decision rule which states that the alternative with the highest utility is chosen. Given that the utility of an alternative cannot be observed, we move to a random utility framework, in which the utility of an alternative is decomposed into a deterministic component and a random component, where the probability of choosing a given alternative increases with its deterministic utility, and where the assumptions made in relation to the random component determine the overall structure of the model.

Specifically, let the deterministic utility of alternative i (out of J) for respondent n (out of N) in choice situation t (out of T) be given by V_{int}, where this is a function of the attributes of that alternative, given by a vector x_{int}, and an associated vector of marginal utilities β, with $V_{int}=f(x_{int}, \beta)$. In the majority of applications, a linear in parameters specification is used for $f()$.

In the simplest structure, the Multinomial Logit model, the probability of choosing alternative i, conditional on β and the vector of attributes x_{nt} (grouping together attributes for all alternatives), is then given by:

$$P_{int}(\beta, x_{nt}) = \frac{e^{V_{int}}}{\sum_{j=1}^{J} e^{V_{jnt}}} \tag{1}$$

The contribution by a single decision maker to the likelihood function of a model is given by $P_n(\beta)$, which itself is equal to:

$$P_n(\beta, x_{nt}) = \prod_{t=1}^{T} \sum_{j=1}^{J} P_{jnt}(\beta, x_{nt})^{\delta_{jnt}}, \tag{2}$$

where δ_{jnt} is equal to *1* if alternative j is chosen by respondent n in choice set t, and *0* otherwise. In other words, the contribution by respondent n is given by the product across choice sets of the computed probability for the actual choices, conditional on β.

The actual log-likelihood function is given by:

$$LL(\beta, x\ \ ) = \sum_{n=1}^{N} ln\left(P_n(\beta, x_{nt})\right), \tag{3}$$

where x groups together the attributes for all observations and alternatives. The estimation of this model thus requires the maximisation of the log-likelihood function with a view to finding the values of β that best explain the observed choices.

In a random coefficients framework, we allow for uncertainty in the vector β. Specifically, in a latent class model, we assume that this heterogeneity can be expressed on the basis of a set of K mutually exclusive classes of respondents, where in each class k, a separate vector β_k is used. The probability of respondent n falling into class k is expressed as $\pi_{nk,}$ where $0 \leq \pi_{nk} \leq 1$, and where $\sum_{k=1}^{K} \pi_{nk} = 1$. Grouping $\pi_{nk,}\ \forall\ k,n$ into π, we then have:

$$LL(\beta, x\ \ , \pi) = \sum_{n=1}^{N} ln\left(\sum_{k=1}^{K} \pi_{nk} P_n(\beta_k, x_{nt})\right), \tag{4}$$

The probability of a respondent n falling into class k is itself determined through a class allocation model, possibly using socio-demographic indicators as explanatory attributes.

Latent class structures are traditionally used to allow for differences across respondents in their sensitivities to changes in the attributes of individual alternatives. Recent work, especially by David Hensher and colleagues in the context of research on information processing, has exploited the latent class structure with a view to allowing the actual utility specification to vary across respondents, with some respondents ignoring certain attributes, while other respondents may treat two related attributes in unison rather than separately (see e.g. Hess and Rose, 2007; Hensher and Greene, 2009; Layton and Hensher, 2010; Scarpa et al., 2009). Work by Hess et al. (2011) has shown how a framework based on a latent class approach can also be exploited to allow for variations across respondents in the actual choice paradigms.

The work presented in this paper falls into the general area outlined in the previous paragraph, but here, we exploit the latent class structure with a view to accommodating respondent uncertainty. Specifically, the difference between the typical latent class specification (e.g. Equation 4) and the one used in our model is that in our case, the differences across classes are not in the vector *β*, but in the vector *x*, i.e. the attribute levels as perceived by the respondent. Specifically, we use:

$$LL(\beta, x \quad, \pi) = \sum_{n=1}^{N} ln\left(\sum_{k=1}^{K} \pi_{nk} P_n(\beta \quad, x_{knt})\right), \tag{5}$$

where x_{knt} is now the vector of attributes for respondent *n* in choice situation *t*, if this respondent falls into class *k*. In the present paper, we exploit this approach to allow for uncertainty in the fare attribute, i.e. we have different classes in our model using different *perceived* values for the fare attribute. Given the small size of the sample available for this analysis, we do not additionally allow for variation in the vector *β*, where this is an important area for future research, for example in terms of different types of bus users having different preferences for specific fare structures.

4. Empirical analysis

This section presents the results of a detailed empirical analysis of the SC data. After initial data cleaning, we obtained a final sample of 95 respondents in Leeds, 98 respondents in Warwick, and 93 respondents in Manchester.

Looking at the market shares (calculated from the choice rates observed in the data) in the various subsamples (Table 2), we can observe a slight preference for *zonal* pricing, which is especially marked in Warwick, with relatively even shares for *as now* and *fixed* overall, but a slightly higher share for *as now* in Manchester, and a slightly higher share for *fixed* in Leeds. These differences between the three survey areas are consistent with what we know of the current fare structures and the specified zone sizes in the three areas. The reasons for the overall higher share for the *zonal* alternative will become clear in the remainder of this discussion.

4.1. Sources of uncertainty

Before looking in detail at model specification, we now turn our attention to the possible sources of respondent uncertainty. Two possible sources of respondent uncertainty arise in this dataset, relating to the fare of the *as now* and *zonal* options.

