Using Stated Preference to Value Annoyance from Aircraft: A Comparison of Approaches

Abigail Bristow and Mark Wardman
Institute for Transport Studies
University of Leeds
Leeds LS2 9JT
abristow@its.leeds.ac.uk, mwardman@its.leeds.ac.uk

1. Introduction

The research reported here was carried out as part of the 5A (Attitudes to Aircraft Annoyance Around Airports) study which aims to gain a greater understanding of attitudes to all forms of annoyance in the vicinity of airports and is funded by the EUROCONTROL Experimental Centre. This exploratory work focussed on noise as a key source of annoyance amongst households around Bucharest, Lyon and Manchester Airports and sought to derive values of noise nuisance using different forms of Stated Preference (SP) method alongside attitudinal data relating to noise, annoyance and broader quality of life issues. In this paper we focus on the stated preference experiments, for comprehensive study results see Bristow and Wardman (2003) and Bristow et al (forthcoming).

SP experiments offer the decision maker hypothetical scenarios and the preferences expressed indicate the relative importance of the attributes that characterise the scenarios. The most common form of evaluation is choice, although ranking exercises are sometimes employed, and typically just two alternatives are compared with between nine and twelve comparisons involved and usually between four and six attributes characterising each alternative. In this paper the term stated preference (SP) is used to mean choice or ranking experiments and hence does not incorporate contingent valuation (CVM).

There has been a gradual appreciation of the advantages of SP methods and a movement towards their application in environmental valuation. Adamowicz (2000, p.353) concluded that, “SP methods are likely to continue to grow in use and popularity in the environmental economics profession”. More recently the Department for Transport has commissioned a manual on the use of the range of SP techniques (Bateman et al 2002). As the method has gained acceptance some practitioners (eg Bates 1998 and Wardman 2003) have argued for greater openness in discussing potential problems in SP. This paper is a contribution in that spirit, exploring the impact of different hypothesised incentives to bias in different experiments and at the same time reporting an innovative application of SP.

Although stated preference methods are being increasingly used in the valuation of environmental externalities, there are still relatively few examples in the valuation of transport externalities. In the context of transport noise a large number of studies have been undertaken using hedonic pricing methods (overviews may be found in Schipper et al 1998 and Nelson 2004) and a much smaller number have made use of CVM or SP methods. In the context of aircraft noise we were able to identify only one significant study using stated preference, by Thune-Larsen (1995) at Oslo Airport, although there are more studies applying CVM (see Navrud (2002) and Wardman and Bristow (2004) for an overview of studies of transport noise valuation). In this context our work was highly experimental and drew on our experience in valuing road traffic noise in Edinburgh and Lisbon (Wardman and Bristow 2004, Arsenio et al 2002).

At the outset we were concerned that naïve SP experiments that were obviously about disturbance from aircraft noise would provide an incentive to strategic or protest response bias. This led us to examine ways of successfully masking the true purpose of the SP experiment. In designing the SP experiments in this context the focus was on two areas:

- the risk of strategic or protest bias in responses; and
- the ability to produce results disaggregated by time period, day of the week, number of movements and aircraft type which is not possible using other techniques.

These issues are addressed in section 2 on the design of the SP experiments. Section 3 outlines the detailed design of the SP exercises. Section 4 contains results from the individual SP exercises, while Section 5 compares the outcomes and discusses the extent to which bias may have influenced the values in the different experiments. Conclusions are drawn in section 6.

2. SP Designs – Why Three Experiments?

A key advantage of SP lies in its ability to discriminate between characteristics of a good or bad. In this case we chose the number of planes, plane type, time of day and day of week which directly influence noise levels and their perception and valuation. These are elements that are to some degree under the control of the airport or regulator, and therefore the results will be relevant to operational decision making. However, an SP that is clearly focused on a particular issue such as aircraft noise is open to strategic or protest bias in responses, thus potentially losing one of its key advantages over other techniques, namely the ability to "mask" the true purpose of the experiment by placing the characteristic of interest amongst other variables and thereby reducing the incentive to bias.

2.1 Incentives to Bias

Bohm (1971) argued the theoretical case that stated willingness to pay should vary with the incentives to bias. This paper offers experiences from studies in the transport sector in Britain relating to incentives to bias. Evidence from elsewhere in the world is not considered in this paper.

Here we consider empirical evidence from SP studies in the transport sector in Britain that values are higher when the purpose of the survey is transparent. Wardman and Whelan (2001) report a meta-analysis based on 45 values (from 18 studies) of new or improved rolling stock on the UK rail system. The regression model is shown in Table 1. Where the purpose of the SP exercise was clearly stock valuation as opposed to such values being obtained from exercises with a more diverse range of attributes, the effect is to increase values by 202%. It is not unreasonable to interpret this as strategic bias. Other variables in the regression, familiarity with the type of rolling stock on offer, travelling for leisure purposes, valuing MarkIV stock relative to Mark3 stock and location on London suburban routes all serve to reduce the value.

Table 1 Regression model of rolling stock values

Variable	Coefficient (t)	% effect
Constant	1.846 (9.16)	Base = 6.33 of fare
Stock valuation	1.106 (5.53)	+202
Familiar	-0.576 (3.44)	-44
Leisure	-0.267 (1.69)	-23
MarkIV	-0.310 (1.81)	-27
London suburban	-0.439 (2.71)	-36
Adjusted R^2	0.54	

Source: Wardman and Whelan 2001

Wardman (2003) reports results from nine studies which sought to value overcrowding on trains in terms of time spent standing. In most of these studies the purpose was clear, but in the only two studies where overcrowding was an element of a broader study looking at aspects of mode choice and interchange, the values were clearly at the lower end of a wide range from 1.2 to 7 times the value of seated in-vehicle time. The values for overcrowding from SP exercises where the purpose of the study was clear appear to be high, but there is little RP evidence with which to compare them. Another contentious issue is arriving late at the destination. Wardman (1998) obtained 11 values of late arrival time. On average late arrival time was valued 7.4 times more highly than in-vehicle time. Again this looks very high and may be the result of response bias.

The introduction of tolling on new or existing roads is an issue that might be expected to be controversial. All other things being equal, the coefficients on different aspects of the money cost of motoring would be close to equal. In this case the hypothesis is that tolls attract a protest response. Wardman (2001) found that the value of time was 19% lower if the numeraire was a toll from a model estimated to 1167 values of time. Table 2 shows results from a subset of studies illustrating different choice contexts. Where a road is already tolled the difference between the coefficient on fuel costs and the toll is fairly small, at around 14%. A toll on a new road increases the difference to 23% on average. The clear protest occurs when a toll is introduced on an existing untolled road, where the coefficient is on average 240% higher. The difference in the earlier, more acceptable scenarios may be partly down to fuel costs not being fully taken into account in decision making. However, all these studies predate the 2000 fuel protests and the perception of the importance of fuel costs may have shifted.

Table 2 Variations in Fuel and Toll Coefficients from SP studies

Context, study and journey purpose		Fuel	Toll	Ratio
Existing tolled road – route choice				
1. 1996	Employers business	1.0	1.12	1.12
	Commercial	1.0	1.09	1.09
	Other	1.0	1.21	1.21
6. 1988	Commercial	-0.044	-0.049	1.11
	Leisure	-0.036	-0.043	1.19
Toll on a new road				
5. 1990	Commercial	1.0	1.30	1.30
	Leisure	1.0	1.15	1.15
Existing road – no toll at present				
3. 1994	Employers business	-0.0012	-0.0070	5.83
	Commercial	-0.0062	-0.0130	2.10
	Other	-0.0026	-0.0146	5.61
4. 1985	Urban	-0.0181	-0.0378	2.09
	Inter-urban	-0.0038	-0.0052	1.37

Adapted from Wardman 2003

Wardman and Shires (2001) estimated interchange penalties from revealed and stated preference data. Four different types of penalty were considered: transfer between trains to arrive earlier, departure time choice, route choice and whether to catch the next train and interchange or wait for a subsequent through train in the context where the intended train had been missed. The results are shown in Table 3. All values are expressed in terms of the equivalent length of in-vehicle time except departure time choice which is in units of wait time.

