# ANASE

## Attitudes to Noise from Aviation Sources in England

Final Report for Department for Transport

In Association With John Bates Services, Ian Flindell and RPS

October 2007

**mva**consultancy

# Document Control

| | |
|---|---|
| Project Title: | Attitudes to Noise from Aviation Sources in England (ANASE) |
| MVA Project Number: | C34351 |
| Document Type: | Final Report |
| Directory & File Name: | I:\Tep\General\Maria\C34351 ANASE 1st October\Final Report\Attitudes To Aircraft Noise In England_Final Report 231007.Doc |

# Document Approval

| | |
|---|---|
| Primary Author: | Paul Le Masurier, John Bates, Jenny Taylor |
| Other Author(s): | Ian Flindell, Darran Humpheson, Chris Pownall, Alice Woolley |
| Reviewer(s): | Mick Roberts, Ian Flindell, John Bates, Martin Dale |
| Formatted by: | Maria Michael |

# Erratum for the ANASE Report

## 1 November 2007

As a result of the Non SP review group comments (Environmental Research and Consultancy Department, Civil Aviation Authority and Acoustics and Vibration Group, Bureau Veritas, dated 31 October 2007, paragraphs 4.13 – 4.15), the following paragraph and figure should replace paragraph 7.3.4 and Figure 7.2. (The figure in the main report in fact presents the percentage of respondents at least moderately annoyed.)

7.3.4   Figure 7.2 shows the percentage of respondents who were at least very annoyed in each of the sites by LAeq:

- the proportion of respondents who are at least very annoyed is less than 10% for areas with LAeq less than 43dB;

- the proportion of respondents at least very annoyed generally increases with LAeq for values of LAeq over 43dB, although there is a relatively large spread in percentages for most LAeq values; and

- at least 40% of respondents were at least very annoyed for all except one of the areas with LAeq greater than 57dB.
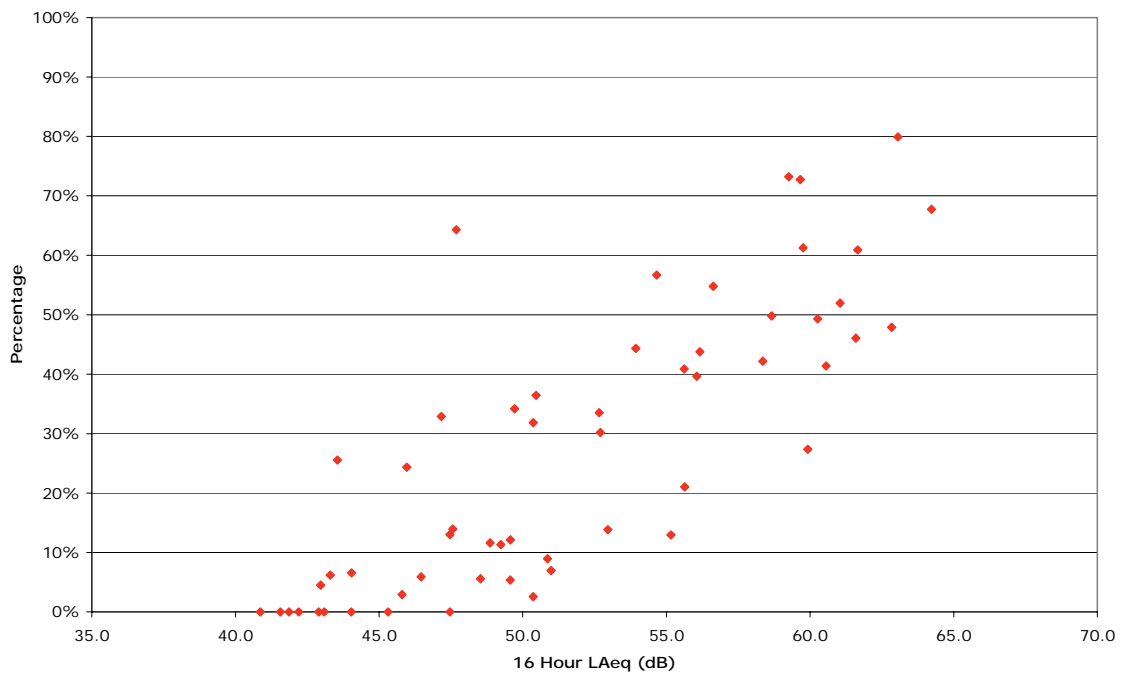