As noted earlier, respondents were not told the true value of the *as now* fare in the stated choice survey. In fact, only a minority of respondents had indicated that they felt certain about the fare level currently in use (12% across the entire sample, 21% in Leeds, 7% in Warwick and 9% in Manchester). Respondents who indicated that they were uncertain about the current fare were asked to indicate a lower and upper boundary on their estimate of the current fare. The uncertainty, also across respondents (noting that these fares are for a trip of the same length), is highlighted in Figure 3, which shows the distribution of respondents' estimates of the current fare on the basis of the lower estimate, middle estimate, and upper estimate.

Additionally, while it should have been clear from the map shown to respondents that the specified journey involved two zones, there was a suspicion that a large share of respondents had assumed that, because it crossed only one zonal boundary, it ought to be priced as a single zone trip. Such respondents would have underestimated the zonal fare by 50%. Since no direct questions were asked in relation to this attribute, the information processing strategy used in this context had to be inferred.

As a first step, we sought to overcome this by using an approach based on conditioning. Specifically, for each respondent in the data, and for each choice set, we calculated the choice probability using the estimated coefficients, but with two different fare levels for the *zonal* alternative, corresponding to a treatment of the trip as a one zone or two zone journey. This equates to using a multiplier γ on the *correct* fare that is equal to either 0.5 or 1. Our approach then worked out the likelihood for the actual observed sequence of choices for each respondent, making use of both π =0.5 and π =1. From this calculated likelihood, it is possible to infer the most likely value of γ for each respondent (cf. chapter 11, Train, 2003). This approach produced conclusive evidence to suggest the presence of a large share of respondents who underestimated the fare by 50% due to misunderstanding the zonal fare structure. Support for the use of 0.5 and 1.0 in this approach is provided by the fact that models built with other pairs of points (e.g. 0.25 and 1.25) provided significantly less explanation of the data (this is unsurprising; there is no intuitive reason for expecting values other than 0.5 and 1.0). This also gives further support to the assumption that the differences found across respondents are a result primarily of misunderstanding rather than standard heterogeneity in the cost sensitivity.

While the above discussion has suggested that there is a certain level of respondent uncertainty in the data, the question remains as to how to treat this during the analysis. For respondents indicating uncertainty in relation to the *as now* fare, the question remains as to whether they used their mean estimate, or their lower or upper estimate, or a function of all three values. Similarly, for the uncertainty in the *zonal* fare, the above analysis only tells us that a certain respondent may be more likely to evaluate the alternatives on the basis of a one-zone journey rather than a two-zone journey, but here the emphasis is on the word *more likely*, with the allocation to either group not being deterministic. This is the purpose of the next section.

4.2. Model specification

The analysis of the SC data made use of four different models. We first estimated a basic MNL model on the data. This was followed by three different specifications of the model discussed in Section 3.

In the first model, we made use of three different classes, representing the uncertainty in the *as now* fare. Specifically, in class 1, the fare for the *as now* alternative is given by a respondent's lower estimate for the current fare. In class 2, the middle value is used, while, in class 3, the upper estimate is used. This approach is motivated by the fact that, for those respondents who are uncertain about the current fare, it is not known to the analyst whether they use their best estimate or the lower or upper bound on that estimate. The fare actually used in evaluating the alternatives is thus unknown, and the probability for each of the three possible values is estimated.

The second model deals with uncertainty in the fare for the *zonal* alternative. Here, we make use of two classes, where in the first class, the correct fare is used, i.e. the fare per zone multiplied by two (the number of zones), while, in the second class, half the correct fare is used, i.e. the fare per zone multiplied by one. This approach is motivated by the fact that the analyst has no a priori knowledge as to a respondent's interpretation of the zonal fare system, and hence whether they treat the journey as a one zone or two zone journey.

In the third model, we jointly account for the two types of uncertainty. For this, we made use of six classes, namely three classes for respondents recognising the fact that it is a two zone journey, and three classes for respondents who treated the journey as a one zone journey. Each time, the three subclasses make use of the lower, middle, and upper estimates for the *as now* fare. We initially estimated the probability for the six classes separately, thus allowing for correlation between the two dimensions of uncertainty. Despite making use of appropriate normalisation, this however led to issues with convergence especially in some of the city-specific models, and we thus reverted to a specification assuming that the two types of uncertainty are independent. On tests using the full data, we observed almost no difference between the two specifications, suggesting that the two types of uncertainty are indeed independent.

In terms of the specification of the deterministic utilities, we made use of a purely linear specification, looking at the marginal sensitivity to changes in travel cost (β_{cost}) and travel time (β_{time}), where the overall preference for different structures was also estimated through constants for *fixed* (δ_{fixed}) and *zonal* (δ_{zonal}), i.e. applying the normalisation to the base fare structure. For travel time, the actual value presented in the survey was used, while the approach used for travel cost varied across models.

In the base model, the actual fare value was used for the *fixed* option, the respondent's own estimate of the current fare was used for the *as now* option[7], and the fare per zone was multiplied by two for the *zonal* option (the specified journey involved two zones).

