Published paper
Wardman, M.; Bristow, A. (2006) Valuation of Aircraft Noise: Experiments in Stated Preference. Journal of Environmental and Resource Economics, 39(4), pp.459-480.

Valuations of Aircraft Noise:

Experiments in Stated Preference

Mark Wardman[1] and Abigail Bristow[2]

[1]Institute for Transport Studies, University of Leeds, Leeds LS2 9JT

[2]Transport Studies Group, Department of Civil and Building Engineering, Loughborough University, Loughborough LE11 3TU

mwardman@its.leeds.ac.uk, a.l.bristow@lboro.ac.uk

Abstract

The aim of this paper is to report new evidence relating to residents' valuations of aircraft noise in three countries with an emphasis on a comparison of the valuations obtained using two contrasting approaches. One might be regarded as a standard stated choice approach offering pairwise comparisons of two alternatives characterised by a limited number of attributes. The other choice format adopted is innovative in drawing inspiration from the priority evaluator approach to embed aircraft movements alongside a wide range of other local factors that impact on residents' quality of life. The paper addresses the differences in the results of the two approaches and explores the possible explanations for these variations. Although not conclusive, there is a suspicion that strategic bias may have influenced the results and we urge further research regarding incentives to such bias.

Acknowledgements

This work was undertaken as 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 for their support. We would also like to thank the two referees for their insightful comments. The content of the paper and any opinions expressed are the sole responsibility of the authors.

Keywords

Aircraft Noise, Environmental Valuation, Noise Valuation, Stated Preference, Choice Format.

This paper has not been submitted elsewhere in identical or similar form, nor will it be during the first three months after its submission to the publisher.

1. INTRODUCTION

The research reported in this paper was exploratory, involving the novel use of different Stated Preference (SP) methods to estimate valuations of annoyance caused by aircraft movements. Additionally there is an international dimension since surveys were conducted amongst residents around Manchester, Lyon and Bucharest airports. The research builds upon our experiences of valuing road traffic noise in Edinburgh (Wardman and Bristow, 2004) and Lisbon (Arsenio et al., 2006) and provides fresh empirical evidence where very little exists. The only SP studies of aircraft noise of which we are aware are those of Thune-Larsen (1995), Carlsson et al. (2004) and ongoing research funded by the UK Department for Transport (MVA, 2004) with the related contingent valuation method (CVM) finding only limited application (Navrud, 2002).

Two contrasting choice formats were used. One might be regarded to be a standard Stated Choice (SC) approach, and offered pairwise comparisons of two alternatives characterised by a limited number of attributes. In this context these were local tax as the numeraire and movements of three different types of aircraft. The other choice format adopted has not, as far as we are aware, been reported in the environmental economics literature. This innovative experiment, termed priority ranking (PR), draws upon the priority evaluator approach (Hoinville 1971) and embeds aircraft movements alongside a wide range of other local factors that impact on residents' quality of life.

There were a number of reasons for employing these two quite different SP experiments. Firstly, it is common practice to trial different methodologies in studies

which are, as this was, exploratory in nature. Indeed, we have reported elsewhere comparison of the values obtained from the SC approach discussed here, which deals with trade-offs between tax and aircraft within a specified time period, with those obtained from a (third) SP experiment of the ranking format, which involved trading-off aircraft movements across time periods (Bristow and Wardman, 2006a). Secondly, the two SP exercises covered in this paper serve different purposes whilst recovering valuations that can be compared. The PR method includes aircraft noise alongside other factors that impact on the quality of life. This may serve to reduce incentives to response bias by not placing undue emphasis on aircraft noise. Such an approach is felt to be useful in establishing an overall valuation for aircraft noise. Nonetheless, there are risks associated with this largely untried method. On the other hand, the SC method was felt to be lower risk and it lends itself to a more detailed examination of variation in values between time periods and aircraft types that cannot practically be explored within an experiment that includes a wide range of quality of life issues. Finally, there are reasons why valuations might differ according to the choice format, as we discuss in section 5, and we wished to explore these.

The aim of this paper is therefore to report new evidence relating to residents' valuations of aircraft noise in three countries but with an emphasis on comparing the valuations obtained from contrasting methodologies. Section 2 outlines the survey design and the data collection procedure. Sections 3 and 4 respectively focus on the empirical results for the SC and PR methods in isolation from each other and the contribution that these provide to the existing body of evidence relating to aircraft noise valuations. These are followed by a comparison of the findings from the SC and PR methods in section 5. Concluding remarks and recommendations for further research are provided in section 6.

2. SURVEY DESIGN AND IMPLEMENTATION

2.1 The 'Standard' Stated Choice (SC) Exercise

This experiment took the form of a binary choice context containing only a few attributes. At least in the European context, this can be regarded to be a standard SC approach. Two abstract alternatives (A and B) offered trade-offs between local tax and aircraft movements. The number of aircraft was disaggregated into three types: 747 jumbo jets and other large aircraft; two engined jets, such as Airbus and 737's; and turbo-prop aircraft.

Respondents were asked to consider the variations in aircraft movements within a specific time period as annoyance caused by aircraft noise will in part be determined by exposure to it and activities being undertaken when the noise is experienced. The time periods chosen to reflect variations in aircraft movements were: weekdays 6am-9am, 9am-6pm and 6pm-10pm; Saturday 6am-9am, 9am-6pm and 6pm-10pm; Sunday 9am-6pm and night. Although the time periods vary across individuals, they do not differ across the SC scenarios evaluated by any one individual.

An example of a choice which respondents were asked to make is given in Table 1. Option B always involved more or the same number of planes per hour than Option A but it was always cheaper. The number of planes passing by was specified both as the number per hour and as the interval since focus groups conducted for this study at locations around the three airports (Heaver 2002) revealed that some preferred the former representation and others preferred the latter.

Table 1 about here

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 choices produced by the design was sixteen, but any individual was presented with only 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, notably increasing the range of cost variations in order to increase the precision with which its parameters could be expected to be estimated.

In this exercise, the purpose of the study would be quite transparent, a feature of many conventional SC exercises and one that is fuelled in part by a tendency at least in Europe to simpler exercises containing as few as three attributes. In addition, this exercise was preceded by questions concerning noise from various sources including aircraft and the annoyance caused.

2.2 Priority Ranking (PR) 'Quality of Life' SP Exercise

We wished to establish the significance of aircraft annoyance within the much broader dimension of quality of life. One conclusion drawn from the focus groups was that we could frame an SP exercise within a broader quality of life dimension which would avoid placing undue emphasis upon aircraft noise, since respondents seemed to have no suspicion that this was the primary concern of the survey. It would also provide a broader context for the interpretation of the results and an opportunity to develop a novel application of a little used form of SP approach.