**Figure 7.2  Percentage of Respondents at Least Very Annoyed with Aircraft Noise**

# Contents

# Appendices

*The appendices numbering corresponds to the chapter in which each provides supportive information.*

**mva**consultancy

# 1 Introduction

## 1.1 Study Background

1.1.1 Noise from aviation sources can be an important issue for many residents living near major airports. The Government has kept itself informed about airport noise issues by commissioning, from time to time, surveys of attitudes to noise from aviation sources in residential areas around major airports. Similar surveys have been carried out in many other countries around the world.

1.1.2 Over the past 40 years, several UK studies have sought to quantify the relationship between the amount of aviation (primarily aircraft) noise and the degree of community annoyance that it gives rise to. These enable government and planning authorities to be better informed in their decisions regarding the aircraft noise environment. This report - '*Attitudes to Noise from Aviation Sources in England*' (ANASE) - is the outcome of the research commissioned in 2001 by the Department for Transport[1] to contribute to informed decision making in this area.

1.1.3 Prior to this study, the last major survey of attitudes to aircraft noise in the UK was carried out in 1982 and reported in 1985. This was the **ANIS study** (United Kingdom Aircraft Noise Index Study[2]) which assessed the then existing Government method for measuring aircraft noise around airports, using the Noise and Number Index (NNI). The NNI took into account both average sound levels and the numbers of aircraft noise events exceeding a sound level threshold of 80 PNdB (approximately equivalent to 65 dBA) in a defined 12-hour busy summer daytime period. It included a 'noise and number trade-off' factor of 15 which meant that each doubling or halving of the numbers of aircraft noise events was considered equivalent to a 4.5 dB increase or decrease in average sound levels. Based on previous research carried out in 1961 and 1967, values of 35, 45, and 55 NNI had been considered broadly equivalent to low, medium and high annoyance.

1.1.4 ANIS was based on research carried out at Heathrow, Gatwick, Luton, Manchester, and Aberdeen airports. It concluded that the NNI placed too much weight on the number variable and that a trade-off factor of 9 or 10 would provide a better fit to the data. A trade-off factor of 10 means that each doubling or halving of the numbers of aircraft noise events is considered to be equivalent to a 3 dB increase or decrease in average sound levels. Based on the results of the ANIS study, the government concluded that the NNI should be replaced by a different index – Leq. This index accounted also for the duration of noise events.[3] Furthermore, the ANIS study suggested that, on a 24-hour basis, "55 Leq could be used to represent the onset of community disturbance". The study also noted that, although according to some of the measures tested, there was some evidence of a rapid increase in reported response around this value, the decision on the value of Leq for policy purposes needed to be judgemental since there was "a smooth, almost linear, variation of disturbance with Leq". Following consultation, the UK Government in 1990 adopted the current 16-hour

---

[1] Department of the Environment, Transport and the Regions press notice of 8 May 2001

[2] United Kingdom Aircraft Noise Index Study: Main Report (DR Report 6402) January 1985, prepared on behalf of the Department for Transport by the Civil Aviation Authority

[3] LAeq measures the total amount of "acoustical energy" received at a point, averaged over a specified period of time.

**mva**consultancy

(07:00-23:00) basis for Leq (DORA report 9023, 1990[4]).  It defined the 57 dBA Leq contour as being broadly equivalent to the onset of annoyance, superseding the 35 NNI contour which had previously been taken as an indicator of low annoyance.