On the basis of this, the utilities for the three alternatives are specified as:

$$V_{as\ now} = \beta_{\text{cost}} \cdot \text{fare}_{as\ now,\ estimated} + \beta_{\text{time}} \cdot \text{time}_{as\ now} \quad (6)$$

$$V_{fixed} = \delta_{\text{fixed}} + \beta_{\text{cost}} \cdot \text{fare}_{fixed} + \beta_{\text{time}} \cdot \text{time}_{fixed} \quad (7)$$

$$V_{zonal} = \delta_{\text{zonal}} + \beta_{\text{cost}} \cdot \text{fare}_{zonal} + \beta_{\text{time}} \cdot \text{time}_{zonal} \quad (8)$$

Attempts to allow for an interaction between cost sensitivity and income were not successful, possibly due to the aggregate nature of the income data. In the model allowing for uncertainty in the *as now* fare, the specification of V_{fixed} and V_{zonal} was as above. However, three different specifications for $V_{as\ now}$ were used across the three classes in this model, namely:

$$V_{as\ now,1} = \beta_{\text{cost}} \cdot \text{fare}_{as\ now,\ lower} + \beta_{\text{time}} \cdot \text{time}_{as\ now} \quad (9)$$

$$V_{as\ now,2} = \beta_{\text{cost}} \cdot \text{fare}_{as\ now,\ estimated} + \beta_{\text{time}} \cdot \text{time}_{as\ now} \quad (10)$$

$$V_{as\ now,3} = \beta_{\text{cost}} \cdot \text{fare}_{as\ now,\ upper} + \beta_{\text{time}} \cdot \text{time}_{as\ now} \quad (11)$$

In the model allowing for uncertainty in the zonal fare, the specification of V_{fixed} and $V_{as\ now}$ was as in the base model. However, two different specifications for V_{zonal} were used across the two classes in this model, namely:

$$V_{zonal,1} = \delta_{\text{zonal}} + \beta_{\text{cost}} \cdot \text{fare}_{zonal} + \beta_{\text{time}} \cdot \text{time}_{zonal} \quad (12)$$

$$V_{zonal,2} = \delta_{\text{zonal}} + \beta_{\text{cost}} \cdot \text{fare}_{zonal} \cdot \tfrac{1}{2} + \beta_{\text{time}} \cdot \text{time}_{zonal} \quad (13)$$

In the model jointly accounting for the two types of uncertainty, the six classes made use of the six possible combinations of $V_{as\ now}$ and V_{zonal} from the two basic latent class structures.

Table 3 provides an overview of the specification of the different models, in terms of which equations are used for the utility functions as well as the specification of the class allocation probabilities. In the estimation of the advanced discrete mixture models, the repeated choice nature of the data was taken into account by assuming class allocation varies across respondents, but stays constant for the six choices made by a given respondent (cf. Revelt & Train, 1998).

4.3. Estimation results

4.3.1. Base model

The estimation results for the base model are shown in Table 4. From the results, we can see that the estimates for the travel cost and travel time coefficients are negative and highly significant in all three subsamples as well as in the overall model. With the exception of δ_{fixed} in the Manchester model, the two constants are also significant above the usual 95% level of confidence. The positive value for the constants indicates that, all else being equal, respondents have a preference for *fixed* and *zonal* over the *as now* option. This is intuitively reasonable.

From the estimates for this model, we can also note that the valuation of travel time savings (VTTS), calculated on the basis of the time and cost sensitivities ($\beta_{\text{time}}/\beta_{\text{cost}}$), is higher than would normally be

[7] Attempts were also made to use the *actual* fare for the *as now* option, rather than a respondent's estimate, but this produced inferior results, which should not come as a surprise (many respondents had only an inaccurate idea of the actual fare).

expected (cf. DfT, 2009), while the willingness-to-pay (WTP) higher fares in return for a move to a zonal system (i.e. ratio between δ_{zonal} and β_{cost}) and the willingness to accept travel time increases in return for such a move[8] are also higher than would normally be expected when taking into account the market shares and the fact that overall fare levels were comparable across options (after taking account of the variation introduced by the experimental design). These two observations were consistent with the existence of uncertainty in the mind of the respondent. Indeed, if a high share of respondents consistently underestimate the zonal fare by 50%, this would lead to a higher than expected market share for this alternative, which a model not accommodating fare uncertainty would attempt to explain through a larger alternative specific constant, as we see here. It is also worth noting some significant differences in the WTP measures across the three study areas. While these are in part linked to differences in uncertainty, as will be discussed below, there are also differences in socio-demographic composition as outlined in Table 1.

4.3.2. Model allowing for respondent uncertainty in the as now fare

Our first departure from the base model allows for uncertainty in the *as now* fare, leading to the estimation of three additional parameters, relating to the probabilities for the three classes set out in Table 3. Looking at the results in Table 5, we observe improvements in model fit in all four models when compared to the model in Table 4. However, these improvements are only statistically significant at the 90% level for the joint model, the 85% level for the Leeds model, the 69% level for the Warwick model, and the 56% level for the Manchester model. The actual estimates show that none of the estimated probabilities is different from the 1/3 base level, although there is a suggestion that the upper bound receives no consideration in the Leeds data, while the middle estimates receive no or little consideration in the Warwick and Manchester datasets. There is a slight drop in the WTP measures compared to the base model, but they remain unrealistically high.