Table 3 Value of the interchange penalty from revealed and stated preference models (confidence intervals in brackets)

Context	RP	SP
Transfer choice	6.97 (±4.5)	8.85 (±4.0)
Departure time choice	5.71 (±4.1)	8.17 (±2.8)
Route choice	11.87 (±9.1)	17.48 (±12.4)
Train choice	0	13.70 (±4.0)

In all four cases the interchange penalty estimated to the SP data is higher than that estimated to the RP data. The authors conclude that this result is consistent with protest responses against having to change trains.

The available evidence from SP in transport suggests that where the object of the exercise is obvious, especially where the issue is contentious strategic bias is likely to occur. Indeed we would be surprised if this were not the case. What is surprising is the extent to which these somewhat alarming findings have been ignored.

Many practitioners insist that SP exercises must be kept relatively straightforward if reliable responses are to be obtained. This view is regarded as 'urban myth' by some, and according to Louviere (2001) there is no evidence to suggest that simpler SP exercises produce different valuations due to task complexity, but it is the prevailing view in Europe. There is evidence that indicates task complexity can influence valuations, largely through simplifying choice rules or ignoring attributes (Malhotra, 1982; Johnson and Meyer, 1984; Widlert, 1998; Arentze et al., 2003).

2.2 Conclusions

In section 2.1 we discussed the need to minimise incentives to bias by masking the true purpose of the survey by the inclusion of a range of other attributes. This ensures that we do not lose one of the key advantages of SP techniques. However another key advantage is the ability to discriminate between the characteristics of a good or bad. Clearly both cannot be done in one experiment. We designed one SP exercise which masked aircraft noise by including it within a set of quality of life issues. A second SP exercise which explored the different characteristics of the "good" in question, numbers of aircraft movements disaggregated by aircraft type within a time period. The purpose of this exercise would be readily apparent to respondents. This type of exercise is a standard SP. As this was an exploratory piece of work we were able to add a third experiment that looked directly at the influence of time period on values by looking at aircraft movements across time periods and where the purpose of the exercise is again transparent. We hypothesise that the values from the two experiments where the objective is clear will be higher due to strategic bias than in the experiment where the purpose is masked.

3. Survey design and implementation

The social survey was intended to identify the key influential variables on the value of noise nuisance from aircraft. The survey design was informed by a wide range of sources which included:

- Fifteen focus groups conducted in Manchester, Lyon and Bucharest for this study, reported in Heaver (2002), and evidence from focus groups conducted in earlier studies, including those carried out for the Diamond et al (2000) work on aircraft noise sleep and health.
- The body of literature exploring the variables that influence annoyance from noise and the value of noise from transport sources. Review papers such as that by Miedema and Vos (1999) on the influence of demographic and attitudinal factors and Fields (1998) on the role of ambient noise were of particular assistance as were valuation studies that identified key variables.
- Guidance on standardised noise reaction questions for use in community noise surveys (Fields et al 2001) which forms the basis of the new BSI (2003) and previous questionnaires such as that implemented in the large scale 1985 UK study of aircraft noise (Brooker et al, 1985).

By drawing on earlier studies we endeavoured to ensure that no key influential variables were omitted and that where appropriate questions were asked in such a way as to make comparisons between studies possible.

The original intention was to conduct household interviews using a computer based survey. However, as we wished to use aircraft noise simulation software and a range of different SP experiments, it was ultimately decided that the best option was a paper based survey implemented in hall test like conditions where staff could assist respondents as necessary and explain each part of the survey before it started. The survey was distributed in parts to prevent respondents from looking ahead.

The survey was implemented in late 2002 at locations around the airports of Bucharest, Lyon and Manchester. The final sample size was 647 with a minimum of 200 respondents at each airport.

Manchester Airport was selected as representative of a large regional airport exhibiting growth over time. Lyon has some similarities with Manchester, as a large regional airport, but with a much more rural population. The populations around these two airports are also comparable in terms of income levels. Otapeni was chosen as an example of an airport in an accession country. It has a smaller number of aircraft movements and income levels are much lower. This range of airports should enable the identification of differences arising from airport characteristics and cultural variation. As we are interested in incentives to bias in this paper, recent developments at the airports are of interest. A second runway opened at Manchester in 2001. This was highly controversial at the time, but the situation appears largely to have settled down now. However, one adversely affected area is included in our study. At Lyon proposals to build two new runways have been approved and there is an active opposition group ACENAS. Bucharest has seen no significant changes since 1997 when the new international terminal opened. We might then expect the greatest levels of awareness and incentives to bias at Lyon.

In this study we developed one SP exercise where the purpose of the survey was masked and two where the focus was clear, but which enabled disaggregation of characteristics. The three SP experiments are elaborated in sections 3.1 to 3.3.

3.1 SP1: Quality of Life SP Exercise

The objective of this SP exercise was to mask the purpose of the study, at least prior to asking detailed questions about aircraft movements, noise and annoyance, and this was to be achieved through the use of a large number of 'quality of life' variables. The focus groups had shown that discussion of quality of life issues successfully masked the true object of the exercise. The common theme amongst the attributes was that to varying degrees they could be regarded to be under the influence of local authorities. The variables chosen were influenced in part by the focus groups and were:

- Local crime levels, expressed as annual burglaries per 1000 households.
- Local school quality, represented by a percentage pass rate
- The level of area wide traffic congestion, specified as a proportionate variation of 10%, 5%, 0%, -5% and –10% on the current situation
- Street cleanliness, defined on a scale of very dirty and untidy, dirty and untidy, neither clean nor dirty, clean and very clean
- Traffic noise experienced at home, specified as extremely noisy, very noisy, moderately noisy, slightly noisy and not at all noisy
- Neighbourhood air quality, which took the levels of very poor, poor, neither good nor poor, good and very good
- The general condition of local roads and pavements, which could be very poor, poor, neither good nor poor, good and very good
- Recreation facilities, which involved the presence or not of a local library and sports/leisure facilities
- Local amenities, which was whether there were food shops and a doctors within walking distance
- Two means of presentation were used for aircraft noise. The number of aircraft movements varying around the current situation and distinguishing between daytime and evening flights or the level of aircraft noise on the same scale as for traffic noise
- Variations in tax

Bearing in mind the number of evaluations that respondents would have to make, five levels were chosen for each variable. The exception to this was cost which had seven levels in order to introduce more variation into this key variable and to allow for uncertainty as to households' valuations.

The aircraft movements used are listed in Table 4. The central column was our best estimate of a typical number of current 'encounters' in terms of planes per hour in the daytime and evening periods. These were defined as 'planes going by', and hence are half of the total number of movements. No distinction was made between take-offs and landings.

Table 4: Levels of Aircraft Movements per Hour

Location	Time	Minutes between planes going by				
Manchester	Daytime Evening	2 2	4 2	4 4	4 7.5	7.5 7.5
Lyon	Daytime Evening	3 2	6 2	6 5	6 15	10 15
Bucharest	Daytime Evening	5 10	20 10	20 30	20 60	60 60

To further increase the variation in the cost variable, two sets of cost variation were used, with respondents randomly allocated one of them. These levels are given in Table 5. The levels for Lyons and Manchester are roughly equivalent while those for Bucharest are much lower reflecting the lower income levels.

Table 5: Cost Variations Used

Manchester	-£8 -£10	-£3 -£5	-£1 -£2	0 0	£1 £2	£3 £5	£5 £10
Lyon	-€13.00 -€16.25	-€4.88 -€8.13	-€1.63 -€3.25	0 0	€1.63 €3.25	€4.88 €8.13	€13.00 €16.25
Bucharest	-L40000 -L50000	-L20000 -L30000	-L5000 -L10000	0 0	L5000 L10000	L20000 L30000	L40000 L50000

Note: These are weekly cost variations. In the case of Lyon, monthly variations were presented to respondents.

Accommodating such a large number of attributes in a conventional choice experiment is feasible, but the demands placed upon individuals in trying to evaluate two options characterised by eleven attributes would be considerable. Serious doubt would have to be placed on whether respondents could provide reliable answers to such an exercise and they could resort to alternative choice rules in order to simplify the task.

An alternative approach was adopted, which effectively involves a one-dimensional rather than a multi-dimensional consideration of attribute variation. Our view is that if offered a whole series of improvements to specific attributes, respondents can more readily state which attribute variation they would most like to achieve than they can weigh up the net benefit of differences in a whole range of attributes between two alternatives. Table 6 illustrates this form of SP exercise

The starting point is to identify the respondent's current situation. For most variables the current level was predefined and shaded. For the remaining variables: street cleanliness, traffic noise at home, neighbourhood air quality, level of aircraft noise (for the categorical scale), general condition of local roads and pavements, recreation facilities and local amenities respondents were asked to mark their starting position.