1.1.5    Since 1982, however, the overall amount of air traffic has increased significantly whilst the sound levels generated by individual aircraft events have been significantly reduced as older, noisier aircraft types have been replaced by more modern aircraft types with quieter engines and much improved climb performance.  In addition, it is possible that attitudes to aircraft noise may have changed due, for example, to the general growth in personal income, and that the aircraft noise indicator adopted after the 1982 ANIS study (Leq) may be less appropriate for present day conditions.  It was therefore considered timely to see whether the current understanding of the links between reported annoyance and aircraft noise levels still held.

## 1.2    Study Objectives

1.2.1    The stated objectives for this research were as follows:

- ■    re-assess attitudes to aircraft noise in England;

- ■    re-assess their correlation with the Leq noise index; and

- ■    examine (hypothetical) willingness to pay in respect of nuisance from such noise, in relation to other elements, on the basis of stated preference (SP) survey evidence.

1.2.2    Specific issues to be considered in study design included:

- ■    potential to test for differences by locality, socio-economic groups etc;

- ■    distinguishing annoyance from noise at different times of day and night;

- ■    examination of the effect of 'confounding factors' such as airport-related employment or self-selection in housing location;

- ■    the interface between subjective annoyance ratings and valuations derived from stated preference; and

- ■    whether/how attitudes might be affected if cash transfers or, for example, insulating grants were actually made available.

1.2.3    The current study is also intended to inform the policy reported in the Government White Paper '*A new deal for transport*' 2003[5] that the aviation industry should, as far as possible, meet the external costs that it imposes.

## 1.3    A Two-Phase Approach

1.3.1    The ANASE research comprised two main phases.  Phase 1 examined a number of issues relating to the study scope, in the form of a series of pilot studies.  Phase 2 comprised a national survey to explore the attitudes and values of a representative sample of residents in close proximity to some of the major airports around England.

---

[4] The Use of LAeq as an Aircraft Noise Index (DORA Report) 9023, September 1990

[5] Government White Paper, A New Deal for Transport: Better for Everyone, 2003

mvaconsultancy

1.3.2    In addition to these two main phases, there was an interim task, known as the Comparative Performance Trial (CPT), which served as a "rehearsal" for the full set of survey and analysis procedures required in Phase 2. This minimised the risks for the Phase 2 fieldwork.  Hence, not only were the general methods of presentation and analysis carefully developed and piloted throughout Phase 1 of the work, but what was effectively an extended pilot of the finalised survey was carried out prior to the commitment of the full fieldwork programme.

1.3.3    The main issues investigated as part of Phase 1, and the main technical tasks involved in planning and executing Phase 2 are illustrated in Figure 1.1.

**Phase 1**
- Qualitative Pilot
- Noise at Night Pilot
- Personal Aircraft Noise Exposure Pilot
- Low Noise Area Pilot
- SP Pilot
- Presentation Pilot
- Comparative Performance Trial

**Phase 2**

| Acoustic Data Collection | Questionnaire/ Stimulus Material | Fieldwork Procedure | Sample Design | Literature Review |

Full Survey and Restricted Survey (National) Fieldwork

Data Analysis

Figure 1.1  ANASE Study Programme

mvaconsultancy

## 1.4    Review Process

1.4.1    Four distinct advisory project committees were established during the course of the study, with the following roles:

- ■    *a Steering Group*: to oversee the development of the study;

- ■    *an international Peer Review Group:* an assembled group of international experts from whom the DfT obtained advice on the initial technical approach;

- ■    *an SP sub-group:* a subset of steering group members with SP expertise, plus invited SP experts from the transport and environment fields to review the development of the SP approach; and

- ■    *a Non-SP sub-group:* invited technical experts who reviewed the non-SP analytical and modelling results of the study.

