Treatment of reference alternatives in stated choice surveys for air travel choice behaviour

Stephane Hess

Institute for Transport Studies
University of Leeds
s.hess@its.leeds.ac.uk

Abstract

With the need for accurate forecasts of passenger demand, the airline sector is increasingly making use of behavioural models calibrated on data from stated choice surveys that allow for the analysis of hypothetical travel situations. To allow analysts to better frame the scenarios presented to respondents, the choice situations in such stated choice surveys often include a current trip as one of the travel options. Classically, these reference alternatives have been treated in the same way as the hypothetical alternatives. The applications presented in this paper show that this potentially leads to biased results, and that it is important to recognise the differences in the nature of the two types of alternatives.

Keywords: discrete choice, logit, air travel behaviour, stated choice, stated preference

1 Introduction

Accurate forecasts of air travel demand are a crucial requirement not only for airlines and airports, but also for transport authorities in many major metropolitan areas. Aside from long term trends in demand, there is great interest in how this demand could be affected by more short term changes in service levels or service characteristics. Airlines are for example likely to be interested in the potential impact of increases or reductions in air fares on passenger numbers, while airports may want to gauge the impact of reductions in minimum check-in or transfer times on the attractiveness (and hence usage) of an airport. Finally, urban transport planners require accurate forecasts of passenger levels and modal split for the access journeys to airports.

In the context there is growing interest in making use of state of the art modelling techniques to analyse air travel behaviour, with a particular reliance on discrete choice models (DCM) belonging to the family of random utility models (RUM)[1]. RUM structures can be calibrated on two main types of data, revealed preference (RP) data describing actual real world choices, and stated choice (SC) data containing choices from hypothetical scenarios presented in travel surveys. With there being a strong interest in predicting behaviour across a range of hypothetical settings (e.g. fare reductions, new routes, new access modes) there is an increasing reliance on SC data[2].

While SC data have an advantage over RP data in being able to look at choices in hypothetical settings, there has been considerable concern about response quality (Louviere et al., 2000), leading to attempts to increase the realism of SC choice situations. One possibility is to weaken the hypothetical nature of surveys by framing choice situations around a scenario known to the respondent. In a growing number of cases, this is achieved by including the current choice as one of the alternatives in the survey. Evidence suggests that such a framing approach makes preference

[1] For a discussion of discrete choice models, see Train (2003).
[2] Recent examples including Proussaloglou and Koppelman (1999), Adler et al. (2005), and Hess et al. (2007)

revelation more meaningful at the level of the individual (Starmer, 2000) and has the advantage of allowing the analyst to determine what kind of incentives are required to get a respondent to move away from their current travel option. One example of exploiting this type of framing approach is the air travel survey data collected by Resource Systems Group Inc (2003) in the US.

However, while this type of survey design advantages in terms of framing the choice situation, potentially crucial in the context of complex air travel decisions, it is not immediately clear whether standard modelling approaches are appropriate for use on such data. Indeed, the two types of alternatives included in the choice situations in these surveys are inherently different (hypothetical versus actually experienced) and it could be suggested that these differences need to be accommodated in the modelling framework.[3]

2 Data

The analysis makes use of SC data collected via the Internet by Resource Systems Group. Specifically, we make use of the 2005 version of the survey, with a sample of 4,256 observations collected from 532 randomly selected travellers who had recently undertaken a domestic air trip. Prior to the SC survey, information was collected on a traveller's most recent air trip, along with detailed socio-demographic information. The traveller is then faced with 8 binomial choices, where in each case, a choice is offered between the current flight and an alternative. While the attributes of the reference alternative remain fixed across the eight choice sets, those of the second alternative are varied according to an experimental design. The airports and airlines used for this second alternative are selected on the basis of information gathered from respondents in terms of a ranking of the airports and airlines available to them.

Aside from the airport and airline names, the attributes used to describe the alternatives in the SC survey include flight time, number of connections, air-fare, arrival time (used to calculate schedule delays), aircraft type, and on-time performance of the various services. Access cost is not included (in the absence of an actual specification of the mode choice dimension), and no choice is given between travel classes.

3 Methodology

Three main modelling approaches are used, with different degrees of recognising the specific nature of the two alternatives included in the choice sets. All models have a multinomial logit (MNL) structure at the core but normally distributed error components with a zero mean were included to account for each respondent facing with eight choices. These random terms are distributed identically and independently across respondents and alternatives, but not across observations for the same respondent. This allows for an individual-specific effect that can be interpreted as a random scale with the aim of avoiding bias in the standard errors, where such bias commonly exhibits itself in the form of underestimated standard errors when failing to recognise the repeated choice nature of SC data (Ortúzar et al., 2000).