Aircraft noise was therefore considered alongside nine other quality of life variables and local tax. The focus groups assisted in identifying the most significant local quality of life issues to consider. These were local crime levels, local school quality, the level of area wide traffic congestion, street cleanliness, traffic noise experienced

at home, neighbourhood air quality, the general condition of local roads and pavements, local recreation facilities and local amenities. The common theme amongst the attributes was that to varying degrees they could be regarded to be under the influence of local authorities.

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 11 attributes would be considerable. Serious doubt would have to be placed on whether respondents could provide reliable answers to such an exercise. There is evidence to indicate that task complexity can influence valuations, largely through the use of simplifying but inappropriate choice rules or ignoring attributes (Timmermans, 1993; Widlert, 1998; Arentze et al., 2003; Caussade et al., 2005).

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 has been to develop an approach which involves the evaluation of attribute variations one at a time rather than the conventional procedure of multiple trade-offs. Our view is that if offered a whole series of improvements (or deteriorations) to specific attributes, respondents can more readily state which (one-dimensional) attribute variation they would most like to achieve than they can weigh up the net benefit of (multi-dimensional) differences in a whole range of attributes between two alternatives.

The approach adopted is very much along the lines of 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).

Bearing in mind the number of evaluations that respondents would have to make, five levels were chosen for each variable. The exceptions to this were tax, which had seven levels in order to introduce more variation into this key variable and to allow for uncertainty as to households' valuations, and local facilities, which were either present or not. The seven levels varied across individuals with maximum increases and reductions on the current level of tax of £10 in Manchester, €13 in Lyon and 50000 Leu in Bucharest[1]. The levels for Lyon and Manchester are roughly equivalent while those for Bucharest are much lower reflecting the lower income levels.

Respondents were offered aircraft annoyance represented either as the number of aircraft movements or the categorical levels of extremely noisy, very noisy, moderately noisy, slightly noisy and not at all noisy. Whilst the former is more useful in terms of quantifying the value of the externalities associated with aircraft operation, the latter categorical scale was used in order to allow direct comparison with valuations of the same categories of road traffic noise. It also readily allows comparisons to be made across the three locations.

For those for whom aircraft noise was represented in terms of the number of movements, the figures offered varied around our best estimate of the current situation as depicted in column III of Table 2 and distinguished between daytime (6am to 6pm) and evening (6pm to 10pm) flights. The variations in total aircraft movements and council tax were specified to be broadly similar to those used in the SC exercise.

Table 2 about here

[1] At the time, the exchange rates were €1 equals £0.639 and 1 Leu equals £0.00001973.

An example of what was offered to respondents is given in Table 3. The starting point is to identify the respondent's current situation. Where possible this was predefined and shaded. Such instances were where a level related to the current situation, as with area wide road traffic congestion, council tax and the number of aircraft movements, and where respondents could not reasonably be expected to know the current level, as with the number of local burglaries and local school pass rates. Respondents were asked to identify their current position for the remaining attributes.

Table 3 about here

Having identified the current situation, the respondent was then asked to consider the improvements, which are all the attribute levels to the right of the current, and to state which improvement would be most preferred. They were then asked to disregard this improvement, treating it as if it were no longer available, and asked to state which was now the preferred improvement. This process continued until all the possible improvements had been ranked in order of preference. Having completed the ranking of improvements, the respondent then proceeded to evaluate the deteriorations, commencing by indicating the worst and proceeding in an entirely analogous fashion to improvements but working towards that deterioration which was regarded to be 'least bad'.

2.3 Survey Locations

Surveys were conducted at three European airports since one of the aims of the study was to evaluate cultural and socio-economic variations in attitudes towards and valuations of aircraft annoyance.

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 set in a much more rural environment. The populations around these two airports are comparable in terms of income levels. Bucharest Airport was selected as an airport with a much smaller number of aircraft movements and significantly lower income levels amongst the local population. Recent developments at these airports are of interest. A second runway opened at Manchester in 2001 and, whilst the controversy has now largely died down, 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. Bucharest has seen no significant changes since 1997 when the new international terminal opened. However, there might be much more tolerance of increased aircraft activity in a low income country such as Romania if economic development benefits are perceived to be associated with it.

2.4 Data Collection

The surveys were conducted in late 2002 at locations around each of the airports which were selected to give a range of exposure to aircraft noise and socio-demographic characteristics. Respondents were recruited from these specified locations in accordance with quotas set for age, gender and employment to ensure a reasonable spread (full details may be found in Bristow et al 2003). A small payment was made to participants. The sample sizes achieved were 200 at Manchester, 210 at Lyon and 237 at Bucharest.

The survey was paper based and conducted in hall test conditions where staff could assist respondents as necessary and explain each part of the survey before it started. Survey materials were distributed in sections, with attitudes to the quality of

life variables and the PR exercise conducted first to prevent respondents looking ahead, thereby concealing the main purpose of the survey. Questions specific to aircraft noise and annoyance and the SC exercise were subsequently presented.

3. STATED CHOICE (SC) RESULTS

Modelling Issues

The results for the SC model are reported in Table 4. The ALOGIT package was used (Hague Consulting Group, 2000) and its jack-knife procedure accounted for individuals' repeat observations (Cirillo et al., 2000). The aircraft movements relate to the number of 'planes going by' in each hour of the time period in question. All values are expressed as € per week.

Table 4 about here

Separate coefficients are estimated for the number of movements in each of the eight time periods. In these models, the coefficients are not estimated sufficiently precisely to support disaggregation by time period and plane type simultaneously. However, any distortions as a result of constraining all aircraft types to have the same valuation can be expected to be minimal since two-engined jets dominate the specified movements.

Some aircraft movement coefficient estimates were far from significant and were dropped from the models. The t ratios associated with these coefficients are listed and indicate that they were generally estimated very imprecisely. A contributory factor is no doubt the limited number of observations for some time periods, and with

hindsight it might have been preferable to concentrate on fewer periods given the intended sample sizes. The ρ^2 goodness of fit measures are low, particularly for Bucharest where the respondents struggled with the SP exercises.

Gains and Losses

We allowed the tax coefficient to differ between increases and reductions in tax. The coefficients for reductions and their associated t ratios were -0.027 (1.2) for Manchester, -0.035 (1.1) for Lyon and 0.648 (1.5) for Bucharest. These insignificant coefficients were not retained and the tax coefficients in the reported models relate only to tax increases.