1.4.2    These groups were involved at different stages in the study, and their input has been greatly valued by the study team.  We have also benefited from close involvement of individual members of the Department for Transport's Aviation and Statistics Divisions.

1.4.3    Nevertheless, the views put forward in this report are those of the study team alone.

## 1.5    Report Structure

1.5.1    The following two chapters set out elements of methodology necessary for an appreciation of the following chapters.  A description of sound level indicators, acoustics data collection and modelling is provided in Chapter 2.  Chapter 3 contains a discussion of the approaches to aircraft noise valuation used in the study.

1.5.2    The content and main findings of Phase 1 are described in Chapter 4.  The sample design for Phase 2 is contained in Chapter 5.  Details of the fieldwork procedures adopted in Phase 2 are provided in Chapter 6.

1.5.3    The presentation of reported annoyance is given in Chapter 7, followed, in Chapter 8, by a description of the analytical work conducted to establish quantitative relationships between reported annoyance and the various aircraft sound level measures tested.

1.5.4    Chapter 9 is devoted to an assessment of changing attitudes to aircraft noise, from a comparison of the ANIS and ANASE studies.

1.5.5    Results for the valuation elements of the study are presented in Chapter 10.

1.5.6    We complete each of the analysis chapters, Chapter 7 to 10, with a summary of the main points.  The study conclusions are given in Chapter 11.

1.5.7    A series of appendices contain detailed supporting material; the appendices numbering corresponds to the chapter in which each provides supportive information.

**mva**consultancy

# 2 Sound Level Measurement

## 2.1 Introduction

2.1.1 This chapter describes the methods and procedures used to obtain appropriate indicators of aircraft sound levels to represent physical or objective exposure to aircraft noise for the following purposes:

- defining the total population resident around airports and therefore available for sample selection within the overall scope of the survey (see Chapter 5 for further details);

- plotting the distributions of  average event sound level (Lav - see below) and average number of events (Nav - see below) for all census output areas defined as available for sample selection within the scope of the survey as above (see Chapter 5);

- providing representative values of the outdoor sound level indicators for each selected 'common noise area' (see Chapter 5) as required for the subsequent statistical analyses;

- providing recordings calibrated to be representative of actual outdoor sound levels at each full SP survey site as part of the context to the trade-off exercise; and

- calibrating the modelled aircraft sound levels against field measurements of aircraft sound levels carried out specifically for that purpose.

## 2.2 Sound level Indicators

### Decibel scales

2.2.1 All sound levels used in ANASE are based on decibel scales.  The word 'level' is an indication that a decibel scale has been used.  It is possible to measure sound in basic SI units of sound pressure, power and energy, but it has been industry standard practice for many years to use decibel scales instead.  Decibel scales are used in many fields of engineering, and because they are essentially logarithmic ratios, they do not have physical units in the same way as units of sound pressure (Newtons per square metre) or units of sound power (watts).  In acoustics, decibel sound level differences are calculated by taking 10 times the logarithm (to the base 10)[6] of the squared ratio of one sound pressure to another sound pressure.  Because sound power is proportional to the square of the sound pressure, decibel sound level differences can also be calculated by taking 10 times the logarithm (to the base 10) of the ratio of one sound power to another sound power.  It should also be noted that sound energy is sound power multiplied by time.

---

[6] Note that the use of logarithms to base 10 is conventional in the treatment of sound: this is conceptually distinct from the multiplication of the logarithm by the factor "10", which simply converts the coarser unit "Bel" to decibels in the usual manner of measurement in metric units