4.3.3. Model allowing for respondent uncertainty in the number of zones

Our next model allows for respondent uncertainty in the number of zones, thus making use of a two class structure. Looking at the results in Table 6, we observe highly significant improvements in model fit in all four models when compared to the model in Table 4, with gains in log-likelihood by 25.15, 9.44, 5.21, and 14.41 units for the overall model, the Leeds model, the Warwick model, and the Manchester model respectively. All increases come at the cost of two additional parameters, where the improvement in log-likelihood required at the 99% level of confidence is 4.605 units (leading to a χ^2_2 critical value of 9.21). The actual estimates show that none of the estimated probabilities is different from the ½ base level at the usual 95% level of confidence, though overall, there is slightly higher probability of a respondent treating the trip as a one-zone journey as opposed to a two-zone journey. The drop in WTP measures when compared to the base model is more marked in this model, giving us more realistic values, especially for the trade-offs involving δ_{zonal}, where we now see a much lower base preference for the *zonal* fare system, and in some cases, a preference for the *as now* option (in the form of a negative 'WTP' for the zonal system). In other words, a significant part of the overall preference for this alternative was a result of respondent uncertainty leading to an underestimation of the fare for this alternative.

4.3.4. Model allowing for both types of respondent uncertainty

We finally turn our attention to a model that incorporates both types of respondent uncertainty, with results reported in Table 7. This model is a generalisation of all three preceding structures, and offers statistically significant improvements in model fit not only over the base model, but also over the two models allowing only for a single type of respondent uncertainty. Indeed, against the base, we observe improvements by 33.94 units, 14.24 units, 8.66 units, and 18.93 units for the overall

---

[8] Here, the *as now* option is used as the base.

model, the Leeds model, the Warwick model, and the Manchester model respectively, all at the cost of five additional parameters, where the improvement in log-likelihood required at the 99% level of confidence is 7.54 units (leading to a χ^2_5 critical value of 15.08). In comparison with the model allowing for uncertainty in the as now fare, the improvements are 30.82 units, 11.56 units, 6.86 units, and 17.57 units, each at the cost of two additional parameters, where the improvement in log-likelihood required at the 99% level of confidence is 4.605 units (leading to a χ^2_2 critical value of 9.21). Finally, in comparison with the model allowing for uncertainty in the zonal fare, we observe improvements by 8.79 units, 4.8 units, 3.45 units, and 4.52 units, each at the cost of three additional parameters, meaning they are statistically significant at the 99% level for the overall data, the 98% level for the Leeds data, the 92% level for the Warwick data, and the 97% level for the Manchester data.

The improvement over the model allowing for uncertainty in the treatment of *zonal* fares is interesting as the initial improvements offered by the model allowing for uncertainty in the *as now* fares were not convincing. This suggests that the two types of uncertainty need to be dealt with jointly rather than separately. As was the case in the separate models, there is again little statistical evidence of asymmetry, with the exception of the treatment of *zonal* fares in the Leeds models, where we see a significantly larger probability for the class making use of a single zone, i.e. suggesting high rates of respondents who misunderstood the notion of zonal fares. We also observe a further drop in WTP measures, again leading to more realistic valuations.

4.3.5. Summary of results

A summary of the results from all four models is presented in Table 8. This highlights the improvement in model fit especially when moving to the model incorporating the *zonal* uncertainty and then the model incorporating both types of uncertainty. The results also once again show that the final model produces more realistic results in terms of the various WTP indicators.

Finally, it becomes clear on the basis of the two models incorporating uncertainty in the zonal fares that the high market share for the zonal alternatives was in large part due to underestimation by respondents of the fare for such journeys rather than an underlying preference for a zonal fare system. In the models not accounting for this uncertainty, we as a result retrieve a downwards bias in the fare coefficient and an upwards bias in the constant for the zonal alternatives. In the model accommodating the uncertainty, this bias is avoided.

5. Conclusions and further work

Our analysis has indicated that it is possible to cope with various kinds of respondent uncertainty in a stated choice context where it would be undesirable to attempt to remove this uncertainty by specifying unrealistically precise choice options. In our case we were faced with two types of respondent uncertainty which we have reason to believe exist in real life. The first was respondent uncertainty as to the current fare and the second was respondent uncertainty as to the correct interpretation of zonal fares.

The results from the base model are clearly biased as a result of not accommodating the respondent uncertainty. However, it is not possible to estimate the base model using the information actually used by respondents, as the uncertainty is not observed; i.e. we cannot with certainty say which respondents misunderstood the concept of a zonal fare system, or which respondents used what estimate for the current fare. A random treatment of uncertainty is thus required.

The methodological contribution in the paper comes in exploiting the latent class structure with a view to accommodating respondent uncertainty, where this model has previously been used primarily with a view to accommodating random variations in respondent sensitivities, and more recently to accommodate heterogeneity in information processing strategies. Our analysis has clearly shown that such treatment of respondent uncertainty is yet another promising area of application for the latent class model.

The first type of uncertainty was allowed for by using the respondent's own estimate of the current fare together with their upper and lower bounds on that estimate. Incorporation of the upper and lower bounds led to a small improvement in the model (compared to one based only on the central estimate) but it was interesting to note that different populations seemed to be putting different weights on the central, upper and lower estimates. The second type of uncertainty was allowed for by allowing, for each respondent, two possible interpretations of a zonal fare structure. This produced a very significant improvement over a model which had assumed that every respondent had interpreted the zonal fare correctly.