It is not surprising that tax increases have a larger impact than tax reductions. This finding could stem from loss aversion, which is generally regarded to be intuitive by psychologists and is not inconsistent with conventional economic theory in the form of diminishing marginal utility. Protest response could also have a bearing. However, these points do not explain why tax reductions have no statistically discernible effect at all upon choice. We feel that a contributory factor here is one of credibility. Respondents might not have believed that taxes would ever be reduced in practice and have therefore ignored tax reductions in their choices. With hindsight, there would have been considerable value in supplementary questions relating to attitudes towards tax increases and reductions. Nonetheless, we note that tax reductions did have benefit in the PR exercise and therefore that sign effects could be conditional upon the choice format used.

Given the limited number of observations per time period, there is little sense in attempting to distinguish between increases and reductions in aircraft movements for each time period. Instead, we specified a single incremental term covering all time

periods to detect whether increased aircraft movements were valued differently to reductions. The incremental terms for increased aircraft and their associated t ratios were 0.0004 (0.1) for Manchester, -0.0098 (1.1) for Lyon and 0.0217 (1.2) for Bucharest. Thus none of these incremental terms were retained. This effect is not entirely surprising in the context of previous work on noise valuation in Lisbon and Edinburgh that also failed to identify a significant sign effect (Arsenio et al, 2006; Wardman and Bristow 2004).

We entered constant terms into the models to determine whether there was any residual preference for one option over the other after accounting for differences between them in tax and aircraft movements.

For Manchester residents the constant was far from significant and hence removed. In the case of Lyon, there is an appreciable preference for the quieter option equivalent to around €20. We interpret this as response bias given concerns raised in the focus groups and the existence of opposition groups regarding the proposed construction of two additional runways at Lyon airport. If the constant is not included then the cost coefficient becomes wrong sign, albeit insignificantly so, and this is not plausible. A likelihood ratio test indicates that the model with a constant is statistically superior. The Bucharest SP model exhibited the reverse preferences to the Lyon model but the constant was highly correlated with the tax variable and hence was not retained.

Time of Day Effects

The results provide an important insight into time of day effects which might be important in appraisal but which cannot be gleaned by the hedonic pricing method.

The valuations reported in Table 4 relate to the entire period in question. Variations by time of day are more apparent if we standardise for the duration of the time period and such valuations are reported in Table 5. This indicates that the valuations tend to be higher in the evenings when more people are at home and exposed to noise than in the day time. However, there is also a tendency for valuations to be relatively high for shorter periods where the standardisation is across fewer movements. In some cases these are sensitive time periods where values might be expected to be fairly high, such as Saturday evenings in Lyon. This is a less likely explanation for high values for Saturday mornings in Bucharest and weekday mornings in Lyon and Bucharest. We suspect that there might be a framing effect at work here.

Nonetheless, it does seem that useful insights can be obtained, although it is clear that larger sample sizes per period are required than were here obtained, and this is encouraging given that respondents did not trade-off movements in different periods but instead the valuations for different periods stem from different individuals each evaluating aircraft noise in a single period. We explore this more fully elsewhere (Bristow and Wardman, 2006a) where we find a reasonable degree of similarity between the SC time specific values reported here and the values obtained from a third SP exercise not reported here which did require each respondent to trade-off movements across time periods. Our main concerns surround the responses of Bucharest residents who have very similar coefficients for each time period, regardless of the length of the period or the likely exposure to the noise. Our view is that these respondents were unfamiliar with any type of opinion survey and have not replied as carefully as others and that the desire to reap economic benefits associated with airport expansion may have had a major bearing on responses.

TABLE 5 ABOUT HERE

Country Variations

Inspecting the valuations across countries, rather than across time periods or data sets, the pattern is largely as expected. Manchester and Lyon have similar income levels and, taken together, the values are broadly similar. The Lyon values are a little higher which is expected as aircraft noise was here rated as the major source of dissatisfaction compared with other quality of life indicators. In contrast, the values are much lower in Bucharest, reflecting the much lower income levels in Romania, the confounding effect of a preference for airport expansion on the grounds of economic development, the lower number of aircraft movements at this airport and the lower levels of annoyance.

4. PRIORITY RANKING (PR) RESULTS

Modelling Issues

Respondents have effectively ranked a number of alternatives in order of preference when undertaking the PR exercise. Each alternative contains all attributes at the current level except the attribute that has improved or deteriorated and has been evaluated. Thus if an individual has evaluated 20 improvements, this is the same as ranking 20 alternatives each of which has only one attribute varying from the base situation. The farther a respondent's current situation is to the left (right) in Table 3 then the more improvements (deteriorations) they will evaluate.

These ranked alternatives are analysed using what is termed the exploded logit model (Chapman and Staelin, 1982), and the ALOGIT package and its jack-knife procedure has again been used. The first ranked of n alternatives is specified to be

preferred to the other n-1 alternatives within a multinomial logit framework. This first ranked alternative is then made unavailable, and the alternative ranked second best is specified to be preferred to the remaining n-2 alternatives, and so on until the ranking is exhausted.

Whilst the multinomial logit model possesses the potentially restrictive independence of irrelevant alternatives (IIA) property, this is not a particular cause for concern. The IIA property operates at the level of demand forecasts, forcing cross-elasticities to be equal, and is less of an issue for valuation where we are interested in relative coefficients. Horowitz (1980) concluded on the basis of simulation tests that the IIA feature was not the problem for valuation that it is for forecasting.[2] Nonetheless, more sophisticated modelling of the data might identify whether, at least in some circumstances, it would yield different or more precise valuations.

Separate models have been estimated for each airport for both the improvements and the deteriorations PR exercises. The key indicators of interest are reported in Tables 6 and 7. Results for the other quality of life indicators may be found in Wardman et al (2003). As with the SC exercise, the aircraft movements relate to the number of 'planes going by' in each hour of the daytime and evening periods and all values are expressed as € per week.

Both the improvement and deterioration models achieve goodness of fit measures (ρ^2) in line with those typically achieved in more conventional SP travel choice models. Most respondents had a current situation towards the left in Table 3 and so

[2] We ran some models as a series of binary choices, where the IIA problem does not apply, and the valuations obtained were not materially different

the improvement models have more coefficients than the deterioration models. Different individuals do have different base situations and thus there is no unique base category from which the incremental effects are interpreted. However, this issue does not arise for the aircraft movement and tax variables of interest here and, moreover, the incremental effects for other quality of life variables were not materially altered when we amended the base category.

PR Improvements

As far as improvements are concerned, variations in daytime aircraft movements have a statistically significant effect in all three locations whilst evening aircraft movements have a significant effect in both Manchester and Lyon. The daytime values in Manchester and Lyon are similar, in line with their similar income levels, whilst the higher sensitivity of Lyon residents to evening aircraft noise was also apparent in the attitudinal responses and stems in part from a preference for evenings spent outdoors.