The respondent was asked to consider the improvements to the current situation. These were all the levels of the attributes to the right of the current levels. The respondent was asked to state which improvement would be most preferred. This should logically be an attribute level in the right hand column. They were then asked to disregard this improvement, treating it as if it were no longer available, and asked to state which was the second best improvement. This process continued until all the possible improvements had been ranked in order of preference.

Table 6: Example of SP1: Manchester (Cheadle)

<table>
<tr><td>Local Crime: Burglaries per 1000 Households</td><td>10</td><td>5</td><td>2</td><td>1</td><td>0.5</td></tr>
<tr><td>Local Schools: % GCSE Pass Rate</td><td>10%</td><td>25%</td><td>40%</td><td>55%</td><td>70%</td></tr>
<tr><td>Area Wide Road Traffic Congestion</td><td>10% More Traffic</td><td>5% More Traffic</td><td>As Now</td><td>5% Less Traffic</td><td>10% Less Traffic</td></tr>
<tr><td>Street Cleanliness</td><td>Very Dirty and Untidy</td><td>Dirty and Untidy</td><td>Neither Clean nor Dirty</td><td>Clean</td><td>Very Clean</td></tr>
<tr><td>Traffic Noise at Home</td><td>Extremely Noisy</td><td>Very Noisy</td><td>Moderately Noisy</td><td>Slightly Noisy</td><td>Not at all Noisy</td></tr>
<tr><td>Neighbourhood Air Quality</td><td>Very Poor</td><td>Poor</td><td>Neither Good nor Poor</td><td>Good</td><td>Very Good</td></tr>
<tr><td>General Condition of Local Roads and Pavements</td><td>Very Poor</td><td>Poor</td><td>Neither Good nor Poor</td><td>Good</td><td>Very Good</td></tr>
<tr><td>Planes Go By</td><td>Every 2m Daytime Every 2m Evenings</td><td>Every 4m Daytime Every 2m Evenings</td><td>Every 4m Daytime Every 4m Evenings</td><td>Every 4m Daytime Every 7½m Evenings</td><td>Every 7½m Daytime Every 7½m Evenings</td></tr>
<tr><td>Council Tax</td><td>£8 more a week</td><td>£3 more a week</td><td>£1 more a week</td><td>As Now</td><td>£1 less a week</td><td>£3 less a week</td><td>£8 less a week</td></tr>
<tr><td rowspan="2">Recreation Facilities Locally Available</td><td colspan="2">No Library</td><td colspan="3">Library</td></tr>
<tr><td colspan="2">No Sports/Leisure Facilities</td><td colspan="3">Sports/Leisure Facilities</td></tr>
<tr><td rowspan="2">Amenities Within Walking Distance</td><td colspan="2">No Local Food Shops</td><td colspan="3">Local Food Shops</td></tr>
<tr><td colspan="2">No Local GP</td><td colspan="3">Local GP</td></tr>
</table>

Our initial preference had been to ask individuals to start at the current situation and gradually improve it by moving to the right one attribute at a time. In contrast to the approach that was actually adopted, which essentially involves evaluating the improvements from the right to the as now situation, this means that the reference situation is continually updated to contain the improvements that have been selected. Whilst members of the study team felt this was the simpler procedure for respondents to undertake, there are two problems facing it. The first, and critical problem, is that this procedure would build in linear dependency between the attribute levels. As a result, it would not be possible to analyse the data. Secondly, it might be that an individual selected an improvement that is not actually preferred simply to subsequently achieve an improved level of that attribute which was preferred overall.

Having completed the ranking of improvements, the respondent then proceeded to evaluate the deteriorations. The deterioration which was regarded to be worst was first identified, followed by the second worst deterioration and so on until all the deteriorations had been evaluated.

3.2 SP2: Standard SP Exercise

The second SP exercise took a much more conventional form. It presented each individual with eight comparisons of two options, each of which was described in terms of the number of planes passing by and the local council tax. The number of planes was disaggregated into three plane types of: 747 jumbo jets and other large planes; two engined jets, such as Airbus and 737's; and turbo-prop aircraft, and pictures of the different types of aircraft were shown.

In addition, respondents were asked to consider the variations in a specific time period, on the grounds that the annoyance caused by aircraft movements will depend on the exposure to it and the activities being undertaken when the noise is experienced. Variation in the number of flights to some extent requires different time period designs for the purposes of realism. The time periods used were: weekdays 6am-9am, weekdays 9am-6pm, weekdays 6pm-10pm, Saturday 6am-9am, Saturday 9am-6pm, Saturday 6pm-10pm, Sunday 9am-6pm and night.

An example of a choice which respondents were asked to make is given in Table 7 below. Option A always involved more or the same number of planes per hour but was cheaper than Option B. The number of planes passing by was specified both as the number per hour and as the interval since the focus groups revealed that some preferred one form of representation and others preferred the other.

A standard fractional factorial design procedure was used to combine the levels of the attributes in each scenario to be evaluated. The full number of scenarios produced by the design was sixteen, but any individual was presented only with eight of them. Simulation tests were undertaken on the designs using synthetic data to ensure that they were satisfactory from a statistical perspective. This led to a number of modifications to increase the precision with which the parameters could be expected to be estimated.

Table 7: Example SP2 Choice

	Time Period: Sunday 9am-6pm	
	A	B
Jumbo jets/large 4 engine planes	1 per hour (About every 60 minutes)	2 per hour (About every 30 minutes)
Two engine jets (eg, 737, Airbus)	20 per hour (About every 3 minutes)	30 per hour (About every 2 minutes)
Turbo-prop (propeller) planes	1 per hour (About every 60 minutes)	2 per hour (About every 30 minutes
Total Flights	22 per hour (About every 2½ minutes)	34 per hour (About every 2 minutes)
Weekly council tax	£2 more	£5 less
I would choose ….	A	B

3.3 SP3

The aim of this SP exercise was to examine within a single experiment preferences towards variations in aircraft noise across different time periods. Given the large number of time periods that need to be used, in order to distinguish between different amounts of airport activity and also variations in respondents' exposure levels, a conventional SP choice exercise was not used. Instead, the same sort of procedure was used as in SP1. Table 8 presents the exercise used in Manchester. As in the other two SP exercises, the number of planes passing by was specified, and no distinction was made between take-offs and landings.

As in SP1, two different sets of cost levels were used and randomly offered to respondents (the cost levels are those in Table 5). In this exercise, the current situation is predefined. Respondents first evaluated the improvements, stating which was the most preferred, the second best, the third best and so on, and then evaluated the deteriorations.

Given that the interview included two other SP exercises along with a large number of attitudinal, socio-economic and demographic variables, this SP exercise was only offered to those who had completed the rest of the questionnaire ahead of other respondents.

Table 8: Manchester SP3 Exercise

	Worse			As Now	Better		
Every Weekday 6-9am	60	40	30	20	15	12	10
Every Weekday 9am-6pm	40	30	20	15	12	10	6
Every Weekday 6-10pm	30	20	15	12	10	6	4
Saturday 6-9am	60	40	30	20	15	12	10
Saturday 9am-6pm	40	30	20	15	12	10	6
Saturday 6-10pm	30	20	15	12	10	6	4
Sunday 9am-6pm	40	30	20	15	12	10	6
Every Night	6	4	3	2	1	0	
Tax	£10 more	£5 more	£2 more	0	£2 less	£5 less	£10 less

4. SP Results

4.1 SP1

The data was analysed using an ordered logit model (Chapman and Staelin, 1982) and a jack-knife procedure within the ALOGIT package was used to account for individuals' repeat observations (Cirillo et al., 2000). Separate models have been estimated for each location and for improvements and deteriorations. Crime, school pass rates and traffic are in the units reported in section 3 whilst the tax variable has been converted into €. All other terms are dummy variables, representing the impact of a level of a variable relative to a base category. Given that not all individuals have the same current situation, the base category used can vary across locations, effectively depending on the balance of respondents amongst categories. This makes interpretation of the results less straightforward than if there was a common base category.

The results for improvements are contained in Table 9. Removed from the data set are those individuals who failed to rank the alternatives in logical order, such as preferring, say, a £3 per week tax saving to a £5 per week tax saving. As demonstrated by Bristow and Wardman (2003), removing these individuals does not alter our conclusions.