In the context of the present paper, we thus accounted for the two possible types of uncertainty, making use of the three possible levels that a respondent could have used for the as now fare (we disregard the possibility of intermediary values), and the two possible values for the zonal fare (disregarding nonsensical assumptions, e.g. of a three zone journey). The same approach as discussed here can clearly also be used in numerous other scenarios where respondent uncertainty possibly has a large impact on choice behaviour.

There is significant scope for future work, for example in terms of attempting to link respondent uncertainty to socio-demographic characteristics through an appropriate specification of the class allocation model, as well as looking at the additional incorporation of random variation in sensitivities across respondents. Attempts to look at such additional complexities however arguably require a larger dataset. There is also interest in exploring the relationship between choices and values deduced from the stated choice experiment and the expected responses to different fare structures provided by respondents in the third section of the questionnaire (Bonsall et al., 2009).

Acknowledgements

This paper is based on research that was undertaken for the UK Department for Transport. The opinions expressed in this paper are those of the authors and do not necessarily represent the views or policies of the UK Department for Transport. The first author acknowledges the financial support by the Leverhulme Trust, in the form of a *Leverhulme Early Career Fellowship*.

References

Bliemer, M. and Rose, JM. (2009), Designing Stated Choice Experiments: State of the Art, in R. Kitamura, T. Yoshii and T. Yamamoto (eds), The Expanding Sphere of Travel Behaviour Research: Selected Papers from the 11th International Conference on Travel Behaviour Research, Emerald Group Publishing Limited , Bingley, United Kingdom, pp. 499-538.

Block, H.D. and Marshack, J. (1960), Random orderings and stochastic theories of responses, in I. Olkin, S. Ghurye, W. Hoeffding, W. Madow and H. Mann (eds.), Contribution to probability and statistics, Stanford University Press, pp. 97-132.

Bonsall, P.W. (2004), Traveller Behaviour: Decision-making in an Unpredictable World, Journal of Intelligent Transport Systems, 8(1), pp. 45-60.

Bonsall, P.W., Shires, J.D., Matthews, B., Maule, J. and Beale, J. (2007), Responses to Complex Pricing Signals: Theory, Evidence and Implications for Road Pricing, Transportation Research Part A, 41(7), pp. 672-683.

Bonsall, P. W., Shires, J.D., Hess, S. and Koh, A. (2009), The value of simplicity: an investigation of travellers' response to simplified fare structures, paper presented at the 12th International Conference on Travel Behaviour Research (IATBR), Jaipur, India, December 2009.

Brown, I. (2004), Rail and a growing London, paper presented at the National Rail Conference Presentation, London, February 2004.

Coase, R. H. (1937), The Nature of the Firm, Economica, 4(16), pp. 386-405.

DfT (2009), Value of Time and Operating Costs, TAG 3.5.6, UK Department for Transport, Transport analysis guidance, http://www.dft.gov.uk/webtag/documents/expert/pdf/unit3.5.6.pdf

Ellsberg, D. (1961), Risk, Ambiguity and the Savage Axioms, Quarterly Journal of Economics, 75, pp. 643-69.

Grayling, T. and Glaister, S. (2000), A new fares contract for London, Institute for Public Policy Research. London.

Greene, W. H., Hensher, D. A. (2003), A latent class model for discrete choice analysis: contrasts with mixed logit, Transportation Research Part B, 37(8), pp. 681-698.

Hensher, D. A., Greene, W. H. (2009), Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification, Empirical Economics, 39(2), pp. 413-426.

Hess, S., Rose, J. M. (2007), A latent class approach to recognising respondents' information processing strategies in SP studies, paper presented at the Oslo Workshop on Valuation Methods in Transport Planning, Oslo, Norway, March 2007.

Hess, S., Ben-Akiva, M., Gopinath, D., Walker, J. (2009), Advantages of latent class models over continuous mixed logit, paper presented at the 12th International Conference on Travel Behaviour Research (IATBR), Jaipur, India, December 2009.

Hess, S., Stathopoulos, A., Daly, A.J. (2011), Allowing for heterogenous decision rules in discrete choice models: an approach and four case studies, ITS working paper, Institute for Transport Studies, University of Leeds.

Kahneman, D. and Tversky, A. (1979), Prospect Theory: An Analysis of Decision under Risk, Econometrica, 47(2), pp. 263-291.

Layton, D., Hensher, D. (2010), Aggregation of common-metric attributes in preference revelation in choice experiments and implications for willingness to pay, Transportation Research Part D, 15(7), pp. 394-404.

Liu, X. and Polak, J.W. (2007), Nonlinearity and the specification of attitudes towards risk in discrete choice models, Transportation Research Record 2014, pp. 27-31.

Loomes, G. and Sugden, R. (1982), Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty**,** The Economic Journal, 92, pp. 805-824.

ONS (2001), Neighbourhood Statistics, Office for National Statistics, http://www.neighbourhood.statistics.gov.uk/

Revelt, D. & Train, K. (1998), Mixed Logit with repeated choices: households' choices of appliance efficiency level, Review of Economics and Statistics, 80 (4), pp. 647-657.

Scarpa, R., Gilbride, T. J., Campbell, D., Hensher, D. A. (2009), Modelling attribute non-attendance in choice experiments for rural landscape valuation, Eur Rev Agric Econ 36 (2), 151-174.