Table 6 about here

Standardising with regard to aircraft movements within the period, we find that the daytime and evening improvements are valued at 1.48 and 1.89 cents per aircraft in Manchester and 1.26 and 5.68 cents in Lyon. It seems sensible that evening valuations are higher given the greater exposure to aircraft noise at home, particularly so for the Lyon residents.

PR Deteriorations

The results are not as satisfactory for deteriorations, since it was not possible to discern a statistically significant effect for evening aircraft movements in either Manchester or Bucharest. It is not clear why this should be so, but we do not feel that it represents a genuine zero valuation.

There is evidence that the Lyon residents are more averse to deteriorations than the Manchester sample. This is not simply a protest against airport expansion at Lyon since it was apparent amongst the values of the other quality of life variables.

After standardising for the number of movements per period, the Lyon deteriorations evening value of 4.75 cents per movement is, as in the improvements model, somewhat higher than the daytime valuation which is 1.71 cents. Again this seems entirely reasonable.

Table 7 about here

Removing Illogical PR Responses

We subsequently removed those who supplied any responses which did not place an improvement or deterioration in logical order, such as preferring, say, a £2 per week tax saving to a £5 per week tax saving. For deteriorations, the reductions in the Manchester, Lyon and Bucharest samples were 33%, 26% and 65% respectively. The corresponding figures for improvements were 45%, 38% and 72%.

As would be expected, this procedure improved the model fit. However, there was no change in the models' inabilities to discern statistically significant effects in some evening periods. There was a tendency for the valuations in the improvements to be lower. This is because the illogicality was quite strongly associated with the tax

reductions. This could be related to the issue of the credibility of tax reductions which was clearly manifest in the SC models. In contrast, the removal of the illogical responses has little effect on the monetary valuations in the deteriorations models.

What is rather alarming is the extent to which illogical answers have been given. This conflicts with our expectation that the PR is a relatively straightforward exercise, and indeed some respondents were observed to have difficulties with it. Others have also clearly not followed instructions, and have moved out from the current situation to the larger changes, the reverse of what was required. Moreover, the large number of improvements and deteriorations to be evaluated increases the chances that mistakes are made. Whilst an improved format involving computer presentation would no doubt help matters, a greater insight into the causes of these problems would require debriefing of respondents when such illogicality was observed to have occurred. Nonetheless, the findings are consistent with results presented elsewhere in that the Bucharest respondents had the greatest difficulties in responding to the valuation questions.

Comparing PR Improvements and Deteriorations

We can compare valuations obtained from the improvements and deteriorations models, and this is best done by reference to Table 8 which contains all the main valuations for comparison purposes. For Lyon, the improvements and deteriorations valuations for both daytime and evening movements are broadly similar, and at the overall value they are very similar. Indeed none of the improvement and deterioration valuations are significantly different for Lyon. In the case of Manchester, the improvement based valuations tend to be more highly valued than deteriorations, with notably a zero valuation for evening deteriorations. However, we feel the latter is more a statistical aberration or some problem with responses rather than a true

reflection of these valuations. Where a significant estimate was obtained, the valuations for improvements and deteriorations were not significantly different. The value of increased aircraft movements is very much lower than reductions in Bucharest. This may reflect a 'halo' effect of economic development benefits associated with airport expansion, and also a heightened aversion to tax increases where incomes are low. Economic theory may contribute to an understanding of the relationships that might exist between the different valuations. We return to this in section 5 when undertaking a broader comparison of the valuations we have obtained.

Means of Presentation

We can also compare the values from the two means of presenting aircraft noise. This is most sensibly done for the improvements model since there is no obvious way of translating extremely noisy or very noisy into an equivalent number of aircraft movements.

In the improvements model of Table 6, the base category for Manchester and Lyon is mainly composed of the very noisy category whilst for Bucharest it is moderately noisy. These correspond broadly with the central values in Table 3 since these base categories reflect the current situation. The not at all noisy category will be represented by the removal of all movements, although admittedly smaller changes than this might suffice.

In Manchester, the movement to not at all noisy is valued at €24.54. Removing the current 15 movements per hour in the daytime and evening would be respectively valued at €18.60 and €7.95. The total value based on movements of €26.55 therefore compares favourably with the equivalent categorical valuation. For Lyon,

the categorical value is €24.35 whilst removing 10 daytime and 12 evening flights per hour is valued at €10.60 and €19.08 to give a total value of €29.68. The Bucharest value of not at all noisy is €0.73 compared to €1.79 based on removing 3 daytime and 2 evening movements per hour which are respectively valued at €1.47 and €0.32. The latter is based on an evening value which is pro-rate the daytime value given that a statistically significant evening value could not be estimated.

Bearing in mind that the movement valuations might be too high, because it might not be necessary to remove all movements in order to achieve a not at all noisy level, the valuations obtained from two different means of presenting aircraft noise exhibit an encouraging degree of similarity.

5. COMPARISON OF SC AND PR VALUES OF AIRCRAFT MOVEMENTS

Comparison of Values

Whilst we have so far discussed the valuations of aircraft movement obtained separately from the SC and PR exercises, a key issue that must be addressed here is the extent to which estimated valuations differ between the two choice formats used and the possible reasons for any differences.

Given that the values of the SC models are derived from tax increases and there are no differences between the utility effects associated with improvements and deteriorations in aircraft movements, the most appropriate comparison of the SC results is with the valuations from the PR deteriorations model. This comparison is reported in Table 9 along with t statistics for the differences in the estimated valuations. The latter is specified as:

$$t = \frac{v_{PR} - v_{SC}}{\sqrt{Var(v_{PR}) + Var(v_{SC}) - 2Cov(v_{PR}v_{SC})}}$$

where v_{PR} and v_{SC} are the two valuations. Given the SC and PR values are estimated to the same sample of individuals, a positive covariance can be expected. However, we have no information on what the covariance would be. Hence our test statistic will understate the significance of the differences between valuations.

The SC valuations are higher for all airports and time periods. In two out of the four cases where a statistically significant valuation could be obtained in the PR model the difference in the valuations is significant. In the remaining two cases the difference is not far from significant, and it should be borne in mind that the t statistic will be an under-estimate given the absence of the expected positive covariance term.

The comparison is clouded by the inability to obtain significant valuations in some periods in both the PR and SC models. For example, the SC model would have provided somewhat higher valuations for daytime and evening periods if it had proved possible to estimate significant valuations for all 7 daytime and evening periods. In addition, the PR values here presumably cover Sunday, since no distinction was made by day, but the Sunday valuations of the SC model have not been included in these comparisons. It is therefore noteworthy that the SC valuations for Sunday for Manchester and Bucharest exceed the weekly daytime valuations of the PR models. Clearly if Sunday were included in the SC models the values would be even higher.