The models achieve goodness of fit measures (ρ^2) in line with those typically achieved in more conventional SP choice models and a large number of coefficients are statistically significant. Given that there were reasonable numbers of respondents whose current situation was towards the left in Table 6, where the attribute levels are relatively poor, there are as a result a large number of possible improvements that can be valued.

The purpose here is not to examine the quality of life results in detail, although we note that there is a reasonable degree of consistency with quality of life importance and satisfaction ratings that were collected (Bristow and Wardman, 2003). There is a close degree of correspondence in many of the values of the Manchester and Lyon samples. As expected, and due to the large differences in incomes and the much lower number of aircraft movements, the values for the Bucharest sample are generally much lower than for the other two.

Table 9: Quality of Life Improvements

	Manchester		Lyon		Bucharest	
	Coeff (t)	Value (t)	Coeff (t)	Value (t)	Coeff (t)	Value (t)
Crime	-0.502 (12.8)	3.89 (9.1)	-0.690 (18.6)	3.71 (16.8)	-1.187 (10.2)	0.85 (6.8)
School Pass Rate	0.072 (9.3)	0.56 (5.5)	0.105 (18.1)	0.56 (16.3)	0.101 (4.2)	0.07 (4.1)
Traffic	-0.192 (8.3)	1.49 (8.2)	-0.244 (19.2)	1.31 (16.9)	n.s	
Sport Facilities	1.320 (2.2)	10.23 (2.2)	n.s		n.s	
Shop Facilities	1.449 (1.9)	11.23 (2.4)	0.465 (6.9)	2.50 (6.7)	n.s	
Local Doctor	2.296 (3.9)	17.80 (3.8)	n.s		n.s	
Library	n.s		0.403 (7.1)	1.62 (6.5)	n.s	
Aircraft: Very Noisy:	Base		Base		Base	
Aircraft: Moderately Noisy	0.876 (2.6)	6.79 (2.6)	1.705 (5.6)	9.17 (5.5)		
Aircraft: Slightly Noisy	1.679 (5.6)	13.02 (5.4)	3.299 (11.2)	17.74 (10.5)	0.355 (2.1)	0.26 (2.1)
Aircraft: Not at all Noisy	3.101 (8.3)	24.04 (7.5)	4.256 (13.1)	22.88 (12.1)	0.796 (3.2)	0.57 (3.1)
Air Movements: Current	Base		Base		Base	
Air Movements: 2nd Best	0.351 (1.9)	2.72 (1.9)	1.974 (11.9)	10.61 (11.3)	-	-
Air Movements: Best	1.318 (6.5)	10.22 (6.3)	2.645 (13.6)	14.22 (12.7)	0.872 (5.1)	0.63 (4.4)
Streets: Dirty/Untidy	Base		Base		Base	
Streets: Neither			1.271 (2.8)	6.83 (2.6)		
Streets: Clean	1.169 (8.1)	9.06 (7.1)	2.523 (6.2)	13.56 (6.1)	0.564 (5.9)	0.41 (4.9)
Streets: Very Clean	2.562 (9.8)	19.86 (9.0)	3.713 (8.9)	19.96 (8.7)	2.099 (12.8)	1.51 (7.1)
Traffic: Extremely Noisy	Base		Base		Base	
Traffic: Very Noisy			0.932 (1.9)	5.01 (1.9)		
Traffic: Moderately Noisy			1.871 (2.4)	10.06 (2.4)		
Traffic: Slightly Noisy	1.153 (6.2)	8.94 (5.6)	3.397 (4.0)	18.26 (4.0)	0.450 (2.4)	0.32 (2.4)
Traffic: Not at all Noisy	2.233 (8.7)	17.30 (7.8)	4.521 (5.4)	24.31 (5.5)	1.336 (6.4)	0.96 (5.4)
Air Quality: Poor/Very Poor	Base		Base		Base	
Air Quality: Neither	0.754 (4.7)	5.84 (4.4)	1.445 (3.1)	7.77 (3.1)		
Air Quality: Good	1.443 (7.1)	11.19 (6.7)	3.204 (6.5)	17.23 (6.5)	0.392 (2.6)	0.28 (2.5)
Air Quality: Very Good	2.571 (8.8)	19.93 (8.1)	4.431 (9.1)	23.82 (9.0)	2.278 (10.1)	1.64 (5.4)
Road Conditions: Very Poor	Base		Base		Base	
Road Conditions: Poor			0.691 (1.8	3.72 (1.8)		
Road Conditions: Neither			1.405 (2.8)	7.55 (2.8)		
Road Conditions: Good	1.066 (7.6)	8.26 (7.0)	2.600 (5.3)	13.98 (5.3)	0.232 (2.1)	0.17 (2.1)
Road Conditions: Very Good	2.422 (10.7)	18.78 (9.3)	3.498 (7.1)	18.81 (7.1)	1.738 (12.6)	1.24 (7.0)
Tax(€)	-0.129 (9.2)		-0.186 (19.4)		-1.392 (7.1)	
ρ^2	0.106		0.097		0.119	
Individuals	109		130		67	

Table 10 converts the values for the different aircraft situations into values for a change in frequency of one plane per hour throughout the specified time period, given the levels specified in Table 3. The values are derived from the *Aircraft Movements 2nd Best* and *Aircraft Movements Best* coefficients. The former indicates the value of variations in evening flights and the difference between the two indicates the incremental effect of variations in the number of daytime flights. However, in the case of Bucharest, it was not possible to estimate a statistically significant coefficient for *Aircraft Movements 2nd Best*. There are two ways in which we might interpret the result. One is to conclude that evening flights have no value, whereupon the overall valuation would be divided by the variation in daytime flights per hour to yield a value per flight in each daytime hour of €0.31. The other is to conclude that the result is a statistical aberration and that evening flights do have value. In this latter case, we have to make an assumption about the relative valuation of aircraft during the day and during the evening. The most sensible would seem to be to assume that they have the same value. The overall valuation is €0.63 and this is decomposed to give the values in the final column of Table 10.

Table 10: Value per Week per Plane Passing by Every Hour (Improvements) During the Time Period

	Manchester	Lyon	Bucharest	
Daytime	1.07	0.90	0.31	0.27
Evening	0.39	1.33	0.00	0.09

These household valuations for Manchester and Lyon do not seem unreasonable. The larger value in the evening than daytime for Lyon residents is consistent with the findings of the focus groups and the attitudinal results and stems from the activities being pursued in an evening. The larger daytime value in Manchester is presumably due to the longer time period it represents, the value per individual movement is actually higher in the evening as the time period is much shorter. However, the Bucharest values do appear to be relatively large. It may be that Bucharest respondents are less likely to believe that any reduction in taxes could occur.

The models for deteriorations to the current situation are reported in Table 11. Although this exercise was more difficult for respondents the goodness of fit tends to be somewhat higher.

Given that the current situations on balance were towards the left of Table 6, with relatively poor levels of the attributes, there was less scope for further deterioration than had the current situations been generally more favourable. As a result, fewer coefficients have been estimated for deteriorations than was the case for improvements. Nonetheless, a large number of statistically significant coefficients have been estimated.