**mva**consultancy

2.2.2   The Sound Pressure Level (this is often referred to just as the Sound Level) as measured by a sound level meter is defined as 10 times the logarithm (to base 10) of the squared ratio of the sound pressure being measured to the standard reference sound pressure, which is defined in SI units as 0.00002 N/m$^2$ (20 micro-Pascals).  If $p_t$ is the sound pressure at time t due to some event, and $p_0$ is the reference pressure (20 µPascals), the sound pressure level at time t, L(t), is given, in decibels, by the formula[7]:

$$L_t = 10 \log_{10}\left( \frac{{p_t}^2}{{p_o}^2} \right)$$

### LAmax

2.2.3   At some time within the duration of the noise event, $L_t$ will have a maximum value.  LAmax is defined as the maximum A-weighted sound level received during that event.  For sounds with rapid fluctuations such as aircraft events the actual value of LAmax recorded by a sound level meter will vary depending on the meter averaging time or 'time weighting' used.  For aircraft sound level measurement the S (or slow) time weighting is used because this averages out for the effect of the rapid fluctuations in sound level caused by atmospheric turbulence.

### Lav

2.2.4   For a given location, the LAmax values for a series of aircraft events can vary over a wide range, so for a series of events we could define an average LAmax, a maximum LAmax, a minimum LAmax, and a distribution of LAmax.  For ANASE, we defined the average event sound level or Lav as the arithmetic average of the separate LAmax values for each of the events contributing to the average.  Lav is calculated as follows;

$$L_{av} = \frac{1}{n} \cdot \sum_{i=1}^{n} L_{A\max, i} \qquad \text{where 'n' is the number of events}$$

2.2.5   The distribution of LAmax at any measurement site always extends downwards to the steady background sound level and below because of the residual effect of increasingly distant aircraft events. The extent to which any particular aircraft noise event will be heard on the ground depends on the margin by which the aircraft sound level exceeds the steady background sound level, and also on what the listener is doing at the time.  Clearly, listeners are unlikely to hear aircraft noise events which do not intrude above the steady background sound level, and may not hear aircraft noise events which only intrude above the steady background sound level by small amounts.  For this reason, it is always necessary to employ a cut-off point to prevent the increasing numbers of quieter events which are not heard from biasing the average downwards.  For ANASE, we tested a range of cut-off points from 50 to 75 LAmax.  The lower cut-off means that all aircraft events down to 50 LAmax are included in the averaging.  The higher cut-off means that only those aircraft events which exceed 75

---

[7] Aircraft sound levels are normally measured using an A weighting filter.  This filter reduces the sensitivity of the measuring instrument to low frequency and to very high frequency sound to approximately correspond to the frequency sensitivity of the average human listener.  Typically the $L_t$ measure will be the A-weighted, and can be written as $LA_t$, though the "A" is often assumed rather than actually written.

**mva**consultancy

LAmax are included in the averaging.  A cut-off of 65 LAmax was used for the initial sample selection, but both higher and lower cut-offs were tested in the subsequent statistical analyses.

2.2.6   The 65 LAmax cut-off has no effect at the noisiest (closest in) receiver sites where only the very quietest aircraft types would be excluded by its adoption.  At receiver sites which are increasingly distant from the airport, the 65 LAmax cut-off excludes an increasing proportion of the quieter aircraft types from the calculated sound levels.  At the most distant receiver sites, only the noisiest aircraft types are included in the calculated sound levels.  These effects are shown in more detail in Appendix A2.

2.2.7   Over the range of modelling cut-off sound levels tested, the effect on the calculated logarithmic average sound level over a defined time period, or LAeq (for a definition of LAeq, see paragraph 2.2.11 below) is negligible.  This is because changing the modelling cut-off sound level has no effect on the noisiest aircraft type and route combinations included in the calculation.  The noisiest aircraft type and route combinations dominate the LAeq calculations at all sites.  However, increasing the LAmax modelling cut-off sound level clearly has an effect on both the number of aircraft events or Nav (for a definition of Nav, see paragraph 2.2.8 below) and Lav at most sites, with Lav increasing and Nav reducing.  This is because as the LAmax modelling cut-off sound level is increased, the number of quieter aircraft type and route combinations taken into account in the calculation reduces.