We must therefore conclude that there is strong evidence to indicate that the valuations delivered by the SC method exceed those of the PR method by an appreciable amount. We now turn to a consideration of the possible causes of this

and speculate as to which set of results we expect to provide the closer account of individuals' actual preferences.

The divergence is unlikely to be the result of different levels of tax or aircraft movements in the two exercises since we selected these to be broadly similar in each. The following sections examine issues that may help in explaining the differences.

Incentives to Strategic Bias

It has long been recognised that some individuals will not reveal their true preferences when there is a benefit to be gained from not doing so. Samuelson (1954, p388) stated that, "now it is in the selfish interest of each person to give *false* signals, to pretend to have less interest in a given collective consumption activity than he really has" whilst Bohm (1971, p.94) comments that, "potential consumers of a proposed output of a public good have stated preferences which can only be expected to *over*estimate their true valuations. The simple reason is of course that the consequences as to their payments (eg a tax increase) have been left out of the process." These statements illustrate the classic free-rider problem and the reverse incentive to overstate values where payment is not expected.

Adamowicz et al. (1999, p.467) stated that "Strategic behaviour should be minimal in SP tasks since the choices are made from descriptions of attributes and it will not be clear which choice will over- or underrepresent a valuation". However, this view is not entirely supported by empirical evidence, with unreasonably large willingness to pay valuations obtained where the purpose of the study is transparent and/or contentious and the likelihood of paying for the improvement is small (Wardman and Whelan, 2001; Wardman and Shires, 2003).

Carson et al (2000) suggest that in order for a valuation experiment to be potentially incentive compatible, the respondent should view them as consequential. In this case, although the true purpose of the PR experiment is hidden from the respondent, this does not make it any less consequential than the SC experiment. In each case the respondent is asked to trade changes in local taxation against improvements or deteriorations in factors that might be logically assumed to be within the control of local government. The two experiments are equally consequential.

If all real-world outcomes are perceived to be equally likely, there is generally no incentive not to reveal true preference in response to hypothetical questioning. An exception is to counter the preferences of others, although the net effect of this is indeterminate. However, all outcomes are not necessarily perceived as equally likely. We would expect a variable to attract a greater amount of attention in terms of strategic behaviour the more likely it is to change in practice.

If respondents perceive that the purpose of the survey is related to local tax, and there is little chance of influencing aircraft movements, then there will be an incentive to favour tax reductions and to be averse to tax increases to a greater extent than is consistent with actual preferences. The implied values of aircraft movements will therefore be too low.

If, on the other hand, the purpose of the survey is perceived to be related to the annoyance caused by air traffic, but that there is little chance of influencing local taxes, there is an incentive to bias responses in order to overstate the importance of aircraft.

We noted in the focus groups some protest against airport expansion. This was particularly noticeable amongst Lyon residents, which worked its way through in the SC model in terms of a constant term in favour of the quieter option. Whilst it is possible that the responses to the PR exercise could be affected by protest response, it seems reasonable to claim that the impact would be greater where the purpose of the study was clear. The ratio of the SC and PR valuations from Table 9 for Lyon and Manchester lie between 1.5 and 3.0. The Bucharest ratios are extremely high, a function of the very low values in the PR experiment.

It is reasonable to conclude that the values of the quality of life variables in relation to each other are unbiased as a result of the masking of the purpose of the exercise and given that respondents perceive that each of these attributes is equally likely to be improved or to worsen. However, the absolute money values even in PR could be biased. Council tax can be a contentious issue and attract protest responses, and it is clearly under local authority control, and thus tax changes might attract strategic bias, resulting in too low monetary values. On the other hand, if respondents perceive the purpose of the exercise to value collectively a wide range of quality of life attributes, there will be an incentive to inflate the values of them.

It is our opinion, based on evidence from the transport sector and supported by our interpretation of the constant in the Lyon SC model, that the SC is more susceptible to strategic response because of the transparency of purpose and (in this case) the relative contentiousness of the subject. This seems to be the most likely explanation for the higher values estimated by the SC model.

Economic Theory

This study is unusual in that the two experiments used provide estimates of values of noise based on the four Hicksian measures of consumer's surplus. The SC experiment is based on compensating loss (CL) and compensating gain (CG), which are also the norm for CVM experiments. The PR experiment provides estimates based on equivalent gain (EG) and equivalent loss (EL). CG and EL would be expected to be similar and lower than CL and EG due to income and substitution effects.

In the SC the cost coefficient for the willingness to accept (CL) was insignificantly different from zero. The cost coefficient on willingness to pay (CG) was significant.

Within the PR we have six potential values for aircraft movements in the evening and day time for the three cities. Values for evening could not be estimated for Bucharest or Manchester, leaving four values. EG is higher than EL in three cases and lower in one (Lyon day time values). The EG and EL values are not significantly different from each other, with the exception of the Bucharest values where the extremely low value for deteriorations is just significantly different from the value for improvements. The ratios of EL to EG are close to one, ranging from 0.74 to 1.57. The extremely low value obtained for the Bucharest deteriorations model does not make a rational comparison. The limited evidence here is reasonably supportive of a Hicksian interpretation. Bateman et al (2000) also found EG to exceed EL though by a larger amount with a ratio of 6.64 albeit with an extremely large confidence interval around the EG estimate. They also found CL to exceed CG implying that the difference may be between measures of WTP and WTA. Interestingly Bateman et al (2000) suggest that EL and EG may be the most appropriate measures in precisely the tax v services trade-off envisaged here as they could avoid falsely identifying gain and loss effects that are really based on caution and in more accurately reflecting the context within which policy choices are made.

Alternative arguments based on loss aversion and reference dependency (Tversky and Kahneman 1991) would expect values of CL to exceed CG with EL and EG lying between the two (Bateman et al 2000). In the PR as the cost and aircraft noise coefficients are both moving in the same direction we cannot be sure if any loss aversion is attached to cost or aircraft movements or both. However, as we find that on balance the EG exceeds EL the results do not suggest the presence of loss aversion.

In looking across the experiments for a Hicksian interpretation we would expect CG to be equal to EL and less than EG. Reference dependence would suggest that CG be less than both EG and EL (Bateman et al 2000). Neither interpretation explains the results across the experiments.

Package Effects

What is termed a package effect denotes that the sum of the valuations of the components of a package differs from the valuation of the entire package. It is common in transport SP applications dealing with packages such as vehicle type and terminus facilities which are made up of a large number of relatively minor attributes and it is almost always found that the sum of the component valuations far exceeds the directly estimated total package valuation (Wardman and Whelan, 2001). It is generally assumed, with only a little empirical evidence in support, that SP provides an accurate estimate of the valuations of each component in relation to each other and so these can be scaled to the total package value.