Table 11: Quality of Life Deteriorations

	Manchester		Lyon		Bucharest	
	Coeff (t)	Value (t)	Coeff (t)	Value (t)	Coeff (t)	Value (t)
Crime	-0.292 (14.4)	3.79 (8.0)	-0.274 (14.2)	4.22 (8.1)	-0.229 (10.9)	0.09 (8.5)
School Pass Rate	0.031 (4.8)	0.40 (4.7)	0.019 (3.4)	0.29 (3.4)	0.029 (4.7)	0.01 (4.6)
Traffic	-0.103 (9.2)	1.34 (5.5)	-0.120 (13.1)	1.86 (6.6)	-0.053 (4.7)	0.02 (4.0)
Sport Facilities	-0.387 (1.9)	5.03 (1.8)	-0.147 (1.8)	2.26 (1.8)	-0.399 (1.9)	0.15 (1.9)
Shop Facilities	n.s		n.s	-	-0.012 (1.7)	0.005 (1.7)
Library	-0.435 (3.3)	5.65 (2.9)	-0.471 (3.9)	7.25 (3.7)	-0.419 (1.5)	0.16 (1.5)
Aircraft: Moderately Noisy	Base		Base		Base	
Aircraft: Very Noisy			-0.838 (3.6)	12.89 (3.4)		
Aircraft: Extremely Noisy	-0.903 (6.3)	11.73 (5.4)	-2.311 (8.9)	35.50 (5.8)	-0.441 (2.7)	0.17 (2.7)
Air Movements: Current	Base		Base		Base	
Air Movements: 2nd Worst	n.s		-1.390 (8.4)	21.38 (6.4)	n.s	
Air Movements: Worst	-0.933 (5.1)	12.12 (4.8)	-2.226 (12.7)	34.25 (7.5)	-0.764 (3.5)	0.29 (3.4)
Streets: Not Very Dirty/Untidy	Base		Base		Base	
Streets: Very Dirty and Untidy	-0.946 (8.6)	12.29 (6.8)	-0.851 (10.5)	13.09 (7.3)	-0.832 (7.0)	0.32 (6.1)
Traffic: Not Noisy	Base		Base		Base	
Traffic: Very Noisy			-0.839 (9.2)	12.91 (6.8)		
Traffic Extremely Noisy	-0.872 (7.7)	11.32 (6.4)	-1.978 (16.7)	30.43 (8.3)	-0.747 (9.8)	0.29 (7.5)
Air Quality: Not Poor	Base		Base		Base	
Air Quality: Poor			-0.709 (9.1)	10.91 (6.8)		
Air Quality: Very Poor	-0.711 (5.4)	9.23 (5.1)	-1.542 (13.6)	23.72 (7.9)	-1.068 (8.0)	0.41 (6.7)
Road Conditions: Not Very Poor	Base		Base		Base	
Road Conditions: Very Poor	-0.585 (4.8)	7.6 (4.6)	-0.658 (8.2)	10.12 (6.5)	-0.559 (4.9)	0.22 (4.6)
Tax (€)	-0.077 (8.3)		-0.065 (8.6)		-2.590 (9.8)	
ρ^2	0.133		0.119		0.131	
Individuals	133		153		84	

As expected, the Bucharest values are again somewhat lower than the values for Manchester and Lyon. As for the latter two groups of residents, there are several similarities but some notable instances where the Lyon values are higher. These include the categorical values of aircraft noise and traffic noise and the value of air quality, which is in line with the findings for improvements.

Table 12 converts the values for different aircraft situations into values per plane passing by. In this case evening values had to be estimated for Manchester and Bucharest.

Table 12: Value per Week per Plane Passing by Every Hour (Deteriorations) During the Time Period

	Manchester		Lyon	Bucharest	
Daytime	0.81	0.61	1.29	0.032	0.028
Evening	0.00	0.20	1.19	0.00	0.009

The values for Manchester and Lyon seem reasonable, although now the daytime value for Lyon is slightly higher. The overall value for Manchester lies between the values for improvements in daytime (1.07) and evening (0.39) flights whilst there is a very encouraging degree of correspondence between the values of deteriorations and improvements in the case Lyon residents. The most disappointing result is the very large difference in values for Bucharest. Whilst we have argued that the large values for improvements could stem from a perception that tax reductions would not occur, and hence are given little weight and thereby imply large values, it could be that the value for deteriorations is lower because it contains the perceived broader economic benefits of airport expansion and an aversion to higher tax levels. This is apparent in the results of the second SP exercise and was apparent in the focus groups and attitudinal results.

4.2 SP2

The SP choice data was analysed using a standard logit model and allowance was made for repeat observations using the jack-knife procedure as for SP1. The results of this SP experiment are reported in Table 13.

Table 13: SP2 Results

	Manchester		Lyon		Bucharest	
	Coeffs (t)	Values (t)	Coeffs (t)	Values (t)	Coeffs (t)	Values (t)
Constant-Quieter	-		1.2899 (5.0)		-1.2064 (6.4)	
Flights - Weekday 6am-9am	-	-	-0.0635 (1.9)	1.29 (1.8)	-0.0895 (2.9)	0.28 (1.8)
Flights - Weekday 9am- 6pm	-0.0277 (1.4)	0.55 (1.5)	-0.0303 (1.2)	0.61 (1.2)	-0.0984 (2.6)	0.31 (1.7)
Flights - Weekday 6pm-10pm	-0.0686 (3.5)	1.37 (3.9)	-0.0821 (3.2)	1.66 (2.9)	-0.0865 (2.5)	0.27 (1.7)
Flights - Saturday 6am-9am	-	-	-	-	-0.1061 (3.5)	0.33 (1.9)
Flights - Saturday 9am-6pm	-0.0726 (4.3)	1.45 (4.1)	-0.0250 (1.0)	0.51 (1.0)	-	-
Flights - Saturday 6pm-10pm	-	-	-0.0463 (1.7)	0.94 (1.7)	-	-
Flights – Sunday	-0.0869 (3.2)	1.73 (3.5)	-0.0256 (1.0)	0.52 (1.0)	-0.0914 (2.3)	0.29 (1.6)
Flight – Night	-0.1921 (2.1)	3.83 (2.3)	-0.0761 (1.8)	1.54 (1.7)	-0.1032 (1.9)	0.32 (1.5)
Tax (€)	-0.0501 (4.7)		-0.0494 (7.2)		-0.3204 (2.3)	
ρ^2	0.070		0.059		0.032	
Obs	1545		1647		1895	

The goodness of fit measures are low, particularly for Bucharest where the respondents struggled more with the SP exercises, and they are lower than for SP1. However, the identification and removal of irrational responses is not possible in this exercise.

Although some successful segmentation by plane type had been conducted (Bristow and Wardman, 2003), the sample sizes are such that segmenting both by plane type and time period is not worthwhile. The models reported here relate to the valuation of the total number of planes per hour with segmentation of the flight coefficient according to time period. For a few time periods, it was not possible to obtain coefficient estimates that were remotely significant and we conclude that this is due to the relatively small sample sizes.

For Lyon residents, the best fitting model contained a constant term denoting a dislike of the option which involved more flights. We interpret this as a protest response. On the other hand, the Bucharest residents had a preference for the option with more flights, other things equal, and it may be that this is linked to the broader economic benefits that result from airport expansion.

Interpretation of the results is hampered by the relatively low level of precision with which the coefficients are estimated. Nonetheless, there are some plausible relativities. Not surprisingly, night values are the highest. Lyon residents had registered their dislike of evening aircraft movements in the focus groups, indicating that it disturbed activities, such as dining outdoors, and hence the relatively high value for evenings. Similar arguments may apply to Manchester residents on Sundays. The weekday daytime values are relatively low, and this is likely to be because there are fewer at home at these times. However, the Bucharest values hardly vary by time period, despite the relatively large confidence intervals, and we conclude that this may stem from difficulties faced with this task and the greater prevalence of random responses.

4.3 SP3

SP3 is analysed in the same way as SP1. Table 14 presents the coefficient estimates and the implied monetary values. A number of interesting findings have emerged.