Simon, H.A. (1957), Models of Man, John Wiley and sons, New York.

Strategic Rail Authority (2003), Fares Review Conclusions, London. www.dft.gov.uk/pgr/rail/faresreview

Tam, M.L., Lam, W.H.K., Lo, H.P. (2010), Incorporating passenger perceived service quality in airport ground access mode choice model, Transportmetrica, 6(1), pp. 3-17.

TfL (2004), London’s Railways – response to Government’s rail review, Transport for London, London.

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

Transit (2002), Brighton and Hove flat fare growth hits 9%, No. 174, p. 1, Landor, London.

von Neumann, J. and Morgenstern, O. (1944), Theory of Games and Economic Behavior, Princeton University Press.

Zito P., Amato G., Amoroso S. (2010), The effect of Advanced Traveller Information Systems on public transport demand and its uncertainty, Transportmetrica, 7(1), pp. 31-43.

Table 1: Socio-demographic characteristics of survey areas (source: ONS, 2001)

Category	Leeds	Manchester	Warwick	England
Population	715,402	392,819	125,931	49,138,831
Full Time Employed	40.44%	32.98%	44.25%	40.81%
Households without Cars	34.48%	47.79%	19.38%	26.84%
Gross Disposable Household Income Index	94.9	90	107.3	102.1

Table 2: Sample shares and choice rates

	Overall	**Leeds**	**Warwick**	**Manchester**
as now vs. *fixed*	47.2%/ 52.80%	40.00%/ 60.00%	46.94%/ 53.06%	54.84%/ 45.16%
as now vs. *zonal*	40.56%/ 59.44%	37.89%/ 62.11%	37.24%/ 62.76%	46.77%/ 53.23%
fixed vs. *zonal*	36.36%/ 63.64%	40.53%/ 59.47%	33.16%/ 66.84%	35.48%/ 64.52%
Overall share: *as now*	29.25%	25.96%	28.06%	33.87%
Overall share: *fixed*	29.72%	33.51%	28.74%	26.88%
Overall share: *zonal*	41.03%	40.53%	43.20%	39.25%

Table 3: Class structure, and equations used for utility specification within classes

	Base model				Uncertainty in as now				Uncertainty in zonal				Both types of uncertainty			
Class	$V_{as\ now}$	V_{fixed}	V_{zonal}	π	$V_{as\ now}$	V_{fixed}	V_{zonal}	π	$V_{as\ now}$	V_{fixed}	V_{zonal}	π	$V_{as\ now}$	V_{fixed}	V_{zonal}	π
1	(6)	(7)	(8)	1	(9)	(7)	(8)	$\pi_{lower\ estimate}$	(6)	(7)	(12)	$\pi_{2\ zones}$	(9)	(7)	(12)	$\pi_{2\ zones} \cdot \pi_{lower\ estimate}$
2					(10)	(7)	(8)	$\pi_{middle\ estimate}$	(6)	(7)	(13)	$\pi_{1\ zone}$	(9)	(7)	(13)	$\pi_{1\ zone} \cdot \pi_{lower\ estimate}$
3					(11)	(7)	(8)	$\pi_{upper\ estimate}$					(10)	(7)	(12)	$\pi_{2\ zones} \cdot \pi_{middle\ estimate}$
4													(10)	(7)	(13)	$\pi_{1\ zone} \cdot \pi_{middle\ estimate}$
5													(11)	(7)	(12)	$\pi_{2\ zones} \cdot \pi_{upper\ estimate}$
6													(11)	(7)	(13)	$\pi_{1\ zone} \cdot \pi_{upper\ estimate}$

Table 4: Estimation results for base models

	Full data		Leeds		Warwick		Manchester	
Observations:	1,716		570		588		558	
Respondents:	286		95		98		93	
LL(0):	-1,189.44		-395.09		-407.57		-386.78	
LL(β):	-1,092.24		-361.93		-367.51		-350.25	
adj. $\rho^2(0)$:	0.078		0.074		0.088		0.084	
Parameter	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
β_{cost}	-0.5310	-8.13	-0.9390	-4.19	-0.3820	-4.73	-0.8310	-5.83
β_{time}	-0.0753	-8.34	-0.1130	-4.74	-0.0672	-5.03	-0.0774	-5.32
δ_{fixed}	0.2840	3.72	0.3710	2.96	0.3340	2.43	0.2450	1.65
δ_{zonal}	0.7080	8.98	0.6080	4.65	0.8880	6.25	0.7230	4.67
	est.		est.		est.		est.	
VTTS (pence/min)	14.18		12.03		17.59		9.31	
VTTS (£/hr)	8.51		7.22		10.55		5.59	
WTP fixed (p)	53.48		39.51		87.43		29.48	
WTP zonal (p)	133.33		64.75		232.46		87.00	
Fixed vs time (min)	3.77		3.28		4.97		3.17	
Zonal vs time (min)	9.40		5.38		13.21		9.34	