A package effect could influence the comparison of PR and SC based results since the daytime and, to a lesser extent, evening SC valuations are the sum of valuations

for sub-time periods. In line with other evidence, we might expect this to inflate the composite SC valuation. Strategic bias could be a source of package effects other commonly cited causes are unaccounted for interactions, halo effects and budget effects.

An interaction effect denotes that the marginal utility of a particular attribute is influenced by the level of or a change in some other attribute. For example, reductions in aircraft noise might be less highly valued if road traffic noise is also being reduced or if it were felt that there is little point in reducing aircraft noise when road traffic noise levels are very high. The latter contextual interaction applies to both PR and SC. Although the PR method offers changes in both air and road traffic noise it was designed so that there were no interactions. It is therefore unlikely that this type of effect has contributed to the divergence between the PR and SC values.

A halo effect is present when respondents interpret a specified change in an attribute to imply other changes as well. For example, respondents may assume that changes in the specific time period that they evaluated would also occur in other time periods, thereby inflating the valuations. If this were so, it would cause the SC values to be higher than those of the PR exercise not reported here where variations in aircraft movements were explicitly traded-off across time periods. This was not, however, the case (Bristow and Wardman, 2006a). A negative halo effect would arise where tax reductions are not selected in the PR exercise because they imply undesirable cuts in services. However, this pattern of responses is identical to that of someone who has high valuations. The true difference between the PR and SC values would therefore be greater than here estimated.

It could be argued that it is not in fact the SC results that are biased upwards but that the PR values are too low. It might be argued that budget effects are at work in the

PR exercise 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 PR values are almost always lower than the SC values even when the latter excludes Sundays. Secondly, we are not in fact asking respondents to purchase all improvements, but merely evaluate them separately.

With hindsight, we should have obtained more aggregate valuations using the SC method to test for the presence of a package effect within that method. Nonetheless, we should note that there are several instances where the values for a single time period in the SC exercise exceed the values for daytime or evening in the PR exercise. Thus whilst there may be a package effect in operation, it could not reconcile the whole difference between the values obtained by the two methods.

The Choice Format

The SC and PR exercises use different preference elicitation formats and studies have found valuations to differ according to the method used. For example, Boyle et al. (2001) examined ratings, rankings and chose one formats and found the latter yielded the lowest values whilst Caparrós et al. (2006) in a highly controlled experiment also found the values between choice and rank to be similar, though with the choice values slightly lower. There is some evidence that might lead us to expect the SC format to yield lower values than the PR. This is quite the reverse of the case here.

Our study can also contribute to this issue since we undertook a third SP experiment, reported elsewhere (Bristow and Wardman, 2006a), where the same PR method as

illustrated in Table 3 was used but where, as with the SC exercise, the emphasis was very much upon aircraft noise. The quality of life variables were replaced with improvements and deteriorations in the number of aircraft movements in each of the time periods specified in the SC exercise. If we compare the values obtained for deteriorations in this third experiment with the SC results in Table 9, we find that with one exception the values are not significantly different. However, the values in the third experiment are lower in all six comparisons. When compared with the PR results those from the third experiment are in all cases higher and significantly so in 2 of the 4 possible comparisons and nearly significant in a third.

It is evident that the PR experiment based on quality of life variables yields lower values than the SC. However, when a PR experiment is focused on aircraft movements it yields similar values to those in the SC. Our conclusion is that the PR format does not inherently lead to low valuations and that the preference elicitation format is not the reason behind the large differences in valuations obtained by the quality of life PR method reported here and the conventional SC approach.

Task Complexity

Where the SP task is complex, we might expect some respondents to ignore some attributes, in order to simplify what is an artificial choice task, or to resort to choice rules which are less demanding than would be used in practice, as discussed in section 2.2. The consequences of the latter are indeterminate, but the former might be expected to produce values lower than otherwise for the less important attributes that are more likely to be ignored (Hensher et al., 2005).

Undoubtedly the PR experiment did cause respondents some problems. An advantage of the PR method over the SC approach is that it can more readily identify

lexicographic or other decision rules or illogical choices. The removal of the large number of illogical PR responses can be regarded to have isolated those observations where task complexity had been a factor. Whilst this process improved the PR models in terms of fit and precision, the values of aircraft movements do not vary greatly. It therefore seems that task complexity is not the cause of the difference between the PR and SC values.

6. CONCLUSIONS

The contribution of the research findings presented here are twofold. Firstly, they add to the body of evidence relating to the valuation of transport related noise, of which there is very little which has addressed either aircraft noise or the use of SP methods, both of which are covered here. Valuations are reported for different time periods whilst elsewhere we have reported how valuations vary with measured levels of aircraft noise and according to a range of socio-economic factors and made comparisons with other estimates of aircraft noise values, concluding that our results are reasonable (Bristow and Wardman, 2006b). Such detail is important for welfare appraisal but cannot be provided by the more conventional hedonic pricing method. Secondly, they provide insights of a methodological nature arising from the application of different choice formats for preference elicitation, and these may have far-reaching implications for the application of SP methods in the valuation of non-market goods.

In comparing the valuations obtained from a standard stated choice (SC) method and from a novel approach that we have termed the priority ranking (PR) method, there is a broad degree of consistency in the extent to which the SC values exceed the PR values, with ratios between 1.5 and 3.0 for Lyon and Manchester.

We have considered a range of possible reasons why the values differ including: strategic bias, economic theory, package effects, choice format, and task complexity. Our conclusion is that the most likely explanation for the higher values in the SC experiment is the increased incentive to strategic bias since the purpose of the survey would have been more transparent. Whilst this study has not formulated the SP exercises in a sufficiently controlled manner to provide a robust test of the presence and consequences of strategic bias, the results do send out a warning and we would do a disservice by ignoring them. There is also a possibility that package effects are present, and with hindsight it would have been desirable to obtain directly an overall all day valuation using the SC method. There is also some evidence that a framing effect exists, whereby respondents do not fully take into account the time period over which they are required to evaluate aircraft movements.

Despite serious concerns in the early literature, the strategic biasing of responses tends to have been overlooked in recent times, particularly within the SP methodology and recently a more cavalier approach has been adopted based on past successes in SP applications. We believe that more research is required into incentives to bias.

Our conclusion at this stage is that the PR has advantages with respect to the derivation of top level values of noise, the exploration of a range of quality of life issues and its ability to identify illogical preferences. However, if its attraction lies in its ability to mask the purpose of the exercise, it cannot then be used to obtain disaggregated time of day values since the purpose would then become transparent. It is then that the SC method proves its worth, by allowing an overall valuation to be disaggregated into time of day valuations dependent upon differential exposure levels.