Table 14: SP3 Results

	Improvements			Deteriorations		
	Manchester	Lyon	Bucharest	Manchester	Lyon	Bucharest
Coefficients						
Weekday 6-9am	-0.192 (3.4)	-0.229 (4.3)	-0.998 (2.6)	-0.057 (3.5)	-0.100 (4.2)	-0.201 (8.1)
Weekday 9am-6pm	-0.225 (3.7)	-0.316 (3.9)	-0.706 (3.9)	-0.069 (3.1)	-0.062 (1.7)	-0.204 (8.9)
Weekday 6-10pm	-0.357 (5.4)	-0.255 (7.2)	-1.489 (5.4)	-0.109 (3.1)	-0.080 (3.7)	-0.214 (9.8)
Saturday 6-9am	-0.244 (4.9)	-0.441 (7.6)	-1.766 (6.5)	-0.034 (3.4)	-0.094 (6.3)	-0.207 (14.1)
Saturday 9am-6pm	-0.283 (5.3)	-0.500 (5.9)	-1.009 (6.9)	-0.071 (3.0)	-0.098 (3.8)	-0.219 (11.2)
Saturday 6-10pm	-0.304 (5.4)	-0.768 (6.5)	-1.993 (7.4)	-0.090 (3.3)	-0.121 (4.0)	-0.229 (11.0)
Sunday	-0.264 (4.6)	-0.684 (7.0)	-1.076 (6.4)	-0.059 (3.4)	-0.153 (7.6)	-0.257 (14.3)
Night	-0.828 (2.5)	-1.218 (2.0)	-2.958 (4.9)	-0.500 (4.5)	-0.999 (5.5)	-0.749 (11.4)
Tax (Euro)	-0.141 (3.5)	-0.072 (4.4)	-4.210 (6.0)	-0.227 (5.8)	-0.092 (4.0)	-6.801 (6.7)
Values						
Weekday 6-9am	1.36 (3.7)	3.18 (3.8)	0.24 (3.0)	0.25 (3.6)	1.09 (3.4)	0.03 (6.4)
Weekday 9am-6pm	1.60 (3.9)	4.39 (3.7)	0.17 (4.7)	0.30 (3.2)	0.67 (1.8)	0.03 (6.4)
Weekday 6-10pm	2.53 (4.3)	3.54 (4.6)	0.35 (6.0)	0.48 (3.2)	0.87 (3.2)	0.03 (6.7)
Saturday 6-9am	1.73 (4.3)	6.13 (4.6)	0.42 (6.8)	0.15 (3.5)	1.02 (3.9)	0.03 (6.7)
Saturday 9am-6pm	2.01 (4.3)	6.94 (4.4)	0.24 (7.1)	0.31 (3.1)	1.07 (3.2)	0.03 (7.1)
Saturday 6-10pm	2.16 (4.3)	10.67 (4.5)	0.47 (7.2)	0.40 (3.4)	1.32 (3.3)	0.03 (6.8)
Sunday	1.87 (4.2)	9.50 (4.6)	0.26 (6.8)	0.26 (3.5)	1.66 (4.1)	0.04 (7.1)
Night	5.87 (2.9)	16.92 (2.0)	0.70 (5.4)	2.20 (4.5)	10.86 (3.7)	0.11 (7.1)

Note: The values are per week and relate to a one unit change in the number of planes per hour for each hour during the time period specified

Firstly, improvements are valued considerably and consistently more highly than deteriorations across all three locations. The effect is much stronger than anywhere else in this study. It will now have become very evident what the purpose of the study is and this could have led to strategic biases which have operated in favour of reductions in the number of aircraft.

Secondly, the monetary values are generally very large, particularly for the improvements, which is again consistent with the operation of strategic bias. Whilst package (part-whole) effects might be present, in that the valuations depend on how many improvements are bought such that improvements during weekday evenings would not have the same value if improvements on Sundays and at night were also being purchased, even the values for specific improvements are often high.

Finally, although the absolute values generally do not seem plausible, the values in relation to each other do seem to contain useful information. For example, variations in the number of aircraft at night are invariably valued more highly, as expected, than variations in daytime flights whilst it is the evening period amongst weekdays where the values tend to be highest and Saturday has values as high or higher than weekdays even though it forms a lower proportion of time since people are then more likely to be at home.
The exception to this is Bucharest for deteriorations, though even here there is clearly a higher value placed on nighttime movements. A possible explanation is that there has again been an influence from the confounding effect of the economic benefits associated with airport expansion and that respondents would not wish to lose these benefits. There is also a large coefficient on taxation indicating an aversion to increased tax levels. This

could be a result of the absence of a true payment mechanism for improvements, as reductions in aircraft movements are traded against reductions in tax.

5. Comparison of Values from the Three SP Experiments

Table 15 provides weekly valuations of the number of planes (events) per hour for the relevant time periods for each of the three SP exercises and the three locations. For comparability, the values by time period are also aggregated into overall daytime and evening values. All values are expressed in €.

Table 15: SP Monetary Values per Week – € per Plane per Hour within a Time Period (t statistics in brackets)

SP		Period	Manchester		Lyon	Bucharest	
1	Imp	Daytime	1.07		0.90	0.31	0.27
1	Imp	Evening	0.39		1.33	0.0	0.09
1	Det	Daytime	0.81	0.61	1.29	0.032	0.028
1	Det	Evening	0.0	0.20	1.19	0.0	0.009
2	Both	Weekday 6am-9am	-		1.29 (1.8)	0.28 (1.8)	
2	Both	Weekday 9am- 6pm	0.55 (1.5)		0.61 (1.2)	0.31 (1.7)	
2	Both	Weekday 6pm-10pm	1.37 (3.9)		1.66 (2.9)	0.27 (1.7)	
2	Both	Saturday 6am-9am	-		-	0.33 (1.9)	
2	Both	Saturday 9am-6pm	1.45 (4.1)		0.51 (1.0)	-	
2	Both	Saturday 6pm-10pm	-		0.94 (1.7)	-	
2	Both	Sunday	1.73 (3.5)		0.52 (1.0)	0.29 (1.6)	
2	Both	Night	3.83 (2.3)		1.54 (1.7)	0.32 (1.5)	
2	Both	Daytime	2.00 [3.30]		2.41 [2.80]	0.92 [1.14]	
2	Both	Evening	1.37 [1.80]		2.60 [2.73]	0.27 [0.34]	
3	Imp	Weekday 6-9am	1.36 (3.7)		3.18 (3.8)	0.24 (3.0)	
3	Imp	Weekday 9am-6pm	1.60 (3.9)		4.39 (3.7)	0.17 (4.7)	
3	Imp	Weekday 6-10pm	2.53 (4.3)		3.54 (4.6)	0.35 (6.0)	
3	Imp	Saturday 6-9am	1.73 (4.3)		6.13 (4.6)	0.42 (6.8)	
3	Imp	Saturday 9am-6pm	2.01 (4.3)		6.94 (4.4)	0.24 (7.1)	
3	Imp	Saturday 6-10pm	2.16 (4.3)		10.67 (4.5)	0.47 (7.2)	
3	Imp	Sunday	1.87 (4.2)		9.50 (4.6)	0.26 (6.8)	
3	Imp	Night	5.87 (2.9)		16.92 (2.0)	0.70 (5.4)	
3	Imp	Daytime	6.79 [8.19]		20.64 [27.76]	1.07 [1.27]	
3	Imp	Evening	4.69 [5.16]		14.21 [16.59]	0.82 [0.89]	
3	Det	Weekday 6-9am	0.25 (3.6)		1.09 (3.4)	0.03 (6.4)	
3	Det	Weekday 9am-6pm	0.30 (3.2)		0.67 (1.8)	0.03 (6.4)	
3	Det	Weekday 6-10pm	0.48 (3.2)		0.87 (3.2)	0.03 (6.7)	
3	Det	Saturday 6-9am	0.15 (3.5)		1.02 (3.9)	0.03 (6.7)	
3	Det	Saturday 9am-6pm	0.31 (3.1)		1.07 (3.2)	0.03 (7.1)	
3	Det	Saturday 6-10pm	0.40 (3.4)		1.32 (3.3)	0.03 (6.8)	
3	Det	Sunday	0.26 (3.5)		1.66 (4.1)	0.04 (7.1)	
3	Det	Night	2.20 (4.5)		10.86 (3.7)	0.11 (7.1)	
3	Det	Daytime	1.01 [1.21]		3.85 [5.10]	0.12 [0.15]	
3	Det	Evening	0.88 [0.95]		2.19 [2.61]	0.06 [0.07]	

Notes: Figures in square brackets include the Sunday value apportioned to daytime (75%) and evening (25%).

A package (part-whole) effect could be in operation in SP2, such that it is not valid to sum up the values across time periods. However, we should note that there are several instances where the values for a single time period in SP2 exceed the values for daytime or evening in SP1. It is readily apparent that the values obtained by SP2, which might be regarded as a conventional SP approach, are somewhat higher than the values of SP1. We hypothesised that this would be the case since the purpose of the exercise is transparent in SP2 and hence there is a much greater incentive to bias responses. It would have been informative to have specified a more aggregated time period within the SP2 exercise to enable direct comparison of time periods. As plane movements vary by time period and the purpose was to identify variations in values between periods the solution is not straightforward. Respondents could have been asked about more than one time period

Comparing SP1 and SP3, we find the disparities to be even larger. Indeed, in some cases, particularly for improvements, the absolute values in SP3 are alarmingly high. A package effect may again be in operation here, which will have inflated the absolute values in SP3, but in any event we would be looking to this exercise to indicate how values vary by time of day.