Table 5: Model allowing for respondent uncertainty in the *as now* fare

	Full data		Leeds		Warwick		Manchester	
Observations:	1,716		570		588		558	
Respondents:	286		95		98		93	
LL(0):	-1,189.44		-395.09		-407.57		-386.78	
LL(β):	-1,089.12		-359.25		-365.71		-348.89	
adj. $\rho^2(0)$:	0.078		0.073		0.086		0.08	
Parameter	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
β_{cost}	-0.5840	-8.18	-1.0800	-4.71	-0.4360	-4.74	-0.9010	-5.58
β_{time}	-0.0773	-8.43	-0.1200	-5.02	-0.0686	-4.94	-0.0781	-5.30
δ_{fixed}	0.2790	2.86	0.4440	2.99	0.3870	1.92	0.2450	1.08
δ_{zonal}	0.7070	7.17	0.6790	4.57	0.9500	4.63	0.7350	3.17
$\pi_{lower\ estimate}$ [a]	0.2650	-0.48	0.3120	-0.10	0.4940	0.35	0.3970	0.19
$\pi_{middle\ estimate}$ [a]	0.3950	0.28	0.6880	1.65	0.0000	0.00	0.1290	-0.38
$\pi_{upper\ estimate}$ [a]	0.3400	0.04	0.0000	0.00	0.5060	0.36	0.4750	0.46
	est.		est.		est.		est.	
VTTS (pence/min)	13.24		11.11		15.73		8.67	
VTTS (£/hr)	7.94		6.67		9.44		5.20	
WTP fixed (p)	47.77		41.11		88.76		27.19	
WTP zonal (p)	121.06		62.87		217.89		81.58	
Fixed vs time (min)	3.61		3.70		5.64		3.14	
Zonal vs time (min)	9.15		5.66		13.85		9.41	

[a] t- ratios for $\pi_{lower\ estimate}$, $\pi_{middle\ estimate}$, and $\pi_{upper\ estimate}$ calculated in relation to 1/3.

Table 6: Model allowing for respondent uncertainty in the number of zones

	Full data		Leeds		Warwick		Manchester	
Observations:	1,716		570		588		558	
Respondents:	286		95		98		93	
LL(0):	-1,189.44		-395.09		-407.57		-386.78	
LL(β):	-1,067.09		-352.49		-362.30		-335.84	
adj. $\rho^2(0)$:	0.098		0.093		0.096		0.116	
Parameter	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
β_{cost}	-0.8710	-10.87	-1.7500	-6.26	-0.5770	-5.58	-1.2600	-8.23
β_{time}	-0.0876	-8.74	-0.1330	-5.03	-0.0728	-5.00	-0.1010	-5.97
δ_{fixed}	0.4390	5.25	0.4430	3.26	0.4810	3.17	0.5000	3.07
δ_{zonal}	0.1990	1.75	-0.3310	-1.26	0.4510	2.22	-0.0233	-0.09
$\pi_{2\ zones}$ [a]	0.4410	-0.84	0.3540	-1.48	0.4540	-0.33	0.4370	-0.52
$\pi_{1\ zone}$ [a]	0.5590	0.84	0.6460	1.47	0.5460	0.33	0.5630	0.52
	est.		est.		est.		est.	
VTTS (pence/min)	10.06		7.60		12.62		8.02	
VTTS (£/hr)	6.03		4.56		7.57		4.81	
WTP fixed (p)	50.40		25.31		83.36		39.68	
WTP zonal (p)	22.85		-18.91		78.16		-1.85	
Fixed vs time (min)	5.01		3.33		6.61		4.95	
Zonal vs time (min)	2.27		-2.49		6.20		-0.23	

[a] t- ratios for $\pi_{2\ zones}$ and $\pi_{1\ zone}$ calculated in relation to 1/2.

Table 7: Model allowing for both types of respondent uncertainty

	Full data		Leeds		Warwick		Manchester	
Observations:	1,716		570		588		558	
Respondents:	286		95		98		93	
LL(0):	-1,189.44		-395.09		-407.57		-386.78	
LL(β):	-1,058.30		-347.69		-358.85		-331.32	
adj. $\rho^2(0)$:	0.103		0.097		0.097		0.12	
Parameter	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
β_{cost}	-1.05	-10.59	-2.14	-3.03	-0.706	-5.65	-1.51	-7.74
β_{time}	-0.0941	-8.75	-0.152	-2	-0.076	-4.94	-0.11	-5.82
δ_{fixed}	0.479	4.25	0.616	1.01	0.611	2.68	0.564	2.27
δ_{zonal}	0.121	0.86	-0.465	-1	0.464	1.78	-0.239	-0.67
$\pi_{2\ zones}$ [a]	0.44	-0.92	0.318	-2.01	0.451	-0.4	0.396	-0.96
$\pi_{1\ zone}$ [a]	0.56	0.92	0.682	2.01	0.549	0.4	0.604	0.96
$\pi_{lower\ estimate}$ [b]	0.334	0.01	0.382	0.26	0.494	0.55	0.412	0.43
$\pi_{middle\ estimate}$ [b]	0.268	-0.45	0.513	0.66	0	-0.68	0.0852	-0.95
$\pi_{upper\ estimate}$ [b]	0.397	0.61	0.105	-1.4	0.506	0.6	0.503	0.92
	est.		est.		est.		est.	
VTTS (pence/min)	8.96		7.10		10.76		7.28	
VTTS (£/hr)	5.38		4.26		6.46		4.37	
WTP fixed (p)	45.62		28.79		86.54		37.35	
WTP zonal (p)	11.52		-21.73		65.72		-15.83	
Fixed vs time (min)	5.09		4.05		8.04		5.13	
Zonal vs time (min)	1.29		-3.06		6.11		-2.17	

[a] t- ratios for $\pi_{2\ zones}$ and $\pi_{1\ zone}$ calculated in relation to 1/2.