### Nav

2.2.8   For ANASE we defined Nav as the number of aircraft events in a given time period which exceed the defined LAmax cut-off (which is 65 LAmax unless otherwise stated).  A lower cut-off will increase the number of aircraft events contributing to the Nav statistic.  .

2.2.9   In this report, the standard time period for counting aircraft events is 30 days unless otherwise stated.

### SEL

2.2.10  The sound exposure level (SEL) of an aircraft event is defined as the sound level in A-weighted decibels of a one-second burst of steady sound which contains the same total A-weighted sound energy as the whole event.  Because SEL can be used as a proxy measure for the sound energy in a single aircraft event (or a series of events if aggregated together) it is often referred to as the Single Event Level, although this would be incorrect if there were more than one event included within the measurement.  If L(t) is the instantaneous sound level at time t, then the total sound energy E(s) of an event lasting T seconds is proportional to $\int 10^{L(t)/10} dt$ .  SEL is then calculated as 10 $\log_{10}$ E(s), and can be measured using integrating sound level meters. For a steady sound of increasing duration SEL increases in proportion to the logarithm to base 10 of the duration in seconds, whereas LAmax does not change.  For typical aircraft events most of the sound energy is concentrated in the time period during the middle part of the flyover when the instantaneous sound level is within 10 dB of LAmax, and the standard measurement procedures reflect this. For typical aircraft flyover events SEL is often around 10 dB higher than LAmax, reflecting an equivalent duration during the middle part of the flyover event of around 10 to 20 seconds.

**mva**consultancy

**LAeq**

2.2.11   LAeq is defined as the logarithmic average sound level over a defined time period, and as such, it takes into account both the sound levels and durations (SEL) of separate events and the quiet periods in between.  The explanation of LAeq which follows is different from that which is normally found in textbooks, but it is mathematically equivalent and may help to explain some of the more detailed technical content found in Chapter 10. For a series of N equivalent aircraft events, each with total sound energy E(s) (see paragraph 2.2.10 above), the total sound energy is E(T) = N.E(s).  If this is then divided by the total time T during which these events occur, we derive a quantity which is proportional to the rate at which sound energy is received at the measurement site, which is proportional to average sound power and hence directly proportional to average sound level.  Converting this back to decibels by taking the logarithm to base 10 and multiplying by 10 we obtain the LAeq.

2.2.12   Hence:

$$\text{LAeq} = 10 \log_{10} (N.E(s)/T)]$$

$$= 10 \log_{10} N + SEL - 10 \log_{10} (T).$$

2.2.13   For a 16 hour period (ie 57,600 seconds), the formula can be written as:

$$\text{LAeq} = 10 \log_{10} N + SEL - 47.6.$$

2.2.14   Because LAeq takes into account both the sound level (LAmax) and the duration of the separate events included within it, it cannot be calculated from just the average sound level (Lav), the number of events (Nav), and the defined time period because the duration of the events is not taken into account by either of the Lav or Nav variables.  To calculate LAeq we must use average SEL and the number of events because SEL takes into account both the sound level and the durations of the separate events.

2.2.15   In chapter 10, we consider alternative formulations which weight the sound level and number variables differently from the weightings within LAeq.  For these calculations we use Lav and Nav rather than average SEL and Nav as input variables, even though the Lav + k log Nav formulation is not strictly the same as LAeq.  We could not use the alternative formulation of SEL + k log Nav for these calculations because the SEL formulation already assumes a 10 log relationship for the effect of duration, and this precludes the SEL formulation from being used for any other k value.

**2.3   The INM 'Integrated Noise Model' for calculating aircraft sound levels**

2.3.1   Aircraft sound levels vary depending on the type of aircraft and how it is flown; the height of the aircraft, the lateral displacement of the flight track to either side of the measurement point, and the number of aircraft events within any defined period if considering an average or aggregate measure of the average sound level or of the overall amount of sound energy received.  Because of this large variation from one receiver site to the next and because of similarly large variation from one day to the next, in any study of this type it is not practical to be able to determine long term average sound levels by any method of measurement alone.  To solve this problem a number of sound level calculation models have been developed with varying degrees of complexity depending on the application.  It has now

**mva**consultancy

become widely accepted as best practice to first calculate sound levels from basic input data using a standard mathematical model and then to test and if necessary re-calibrate the input data used to construct the model against a limited sample of field measurements which do not need to be carried out at every receiver site.