The PR method is, as far as we are aware, a novel application in the area of environmental economics. We believe it has considerable potential, not only in terms of making the purpose of the exercise less transparent but also as a means of valuing a wide range of attributes in a less demanding fashion than would be implied by more standard SP procedures. However, further developments are required.

The PR format did cause problems for some respondents but it could be considerably improved by using a computer based tool that could: take the respondent through a simple example, show the respondents baseline, black out already selected options, highlight the available options for each choice to be made and only allow selection of the highlighted attributes.

It would now be interesting to undertake a more controlled experiment using the two types of experiment, with debriefing of respondents, to further explore the extent to which differences in values are due to design effects. However, it will be a challenge to construct an experiment that offers the same attribute levels without collapsing the PR approach to a single line and losing its key appeal. The approach also has the scope for further development to include interactions between attributes and to permit trade-off between gains and losses explicitly.

References

Adamowicz W.L., Boxall P.C., Louviere J.J. and Swait J. (1999) Stated Preference Methods for Valuing Environmental Amenities. Chapter 13 pp. 460-482 in I.J. Bateman and K.G. Willis, eds, *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Methods in the US, EU and Developing Countries*. Oxford 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 E* 39 229-244.

Arsenio E., Bristow A.L. and Wardman M. (2006) Stated Choice Valuations of Traffic Related Noise. Transportation Research D. 11(1) pp15-31.

Bateman I.J., Langford I.H., Munro A., Starmer C. and Sugden R. (2000) Estimating Four Hicksian Welfare Measures for a Public Good: A Contingent Valuation Investigation. Land Economics 76 (3), pp.355-373

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

Boyle K.J., Holmes T.P., Teisl M.F. and Roe B. (2001) A comparison of conjoint analysis response formats, American Journal of Agricultural Economics 83(2) 441-454

Bristow A.L and Wardman M. (2006a) Valuation of Aircraft Noise by Time of Day: A Comparison of Two Approaches. Transport Reviews. 26(4) pp417-433.

Bristow, A.L. and Wardman, M. (2006b) What Influences the Value of Aircraft Noise. Paper presented at AET European Transport Conference, 18th-20th September, Strasbourg.

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. (published) July 2003).
http://www.eurocontrol.int/eec/public/standard_page/eec_2003_all_docs.html

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

Caparrós A., Oviedo J.L. and Campos P. (2006) Comparing Structural Models and Welfare measures in Ranking and Choice Experiments. Paper to the 3rd World Congress of Environmental and Resource Economists, 3rd-7th July, 2006, Kyoto.

Carlsson F., Lampi E. and Martinson P. (2004) The marginal values of noise disturbance from air traffic: does the time of day matter? *Transportation Research D* 9 373-385.

Carson R.T., Groves T. and Machina M.J. (2000) Incentive and Informational Properties of Preference Questions. Draft paper 2000.
http://www.econ.ucsd.edu/~rcarson/cgm.pdf

Caussade S., Ortuzar J de D., Rizzi L.I. and Hensher D.A. (2005) Assessing the influence of design dimensions on stated choice experiment estimates Transportation Research B 39 621-640

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 J. de D. Ortúzar, *Stated Preference Modelling Techniques*, PTRC, London.

Hague Consulting Group (2000) ALOGIT 4.0EC. The Hague.

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.A., Rose J. and Greene W.H. (2005) The implications on willingness to pay of respondents ignoring specific attributes *Transportation* 32 203-222

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

Horowitz, J.L. (1980) The Accuracy of the Multinomial Logit Model as an Approximation to the Multinomial Probit Model of Travel Demand. Transportation Research 14B, pp. 331-341.

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

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

MVA Consultancy (2004) Attitudes to Noise from Aviation Sources in England: Pilot of time-of-Day Stated Preference Study and Design of Standard Annoyance Questions. Prepared for the Department for Transport, London.

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

Samuelson P (1954) The Pure Theory of Public Expenditure. *The Review of Economics and Statistics,* 36(4) 387-389

Timmermans, H. (1993) The Impact of Task Complexity on Information in Multi-Attribute Decision Making. *Journal of Behavioural Decision Making* 6, 95-111.

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).

Tversky A. and Kahneman D. (1991) Loss Aversion in Riskless Choice: A Reference Dependent Model. Quarterly Journal of Economics. 106(4) pp1039-1061

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. (2003) Review of Fare Elasticities in Great Britain. Working Paper 573. Institute for Transport Studies, University of Leeds.

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.

Wardman, M., Bristow, A., Murphy, P. and Heaver, C. (2003) Valuation of Aircraft Noise Using Stated Preference Methods Within a Broader Quality of Life Dimension. Paper presented at AET European Transport Conference. 8th-10th October, Strasbourg.

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

Table 1: Example SC Scenario

	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

Table 2: Levels of Aircraft Movements per Hour in PR Exercise

Location	Time	Minutes between planes going by				
		I	II	III	IV	V
Manchester	Daytime	2	4	4	4	7½
	Evening	2	2	4	7½	7½
Lyon	Daytime	3	6	6	6	10
	Evening	2	2	5	15	15
Bucharest	Daytime	5	20	20	20	60
	Evening	10	10	30	60	60

Note: Column III is the current situation. Columns IV and V denote the combinations of aircraft movements used in the improvements PR exercise. The corresponding movements in the deteriorations exercise are given in columns I and II.

Table 3: Example of SP1: Manchester (Cheadle Area)

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

Table 4: SC Model Results

	Manchester	Lyon	Bucharest
Coefficients			
Constant-Quieter	t=0.4	1.177 (7.8)	-
Flights – Weekday 6am-9am	t=0.6	-0.071 (4.9)	-0.094 (5.6)
Flights – Weekday 9am- 6pm	-0.029 (2.4)	-0.139 (5.1)	-0.082 (5.8)
Flights – Weekday 6pm-10pm	-0.071 (5.6)	-0.087 (5.3)	-0.094 (4.9)
Flights – Saturday 6am-9am	t=0.2	t=1.1	-0.121 (6.5)
Flights – Saturday 9am-6pm	-0.073 (6.3)	-0.033 (2.4)	t=1.2
Flights – Saturday 6pm-10pm	t=1.4	-0.049 (3.1)	t=1.3
Flights – Sunday	-0.093 (9.0)	-0.027 (2.0)	-0.103 (5.0)
Flights – Night	-0.201 (4.9)	-0.086 (2.1)	-0.153 (2.9)
Tax Increases (€)	-0.084 (9.6)	-0.057 (4.6)	-0.497 (3.1)
Values			
Flights – Weekday 6am-9am	-	1.25 (2.8)	0.19 (2.7)
Flights – Weekday 9am- 6pm	0.35 (2.2)	2.44 (2.9)	0.16 (2.6)
Flights – Weekday 6pm-10pm	0.85 (4.2)	1.53 (2.9)	0.19 (2.6)
Flights – Saturday 6am-9am	-	-	0.24 (2.8)
Flights – Saturday 9am-6pm	0.87 (4.6)	0.58 (1.9)	-
Flights – Saturday 6pm-10pm	-	0.86 (2.3)	-
Flights – Sunday	1.11 (5.7)	0.47 (1.6)	0.21 (2.6)
Flights – Night	2.39 (3.9)	1.51 (1.7)	0.31 (2.1)
ρ^2	0.080	0.057	0.031
Individuals	196	208	237
Observations	1532	1647	1895
Option A	772	1324	586
Option B	760	323	1309

Note: t statistics in parentheses. Exclusion of the insignificant coefficients did not materially alter the results and their t statistics are reported for completeness.