Table 15 allowed an examination of the extent to which SP2 and SP3 provide similar variations in values by time period. In SP3 all time periods were considered simultaneously but in SP2 each respondent considered only one time period. Given the SP2 models segmented by time period have relatively small samples, it was not possible to obtain sufficiently robust results for all time periods. Hence we have compared across time periods for which coefficients are reported for SP2. For each set of results, Table 16 presents the proportions that each value in a time period forms of the sum of the values across all relevant time periods.

It can be seen that there is a very encouraging degree of similarity between the relative valuations by time period for SP2 and SP3 for Manchester, especially for the improvements in SP3. The same can be said for Lyon and Bucharest, although with some large differences between the figures for night. It appears that as the airports have minimal night time movements subject to restrictions any deterioration is seen as unacceptable and so this time period becomes dominant when looking at a worsening in the situation.

The similarity of the SP2 and SP3 results allows us to conclude that respondents can distinguish between the aircraft annoyance of different time periods and indicates that it is reasonable to estimate values by time period without considering all time periods simultaneously.

Table 16: Variations by Time Periods on SP2 and SP3 (% of total value)

Time period	Manchester			Lyon			Bucharest		
	SP2	SP3I	SP3D	SP2	SP3I	SP3D	SP2	SP3I	SP3D
Weekday 6am-9am	-	-	-	18.2	5.8	6.2	15.6	11.2	11.1
Weekday 9am- 6pm	6.2	11.5	8.5	8.6	8.0	3.8	17.2	7.9	11.1
Weekday 6pm-10pm	15.3	18.2	13.5	23.5	6.4	5.0	15.0	16.4	11.1
Saturday 6am-9am	-	-	-	-	-	-	18.3	19.6	11.1
Saturday 9am-6pm	16.2	14.5	8.7	7.2	12.6	6.1	-	-	-
Saturday 6pm-10pm	-	-	-	13.3	19.3	7.5	-	-	-
Sunday	19.4	13.5	7.3	7.4	17.2	9.5	16.1	12.2	14.8
Night	42.9	42.3	62.0	21.8	30.7	61.9	17.8	32.7	40.8

5.1 Bias?

We have focussed on an examination of whether the values obtained depend on the form of SP exercise used. In particular we have tested whether the hypothesised variations in the incentives to bias responses affect the estimated valuations in the expected manner. Here we discuss the possible biases and their influence on the values.

The findings indeed indicate that the estimated values are somewhat higher in the naïve SP models. This may be due to the transparency of the exercise which facilitates strategically biased responses. We contend that this is consistent with other evidence in the transport literature discussed on section 2. Our recommendation therefore is that every effort should be made, where possible, to avoid making the purpose of the exercise transparent. This can be expected to be particularly important where emotive or contentious issues, such as aircraft noise are being investigated.

Whilst the values of the quality of life variables can be expected to be unbiased in relation to each other, given that all are perceived to be equally likely given the exercise did not focus on any particular quality of life attribute, there remains the possibility that the monetary values are biased if respondents perceive changes in the monetary values to have a different likelihood of occurring or if forms of protest against council tax are present. Whilst we are encouraged by what appear sensible monetary values, nonetheless the use of an additional numeraire, such as journey time, through which monetary valuations could be obtained indirectly using the best available evidence on values of time, would be a useful enhancement to any future application.

It could be argued that in fact it is not that the SP2 results are biased upwards but that the discrepancy between the two sets of results stems from the SP1 values being too low. It might be argued that budget effects are at work in SP1, given the range of attributes under consideration, whereby the value of any aspect of quality of life is higher when purchased separately than when purchased with others. Two points can be made here. Firstly, any budget constraints would not apply in the case of improvements, yet these values for SP1 are lower than for SP2. Secondly, we are not in fact asking respondents to purchase all improvements, but merely evaluate them separately, so in this regard SP1 is no different to conventional SP exercises such as SP2.

It might also be argued that some do not select the tax reductions as much as they might otherwise because it implies undesirable cuts in services. However, this pattern of responses is identical to that of someone who has high valuations. The true difference between the SP1 and SP2 values would therefore be greater than here estimated.

There does remain the issue that a package (part-whole) effect is apparent in SP2 responses, and ideally this approach should, to the extent realistic, consider variations across much longer time periods for consistency with SP1. This would then provide an upper limit to the valuation of aircraft annoyance which could be used to scale the valuations obtained by time period. However, we have observed that the values by time period in SP2 are sufficiently large that this can hardly be expected to remove the difference between the overall values produced by the two forms of SP exercise.

SP1 would appear to have advantages in terms of reducing the likelihood of the use of simplifying decision rules, as each choice is for one of a set of attributes. Any use of simplified decision rules would be obvious, as the only way to simplify the exercise would be to pick one attribute and all the associated changes and then move onto the next and so on. Any such patterns would clearly be identifiable and also clearly not rational choices and so could be excluded from the data. In the Manchester data set we found 5 instances of such behaviour when improvements were offered and 3 in the deteriorations exercise. This is less than 3% of the sample.

The values for aircraft noise in Lyon are somewhat higher than in Manchester although total movements are lower. This could be the result of genuine preference for a quieter environment and this is supported by evidence from the focus groups and the ratings of annoyance from aircraft noise relative to other issues (Bristow and Wardman 2003). It could also be a genuine response to the current controversy over additional runways, which means that people in Lyon have spent more time thinking about the issue and formulating their preferences. Or there could be an element of policy response bias as suggested by the constant term indicating a preference for the option involving fewer flights in the SP2 model reported in Table 13. These issues are probably also affecting the values from Knutsford. The Knutsford area coefficient in a disaggregation of the SP2 model is three times the size of that for any other area even after allowing for income effects. This is an area that experiences noise from the second runway at Manchester and where the issue is still a sore point.

5.2 Preferred Values

Our conclusion is that the results from the quality of life (SP1) exercise are preferred for the overall valuations of aircraft annoyance as represented by variations in aircraft movements. Two sets of values have been obtained, based on improvements and deteriorations. There are differences in the values for gains and losses. This could be explained by variations in either the marginal utility of aircraft movements and/or the marginal utility of cost. However, we cannot distinguish the impact of the two influences. This is because the improvements in aircraft movements were valued relative to corresponding improvements in tax and deteriorations in aircraft movements were valued in corresponding increases in tax but we have not examined willingness to pay for fewer aircraft movements or willingness to be compensated in the case of additional aircraft movements. In such circumstances, the most sensible approach at this stage would seem to be to use an average of the values for deteriorations and improvements.

The preferred values are set out in Table 18. The values in brackets have been adjusted using Purchasing Power Parities (IEA, 2002), relative to the French values. Once adjusted in this way, the Bucharest values appear closer to those obtained elsewhere. We might expect the Bucharest values to be lower as the airport has far fewer aircraft movements than Lyon or Manchester, levels of annoyance from aircraft are lower and there is the confounding factor of anticipated economic benefits associated with airport expansion. If values were wholly dependent on the number of plane movements we might expect the values in Manchester to be higher than Lyon, which is not the case. Clearly respondents in Lyon have expectations of quiet and a strong preference for reduced disturbance in the evenings.

Table 18: Preferred Weekly Values per Plane Passing by Per Hour (€)

	Manchester	Lyon	Bucharest
Per Daytime Hour	0.84 (0.87)	1.10	0.15 (0.58)
Per Evening Hour	0.30 (0.31)	1.26	0.05 (0.19)

6. Conclusions

6.1 Main conclusions

In this paper we have reported results from SP experiments designed to test for the presence of strategic bias in responses. Our conclusion is that a naïve SP exercise is prone to strategic bias. We therefore prefer the results from the quality of life (SP1) exercise for the overall valuations of aircraft annoyance as represented by variations in aircraft movements as these results are less open to biased responses. However, a natural limitation of the quality of life SP is that it cannot focus on detailed issues, such as the annoyance of different types of planes or at different times of day, since this defeats the whole object of trying to mask the purpose of the exercise. The way forward we believe is to use the quality of life form of SP exercise to provide 'top level' values. These values can then be decomposed according to the relativities derived from the more naïve SP approach.

6.2 Future Work

The drawback of masking the purpose of the exercise is that more variables are needed, and this leads to problems of task complexity. The challenge therefore is to be able to cover a wide range of variables yet ensure that the task is manageable. The means by which this has been tackled here has been to use an approach which effectively involves the evaluation of attribute variations one at a time rather than the conventional SP procedure of multiple trade-offs. The approach has similarities with the priority evaluator technique which has been used when there has been a need to evaluate a large number of variables, such as the many different types of rolling stock and station facility attributes (MVA, 1985; 1986) and diverse quality of life issues (Brown, 1996; Hoinville, 1971).