[3] The issues dealt with here are different from similar discussions of studies combining separate RP and SC datasets in a joint analysis (Ben-Akiva and Morikawa, 1990). The difference lies that, with the present data, a real world alternative is included as one of the options in the SC survey, but only a single dataset is used.

3.1 Base specification

A standard specification is used for the base model, with all parameters entering the utility function in linear fashion. A common coefficient is used for all levels of memberships in frequent flier programmes, and no distinction is made between flights with a single connection and flights with two connections[4]. The observed utility (V) for the reference alternative (R) is given by:

$$V_R = \beta_{current} + \beta_{access\ time} \cdot \text{access time}_R + \beta_{air\ fare} \cdot \text{air fare}_R + \beta_{flight\ time} \cdot \text{flight time}_R + \beta_{OTP} \cdot OTP_R$$

$$+\beta_{connecting}\ \delta_{connecting,R} + \beta_{FF}\ \delta_{FF,R} + \beta_{closest\ airport}\ \delta_{closest\ airport,R}, \qquad (1)$$

where all β parameters are to be estimated.

The parameter, $\beta_{current}$, is an alternative specific constant (ASC) for the reference alternative that, amongst other things, captures inertia. Parameters, $\beta_{access\ time}$, $\beta_{air\ fare}$ and $\beta_{flight\ time}$ are marginal utility coefficients that capture the disutility associated with an increase by one unit (1 minute or \$1) in access time, air fare and flight time. β_{OTP}, relates to the on-time performance (in percentage points) of an alternative. For the reference alternative, two levels are used, depending on whether the flight was on time (100%) or not (zero), while, for the second alternative, five levels between 50% and 90% are used. The variable $\delta_{connecting,R}$ is set to unity for flights with at least one connection, while $\delta_{FF,R}$ is set to one if the respondent holds some form of frequent flier (FF) membership with the airline. Finally, $\delta_{closest\ airport,R}$ is set to unity if the airport used for the trip is that closest to the respondent's home.

The utility function for the second alternative is specified in a similar fashion, with the absence of ASC ($\beta_{current}$), and with the hypothetical, as opposed to reference, values for the various attributes and dummy variables.

3.2 Differential response to attribute values of reference alternative

To test the validity of the assumption that respondents treat the attributes of the reference alternative in the same way as those of the hypothetical alternatives, we use a specification in which all coefficients are alternative-specific:

$$V_R = \beta_{current} + \beta_{access\ time,R} \cdot \text{access time}_R + \beta_{air\ fare,R} \cdot \text{air fare}_R + \beta_{flight\ time,R} \cdot \text{flight time}_R + \beta_{OTP,R} \cdot OTP_R + \beta_{connecting,R}\ \delta_{connecting,R} + \beta_{FF,R}\ \delta_{FF,R} + \beta_{closest\ airport,R}\ \delta_{closest\ airport,R}, \qquad (2)$$

The corresponding specification for the second alternative again lacks a constant, with the remaining seven coefficients being specific to the alternative. This specification not only allows for differences in how respondents react to the attribute values of the two alternatives, but also accounts for differences in the on-time performance attributes for the alternatives (a simple distinction between on-time and delayed flights for the reference alternatives, with percentage rates of on-time arrival for the second alternative).

3.3 Asymmetrical preference formation

Finally, we develop a model based on concepts taken from prospect theory, where the attribute levels of an alternative are evaluated relative to those of the base alternative (Hess et al., 2008),

[4] Very few alternatives with two connections were included.

while allowing for a differential response to increases and decreases (gains and losses) compared to these base levels:

$$V_R = \beta_{current}, \tag{3}$$

and

$$V_S = \beta^{+}_{access\ time} \cdot \delta_{access\ time\ inc} \cdot (access\ time_S - access\ time_R) + \beta^{-}_{access\ time} \cdot \delta_{access\ time\ dec} \cdot (access\ time_R - access\ time_S) + \ \ldots \ , \tag{4}$$

where we only show the coefficients associated with access time.

With this specification, the coefficients in the utility function for the second alternative interact with the difference between the attribute values for the two alternatives. Separate coefficients are used for increases and decreases relative to the attribute value for the base alternative, with $\beta^{+}_{access\ time}$ and $\beta^{-}_{access\ time}$, for example, giving the coefficients for increases and decreases in the access time attribute. The additional term $\delta_{access\ time\ inc}$ is set to one only when the access time is longer for the second alternative than for the base alternative, with the same applying for $\delta_{access\ time\ dec}$ in the case of decreases relative to the base alternative. The assumption of a symmetrical response can be tested by looking at the difference between coefficients for increases and decreases, say the difference between $\beta^{+}_{access\ time}$ and $\beta^{-}_{access\ time}$ in the case of access time.