Table 5: SC Value per Single Plane Movement (cents)

Period	Manchester	Lyon	Bucharest
Week 6am-9am	-	8.33	1.27
Week 9am- 6pm	0.78	5.42	0.36
Week 6pm-10pm	4.25	7.65	0.95
Sat 6am-9am	-	-	8.00
Sat 9am-6pm	9.67	6.44	-
Sat 6pm-10pm	-	21.50	-
Sunday	6.94	2.94	1.31
Night	4.27	2.70	0.55

Table 6: PR Improvements Model Results

	Manchester	Lyon	Bucharest
Coefficients			
Aircraft: Extremely Noisy	Base	Base	Base
Aircraft: Very Noisy			
Aircraft: Moderately Noisy	0.711 (3.8)	1.148 (4.3)	
Aircraft: Slightly Noisy	1.489 (6.9)	2.914 (15.4)	0.364 (4.7)
Aircraft: Not at all Noisy	1.669 (6.8)	3.482 (11.4)	0.309 (2.3)
Aircraft Movements: Day	-0.084 (3.3)	-0.152 (5.5)	-0.205 (2.5)
Aircraft Movements: Evening	-0.036 (1.7)	-0.227(13.7)	t=0.4
Tax (€)	-0.068 (6.3)	-0.143 (22.0)	-0.421 (5.2)
Values			
Aircraft: Moderately Noisy	10.46 (3.1)	8.03 (4.2)	-
Aircraft: Slightly Noisy	21.90 (4.2)	20.38 (12.6)	0.86 (3.4)
Aircraft: Not at all Noisy	24.54 (4.0)	24.35 (10.1)	0.73 (2.0)
Aircraft Movements: Day	1.24 (2.9)	1.06 (5.3)	0.49 (2.3)
Aircraft Movements: Evening	0.53 (1.7)	1.59 (12.8)	-
ρ^2	0.083	0.062	0.081
Individuals	196	208	237

Note: **t statistics in parentheses. Exclusion of the insignificant coefficients did not materially alter the results and their t statistics are reported for completeness.**

Table 7: PR Deteriorations Model Results

	Manchester	Lyon	Bucharest
Coefficients			
Aircraft: Not at all Noisy	Base	Base	Base
Aircraft: Slightly Noisy			
Aircraft: Moderately Noisy			
Aircraft: Very Noisy		-0.7184 (4.4)	
Aircraft: Extremely Noisy	-0.681 (5.6)	-2.015 (9.9)	-0.253 (2.1)
Aircraft Movements: Day	-0.049 (4.3)	-0.075 (8.1)	-0.051 (3.8)
Aircraft Movements: Evening	t=0.5	-0.069 (8.5)	t=0.8
Tax (€)	-0..062 (8.2)	-0.052 (7.0)	-1.587(14.4)
Values			
Aircraft: Very Noisy	-	13.82 (4.0)	
Aircraft: Extremely Noisy	10.98 (4.9)	38.75 (6.3)	0.16 (2.3)
Aircraft Movements: Day	0.79 (5.6)	1.44 (5.5)	0.03 (3.8)
Aircraft Movements: Evening	-	1.33 (5.7)	-
ρ^2	0.123	0.112	0.106
Individuals	196	208	237

Note: **t statistics in parentheses. Exclusion of the insignificant coefficients did not materially alter the results and their t statistics are reported for completeness.**

Table 8: Comparison of SP Monetary Values

		Manchester	Lyon	Bucharest
PR	Daytime – improvements	*1.24 ±0.85*	*1.06 ±0.40*	*0.49 ±0.43*
	Evening – improvements	*0.53 ±0.62*	*1.59 ±0.25*	*0.0*
	Daytime – deteriorations	*0.79 ±0.28*	*1.44 ±0.52*	*0.03 ±0.02*
	Evening – deteriorations	*0.0*	*1.33 ±0.47*	*0.0*
	Total – improvements	*1.77 ±1.06*	*2.65 ±0.47*	*0.49 ±0.43*
	Total – deteriorations	*0.79 ±0.28*	*2.77 ±0.70*	*0.03 ±0.02*
SC	Weekday 6am-9am	-	*1.25 ±0.89*	0.19 ± 0.14
	Weekday 9am- 6pm	*0.35 ±0.32*	*2.44 ±1.68*	0.16 ± 0.12
	Weekday 6pm-10pm	*0.85 ±0.40*	*1.53 ±1.06*	0.19 ± 0.15
	Saturday 6am-9am	-	-	0.24 ± 0.17
	Saturday 9am-6pm	*0.87 ±0.38*	*0.58 ±0.61*	-
	Saturday 6pm-10pm	-	*0.86 ±0.75*	-
	Daytime (No Sunday)	*1.22 ±0.50*	*4.27 ±2.00*	*0.59 ±0.25*
	Evening (No Sunday)	*0.85 ±0.40*	*2.39 ±1.30*	0.19 ± 0.15
	Total (No Sunday)	*2.07 ±0.64*	*6.66 ±2.38*	*0.78 ± 0.29*
	Total (with Sunday)	*3.18 ±0.75*	*7.13 ±2.45.*	*0.99 ±0.33*

Note: Values are weekly and expressed in € and relate to variations in the number of movements per hour for every hour in the time period specified. 95% confidence intervals are also provided.

Table 9: Comparison of Estimated PR and SC Values

		PR	SC	t (PR v SC)
Manchester	Day Deterioration Evening Deterioration	0.79 (±0.28) 0.0	1.22 (±0.50) 0.85 (±0.40)	1.50
Lyon	Day Deterioration Evening Deterioration	1.44 (±0.52) 1.33 (±0.42)	4.27 (±2.00) 2.39 (±1.30)	2.74 1.55
Bucharest	Day Deterioration Evening Deterioration	0.03 (±0.02) 0.0	0.59 (±0.25) 0.19 (±0.15)	4.65

Note: 95% confidence intervals in parentheses. t test reported in final column