A limitation of the current method is that it deals only with improvements or only with deteriorations in attributes. Whilst in principle there is no reason why individuals cannot be asked to accept a deterioration in one attribute to achieve an improvement in another, along the lines of the priority evaluator approach, this builds in linear-dependency amongst the attributes and hence the undoubted information in the responses cannot be statistically modelled. Moreover, examination of interaction effects would require more attributes to vary simultaneously. Further research to overcome these problems, yet maintain the essential simplicity of the approach is required.

Although the quality of life SP has worked well, despite the conceptual difficulties respondents had with it, improvements to the presentation could make it easier for respondents to understand, for example, by removing those attribute levels that had already been stated as preferred, and by making it more difficult or impossible to record "wrong" choices. This would most certainly require some form of computerised presentation.

Acknowledgements

This work formed part of the 5A (Attitudes to Aircraft Annoyance Around Airports) study funded by the EUROCONTROL Experimental Centre. The authors would like to thank Ted Elliff, Peter Hullah and Elisabeth Plachinski of EUROCONTROL for their support. We would also like to thank Paul Murphy and Christine Heaver of FaberMaunsell for their contribution to this study. The content of the paper and any opinions expressed are the sole responsibility of the authors.

References

Adamowicz, W. (2000) Environmental Valuation Case Studies. In Louviere, J.J., Hensher, D.A and Swait, J.D., Stated Choice Methods: Analysis and Application. Cambridge University Press.

Arentze, T., Borgers, A., Timmermans, H. and DelMistro, R. (2003) Transport Stated Choice Responses: Effects of Task Complexity, Presentation Format and Literacy. Transportation Research 39E (3) 229-244.

Arsenio E., Bristow A.L and Wardman M. (2002) Marginal Values of Traffic Noise Externalities from Stated Preference Methods. Paper to European Transport Conference 9th - 11th September 2002.

Bateman I.J., Carson R.T., Day B., Hanemann M., Hanley N., Hett T., Jones-Lee M., Loomes G., Mourato S., Ozdemiroglu E., Pearce D.W., Sugden R. and Swanson J. (2002) Economic Valuation with Stated Preference Techniques: A Manual. Edward Elgar.

Bates, J. (1998). Reflections on Stated Preference: Theory and Practice, in Ortuzar, J. D. Hensher and S. Jara-Diaz (editors), Travel Behavior Research: Updating the state of play, chapter 6, 89-103, Pergamon, UK.

Bohm, P. (1971) An Approach to the Problem of Estimating Demand for Public Goods. Swedish Journal of Economics, 73 (1).

Bristow A.L., Wardman M. et al (2003) Attitudes Towards and Values of Aircraft Annoyance and Noise Nuisance. Attitudes to Aircraft Annoyance Around Airports (5A) Survey Report. EEC/SEE/2003/002. EUROCONTROL Experimental Centre, France. http://www.eurocontrol.fr/ba_env/Documents/EEC_SEE_2003_002.pdf

Bristow A.L., Wardman M., Heaver C. and Plachinski E. (2004) Aircraft Noise: annoyance and values a comparative study of airports in three countries. World Conference on Transport Research, Istanbul, July 2004

British Standard Institute (2003) Acoustics – Assessment of noise annoyance by means of social and socio-acoustic surveys. Draft for Development

Brooker P, Critchley J.B., Monkman D.J. and Richmond C. (1985) United Kingdom Aircraft Noise Index Study: Main Report. DR Report 8402, Civil Aviation Authority.

Brown, N (1996) Quality of Life Issues within an Integrated Transport Strategy Framework. PhD Thesis, University of Leeds.

Chapman, R.G. and Staelin, R. (1982) Exploiting Rank-Ordered Choice Set Data Within the Stochastic Utility Model. Journal of Marketing Research 19 (3) 288-301.

Cirillo, C., Daly, A. and Lindveld, K (2000) Eliminating bias due to the repeat measurements problem in SP data. In Stated Preference Modelling Techniques, edited by Ortúzar, J. de D., PTRC, London.

Diamond I., Stephenson R., Sheppard Z., Smith A., Hayward S., Heatherley S., Raw G., and Stansfield S. (2000) Perceptions of Aircraft Noise, Sleep and Health. Report to Civil Aviation Authority.

Fields J.M. (1998) Reactions to environmental noise in an ambient noise context in residential areas Journal of the Acoustical Society of America 104(4) 2245-2261

Fields J.M., De Jong R.G., Gjestland T., Flindell I.H., Job R.F.S., Kurra S., Lercher P., Vallet M., Yano T., Guski R., Felscher-Suhr U. and Schumer R. (2001) Standardized General-Purpose Noise Reaction Questions for Community Noise Surveys: Research and Recommendation, Journal of Sound and Vibration 242(4) 641-679

Heaver C. (2002) Attitudes to Aircraft Annoyance Around Airports (5A) Focus Group Report. EEC/ENV/2002/009
http://www.eurocontrol.fr/ba_env/Documents/EEC_ENV_2002_009.pdf

Hensher D. (2001) Measurement of the Valuation of Travel Time Savings Journal of Transport Economics and Policy 35(1) 71-98

Hoinville, G. (1971) Evaluating Community Preferences. Environment and Planning 3 33-50.

International Energy Agency (2002) Beyond Kyoto: Energy Dynamics and Climate Stabilisation. Table 7 accessed 30/4/03.
http://www.iea.org/public/studies/beyond/table7.pdf

Johnson, E.J. and Meyer, R.J. (1984) Compensatory Choice Models of Noncompensatory Processes: The Effect of Varying Context. Journal of Consumer Research 11 45-57.

Louviere, J.J. (2001) Choice Experiments: An Overview of Concepts and Issues. In Bennett, J. and Blamey, R. (Eds) The Choice Modelling Approach to Environmental Valuation. Edward Elgar, Cheltenham.

Malhotra, N.K. (1982) Structural Reliability and Stability of Nonmetric Conjoint Analysis. Journal of Marketing Research 19 199-207.

Miedema H.M.E. and Vos H. (1999) Demographic and attitudinal factors that modify annoyance from transportation noise. Journal of the Acoustical Society of America, 105(6) 3336-3344.

MVA (1986) Evaluation of InterCity Rolling Stock Improvements. Prepared for British Railways Board.

MVA (1985) Station Modernisation Priorities and Payoffs. Prepared for British Railways Board

Navrud S. (2002) The State-of-the-Art on Economic Valuation of Noise. Final Report to European Commission DG Environment.

Nelson J.P. (2004) Meta-Analysis of Airport Noise and Hedonic Property Values: Problems and Prospects. Journal of Transport Economics and Policy 38(1) 1-28.

Schipper, Y., Nijkamp, P. and Rietveld, P. (1998). Why Do Aircraft Noise Values Estimates Differ? A Meta-Analysis. Journal of Air Transport Management 4, 117-124.

Thune-Larsen H. (1995) Flystøyavgifter basert på betalingsvillighet, TOI report 289/1995. (English language summary report: Charges on Air Traffic Noise by Means of Conjoint Analysis).

Wardman, M. (1998) The Value of Travel Time: A Review of British Evidence. Journal of Transport Economics and Policy, 32(3), 285-315.

Wardman M. (2001) Public Transport values of Time. Working Paper 564, Institute for Transport Studies, University of Leeds.

Wardman M. (2003) Reliability of Values Obtained from Stated Preference Methods. Mimeo.

Wardman M. and Bristow A.L (2004) Noise and Air Quality Valuations: Evidence from stated preference residential choice models, Transportation Research D, 9(1) 1-27.

Wardman, M. and Shires, J. (2001) Comparison of Within Mode Revealed Preference and Stated Preference Choice Models. Paper presented at AET European Transport Conference.

Wardman, M. and Whelan, G. (2001) Valuation of Improved Railway Rolling Stock: A Review of the Literature and New Evidence. Transport Reviews 21 (4) 415-448.

Widlert, S. (1998) Stated Preference Studies: The Design Affects the Results. in Ortuzar, J. D. Hensher and S. Jara-Diaz (editors), Travel Behavior Research: Updating the state of play, chapter 7, 105-123, Pergamon, UK.