There is another difference compared to the previous models. A common factor is used for all non-zero levels of connections and all levels of frequent flier membership in Equations 1 and 2. This would mean, for example, that the willingness-to-pay more for flying on an airline where the passenger holds an elite frequent flier account is the same as for flying on an airline where the passenger only holds a standard account. In the asymmetrical model, such counter-intuitive valuations are avoided by taking the difference between the number of connections, and between the tiers in the frequent flier programmes[5]. This multiplicative approach makes the assumption of linearity in the sensitivities[6], and no evidence is found to suggest that this is not justified.

4 Results

Three models are estimated in BIOGEME (Bierlaire, 2005) and the simulation-based estimation is carried out using 500 Halton draws (Halton, 1960). No distinction is made between separate purpose segments, and no socio-demographic interactions were tested.

4.1 Base model

The results for the base model are shown in Table 1. All marginal utility coefficients have the expected sign and are significantly different from zero with high levels of confidence. The results indicate that increases in air-fare, flight time and access time have a negative effect on utility, while increases in on-time performance have a positive effect. Respondents also have a preference for direct flights, flights on an airline where they receive frequent flier benefits, and flights from the

[5] Four tiers are used; no membership, standard membership, elite membership and elite plus membership.

[6] That is, an increase from one connection to two connections carries the same penalty as a change from a direct flight to a flight with a single connection.

airport closest to their ground-level origin. The positive estimate for the constant associated with the current alternative indicates that, all else being equal, respondents have a strong preference for their current option, implying a high level of inertia. The standard deviation of the error components is significantly different from zero, suggesting the presence of an individual specific effect. In terms of monetary valuations, the implied willingness-to-pay for access time reductions is almost 40% larger than the corresponding figure for flight time reductions. The results also show that frequent flier benefits are valued almost as highly as a reduction in access time by one hour, with the valuation of direct flights being even higher.

Table 1: Results for base model

Respondents	532	
Observations	4256	
LL	-1522.76	
par	9	
adj. ρ^2	0.4808	
	est.	asy. t-rat.
$\beta_{current}$	1.0902	5.67
$\beta_{access\ time}$	-0.0069	-6.14
$\beta_{air\ fare}$	-0.0165	-8.66
$\beta_{flight\ time}$	-0.0050	-6.08
β_{OTP}	0.0109	3.95
β_{FF}	0.3520	2.62
$\beta_{closest\ airport}$	0.5705	5.05
$\beta_{connecting}$	-0.7507	-5.18
σ	1.0932	7.16
WTP indicators		
access time reductions ($/hour)	25.06	
flight time reductions ($/hour)	18.13	
on time arrival ($)	66.06	
FF benefits ($)	21.38	
departure from closest airport ($)	34.65	
direct flight ($)	45.60	

4.2 Model with differential response to reference alternative attribute values

The results for the model allowing for a differential response to the attribute levels of the two alternatives are summarised in Table 2. The base model is a simplified version of this model, so a likelihood-ratio test can be used for comparison. The log-likelihood (LL) improves by 12.52 units, at the cost of seven additional parameters, giving a test value of 25:05, with a χ^2_7 critical value of 14.07, meaning the improvement is statistically significant.

Table 2: Results allowing for differential responses to reference alternative attribute values

Respondents	532
Observations	4256
LL	-1510.24
par	16

adj. ρ^2	0.4826

	RP alternative		SP alternative		
	est.	asy. t-rat.	est.	asy. t-rat.	asy. t-rat (diff)
$\beta_{current}$	1.2018	2.23	-	-	-
$\beta_{access\ time}$	-0.0107	-4.08	-0.0064	-5.49	1.72
$\beta_{air\ fare}$	-0.0156	-8.29	-0.0170	-8.89	1.59
$\beta_{flight\ time}$	-0.0044	-4.67	-0.0056	-5.97	1.40
β_{OTP}	0.0084	2.73	0.0136	3.57	1.10
β_{FF}	0.4468	2.50	0.3197	2.25	0.72
$\beta_{closest\ airport}$	0.3321	1.85	0.6663	4.97	1.57
$\beta_{connecting}$	-0.9144	-4.36	-0.6778	-3.85	0.95
σ	1.0435	7.70	1.0435	7.70	-