[b] t- ratios for $\pi_{lower\ estimate}$, $\pi_{middle\ estimate}$, and $\pi_{upper\ estimate}$ calculated in relation to 1/3.

Table 8: Comparison of results across models

adj. $\rho^2(0)$

	Full data	Leeds	Warwick	Manchester
Base model	0.078	0.074	0.088	0.084
Respondent uncertainty in the as now fare	0.078	0.073	0.086	0.08
Respondent uncertainty in the number of zones	0.098	0.093	0.096	0.116
Both types of respondent uncertainty	0.103	0.097	0.097	0.12

VTTS (pence/min)

	Full data	Leeds	Warwick	Manchester
Base model	14.18	12.03	17.59	9.31
Respondent uncertainty in the as now fare	13.24	11.11	15.73	8.67
Respondent uncertainty in the number of zones	10.06	7.6	12.62	8.02
Both types of respondent uncertainty	8.96	7.1	10.76	7.28

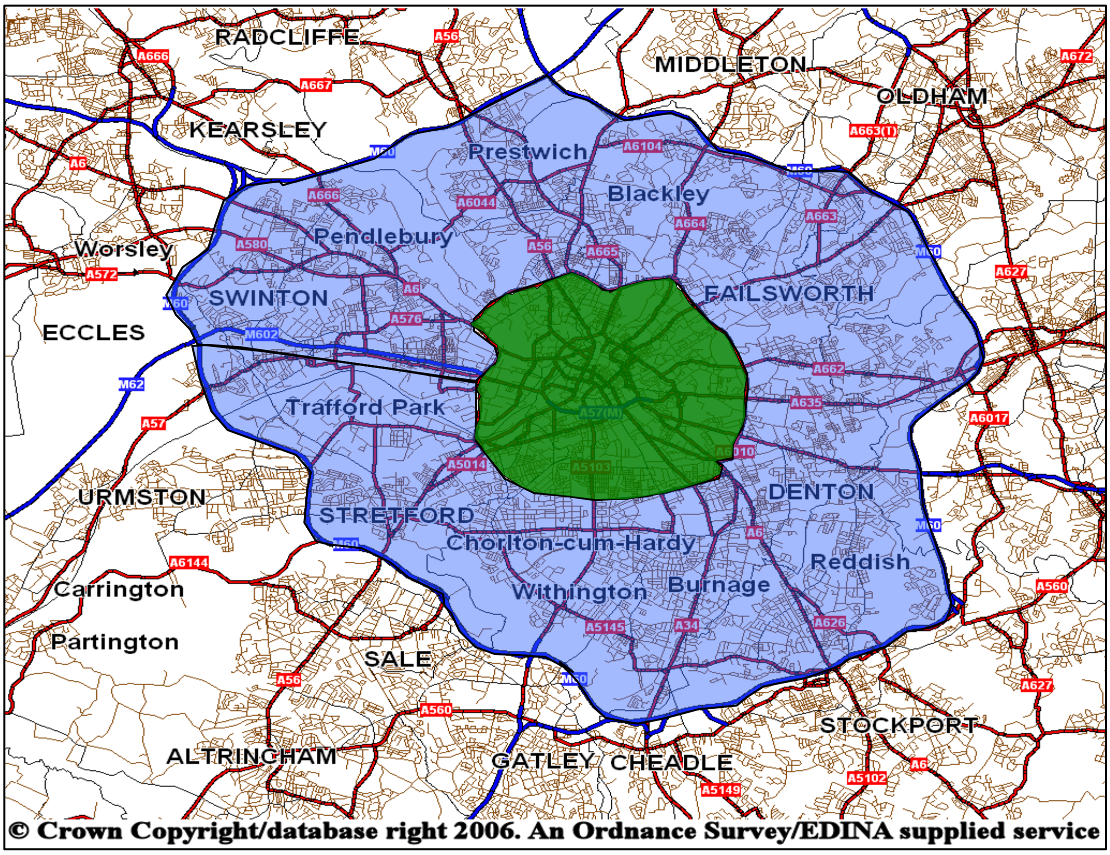
WTP fixed (p)

	Full data	Leeds	Warwick	Manchester
Base model	53.48	39.51	87.43	29.48
Respondent uncertainty in the as now fare	47.77	41.11	88.76	27.19
Respondent uncertainty in the number of zones	50.4	25.31	83.36	39.68
Both types of respondent uncertainty	45.62	28.79	86.54	37.35

WTP zonal (p)

	Full data	Leeds	Warwick	Manchester
Base model	133.33	64.75	232.46	87
Respondent uncertainty in the as now fare	121.06	62.87	217.89	81.58
Respondent uncertainty in the number of zones	22.85	-18.91	78.16	-1.85
Both types of respondent uncertainty	11.52	-21.73	65.72	-15.83

Figure 1: Map used in conjunction with SC exercise



Zone 1 =		Zone 2 =		Zone 3 =	

Figure 2: Example of choice offered in SC survey

	A	B
Fare structure:	**Fixed**	**As now**
Fare level:	**£1.60**	**As now**
Journey time (average at this time of day):	**30** minutes	**20** minutes

Figure 3: Respondent uncertainty in relation to *as now* fare