WTP indicators	RP alternative	SP alternative
access time reductions ($/hour)	41.19	22.40
flight time reductions ($/hour)	16.83	19.81
on time arrival ($)	53.80	79.86
FF benefits ($)	28.62	18.76
departure from closest airport ($)	21.27†	39.10
direct flight ($)	58.57	39.78

† calculation involves parameter significant only at the 93% level of confidence

Additionally, all of the attributes, the difference in the sensitivities for the two alternatives is not significant at the 95% level. However, levels of 91%, 89% and 88% are obtained in the case of $\beta_{access\ time}$, $\beta_{air\ fare}$ and $\beta_{closest\ airport}$, with 84% in the case of flight time. For the remaining three coefficients, the significance levels for differences are lower, at 73% for β_{OTP}, 53% for β_{FF} and 66% for $\beta_{connecting}$.

In terms of differences between the two alternatives, sensitivity to access time changes is 68% higher for the reference alternative. With the air-fare coefficient being 10% higher for the second alternative, there is a much higher monetary valuation of travel time savings on the access journey for the reference alternative. On the other hand, the degree by which flight time increases are valued more negatively for the second alternative overturns the higher air fare sensitivity, leading to a higher monetary valuation of flight time reductions for the hypothetical alternative. Major differences also arise for β_{OTP}, β_{FF} and $\beta_{connecting}$, but the significance levels for these differences are too low to make any inferences. Finally, although the difference is only significant at the 88% level, the sensitivity towards increases in the on-time performance is much higher for the second alternative than for the reference alternative. Here, the range for the levels for the attribute in the two alternatives at least partly explains these differences.

4.3 Model with asymmetrical response formation

The results for the model allowing for asymmetrical preference formation are summarised in Table 3. Separate coefficients were estimated for increases and decreases relative to the base alternative for the seven explanatory attributes. In each case, asymptotic t-ratios for the differences between the coefficients for increases and decreases were calculated, taking into account the differences in sign between coefficients. Given the different treatment used for connections and frequent flier benefits, likelihood-ratio tests cannot be used to compare the model to the others. Preference is given to the adjusted ρ^2 measure, that suggests that the performance of the asymmetrical model is superior to that of the base model, and the model allowing for a differential response to the attributes of the two alternatives. All coefficients are of the expected sign, with increases in desirable attributes being valued positively, and decreases negatively, with the converse applying in the case of undesirable attributes. However, three of the coefficients, $\beta^{-}_{\text{access time}}$, β^{+}_{OTP} and β^{+}_{FF}, are not significantly different from zero at any reasonable level of confidence. This is a result of the design of the survey, where increases in the tier of FF membership and on-time performance were presented relatively rarely, as were reductions in access time. The base model especially is unable to account for this, and the its parameter estimates are potentially biased as a result.

While the low significance levels of some of the parameters need to be taken into account, the results give an indication that, consistent with prospect theory, losses are valued more negatively than gains are positively, i.e. the coefficients associated with an amelioration are not as large as those associated with a reduction in attractiveness. The only exception to this arises in the case of $\beta_{\text{closest airport}}$. The asymmetry is especially noticeable for changes in air fare, where the difference, which attains a high level of statistical significance, is of the order to 2:1. In real terms, this would mean that airlines could expect much larger drops in passenger numbers following increases in air fares than increases in passenger numbers following corresponding drops in air fares.

There is an important difference between symmetrical and asymmetrical models in the calculation of trade-offs. With coefficients associated with increases as well as reductions in attribute values, we can calculate separate indicators for the willingness-to-pay for improvements in an attribute, and the willingness-to-accept a less desirable attribute value in return for a lower air-fare. The differences between these two ratios give an indication of the asymmetries in preference formation. As an example, we can see that a much bigger monetary incentive is required to accept an increase in the flight time by one hour than the corresponding willingness-to-pay for a reduction in flight time by one hour. The latter is lower than the symmetrical trade-off produced in the two previous models, while the former is higher. This gives an indication of the risk of biased results in symmetrical models. The models also suggest that the penalty resulting from a drop in on-time performance from 100% to zero is equivalent to the benefit of a reduction in air-fare by $120.

Table 3: Results allowing for asymmetrical preference formation

Respondents	532
Observations	4256
LL	-1498.8
par	16
adj. ρ^2	0.4865

	decreases		increases		
	est.	asy. t-rat.	est.	asy. t-rat.	asy. t-rat (diff)
β_{current}	0.3978	2.04	-	-	-
$\beta_{\text{access time}}$	0.0023	0.89	-0.0078	-5.81	1.78
$\beta_{\text{air fare}}$	0.0127	7.46	-0.0263	-5.20	2.75
$\beta_{\text{flight time}}$	0.0046	2.88	-0.0053	-5.73	0.39
β_{OTP}	-0.0151	-3.87	0.0058	1.32	1.38
β_{FF}	-0.3982	-3.25	0.0689	0.29	1.20
$\beta_{\text{closest airport}}$	-0.4706	-3.10	0.7661	3.49	-1.02

$\beta_{connecting}$	0.6211	3.52	-0.6666	-3.92	0.19
σ	1.0050	8.18	1.0050	8.18	-
Willingness to pay for improvements					
access time reductions ($/hour)	5.20[†]				
flight time reductions ($/hour)	10.39				
on time arrival ($)	22.19[†]				
gaining tier of FF benefits ($)	2.62[†]				
moving to closest airport ($)	29.10				
reduced number of connections ($)	23.59				
Drop in fare required to accept poorer conditions					
access time increases ($/hour)	-36.96				
flight time increases ($/hour)	-24.99				
late arrival ($)	-119.54				
drop in tier of FF benefits ($)	-31.42				
moving away from closest airport ($)	-37.14				
increased number of connections ($)	-52.60				

[†] calculation involves parameter not significant at the 95% level of confidence

5 Conclusions

With an increasing reliance on models of air travel behaviour estimated on SC data, there have been repeated efforts to raise the bar in survey design with the aim of improving response quality. One avenue has been to frame the choice situations around a real-world trip allowing the respondent to better relate to the scenarios that he or she is faced with. We have looked specifically at the situation where an observed trip is included as one of the alternatives in the SC survey. While not doubting the benefits of such a framing approach, we questioned the wisdom of using standard modelling approaches on such data. Specifically, the results give a strong indication that respondents in SC surveys do react differently to the attributes of reference, as compared to hypothetical, alternatives. The analysis suggests further that respondents may evaluate the attribute levels of purely hypothetical alternatives relative to those of the base alternative, with an asymmetrical response to gains and losses.

In terms of practical implications, a standard modelling approach is not appropriate when dealing with datasets that include a current trip as one of the alternatives. However, the results produced with the asymmetrical models are arguably more useful than those from a standard model estimated on datasets with no reference alternative. Indeed, results such as those produced here would for example allow airlines to forecast not only the increases in passenger demand following reductions in air fares, but would also allow them to look at likely reductions in passenger demand with higher fares. In datasets that are not framed around a current trip, such a treatment is diff-cult and the models would most likely have to work under the assumption of a symmetrical response to gains and losses.

Acknowledgements

The author would like to thank Tom Adler and Resource Systems Group for making the air travel survey data available. Part of the work was carried out while the author was working at Imperial College London and ETH Zurich. The comments of two anonymous referees helped to improve the paper.

References

Adler, T., Falzarano, C. S., Spitz, G., 2005. Modeling Service Trade-offs in Air Itinerary Choices. paper presented at the 84th Annual Meeting of the Transportation Research Board, Washington, DC.

Ben-Akiva, M., Morikawa, T., 1990. Estimation of Switching Models from Revealed Preferences and Stated Intentions. Transportation Research A 24, 485-495.

Bierlaire, M., 2005. An introduction to BIOGEME Version 1.4. biogeme.ep.ch.

Halton, J., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numerische Mathematik 2, 84-90.

Hess, S., Adler, T., Polak, J. W., 2007. Modelling airport and airline choice behavior with stated-preference survey data. Transportation Research E 43, 221-233.

Hess, S., Rose, J. M., Hensher, D. A., 2008. Asymmetrical Preference Formation in Willingness to Pay Estimates in Discrete Choice Models. Transportation Research E, forthcoming.

Louviere, J. J., Hensher, D. A., Swait, J., 2000. Stated Choice Models: Analysis and Application. Cambridge University Press, Cambridge.

Ortúzar, J. de D., Roncagliolo, D. A., Velarde, U. C., 2000. Interactions and independence in stated preference modelling. In: Ortúzar, J. de D. (ed.), Stated Preference Modelling Techniques: PTRC Perspectives 4. PTRC Education and Research Services Ltd, London.

Proussaloglou, K., Koppelman, F. S., 1999. The choice of air carrier, flight, and fare class. Journal of Air Transport Management 5, 193-201.

Resource Systems Group Inc., 2003. Air Travelers 2003: The New Realities? Annual Air Survey project report.

Starmer, C., 2000. Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. Journal of Economic Literature 38, 332-382.

Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, MA.