2.3.2   The two most commonly used aircraft sound level calculation models in the UK are the CAA's proprietary model known as ANCON which is used to produce 'official' annual aircraft noise contours at the 'designated' airports and at many others and the United States Federal Aviation Administration's INM or 'Integrated Noise Model' which is widely used both in the UK and internationally.   Both models, providing that they are used correctly, are fully compliant with current best practice international guidance as set out in ECAC.CEAC Doc 29[8].

2.3.3   For ANASE we used INM version 6.2 as the basic aircraft sound level calculation engine because this approach provided the maximum flexibility for calculating alternative sound level indicators for use in subsequent statistical analyses.  A separate comparison against spot values for Heathrow calculated using ANCON and provided to the research team by the CAA showed that although there were small differences between the results obtained using the two models, the differences were too small to have any material effect on the overall conclusions.  It should be noted that all calculation models are subject to varying degrees of uncertainty which increase at lower sound levels.  There is no a priori reason to assume that the residual levels of uncertainty associated with either ANCON or INM are not similar.  The basic aircraft sound level data calculated using INM was then aggregated together to produce a range of composite sound level indicators using simple EXCEL spreadsheets as required.

2.3.4   A series of field measurements was carried out at 19 sites during the main Phase 2 survey period during the summer and autumn in 2005 and the results used to re-calibrate the input assumptions used to inform the INM models wherever necessary.  This part of the work is described in more detail in section 2.8 below.  It should be noted that it was not feasible to carry out any comparisons of the model outputs against field measurement data at any earlier stage before the actual survey sample sites had been selected. There is thus an implication that the sound level data used for selecting the survey sample sites was subject to marginally greater uncertainty than the re-calibrated sound level data used for the subsequent statistical analyses.

## 2.4   Calculations of LAmax for defining overall area in scope

2.4.1   While most people have some experience of aircraft overflights, wherever they live, the main focus of the ANASE study was on people living sufficiently near to major civil airports that aircraft noise either is, or could be, a significant cause of annoyance.  It was necessary to limit the proportion of the overall population that would be included within the scope of the survey to prevent the majority of the survey resources from being effectively 'wasted' by being expended in areas which were not materially affected by aircraft noise. This required a procedure for setting an outer boundary around each airport that would be included within the overall study area.   An outer LAmax boundary was calculated according to a set of assumptions which are described below.

---

[8] **ECAC.CEAC Doc No. 29,** 3rd Edition, Report on Standard Method of Computing Noise Contours Around Civil Airports Volume 1: Applications Guide & Volume 2: Technical Guide, Dec 2006

**mva**consultancy

2.4.2   First, the twenty largest commercial airports were selected according to the procedures set out in Chapter 5.  For each of these airports a basic INM model was set up to model the outer boundary within which the noisiest type of aircraft known to fly regularly at that airport would exceed 65 LAmax on the ground having 'flown' any of the known arrivals and departures routes as shown in the UK Aeronautical Information Package (AIP – 'Air Pilot'). This exercise was repeated for a range of higher LAmax cut-offs (leading to successively smaller outer boundaries) and the overall populations resident within each LAmax cut-off value calculated using standard GIS procedures.  The overall population resident within the 65 LAmax outer boundary was found to be in excess of 3,000,000, with successively smaller populations resident within each higher LAmax outer boundary.  Based on this information, the Department decided that all residential areas within the 65 LAmax outer boundary as calculated in this way should be included within the scope of the survey.

2.4.3    The 65 LAmax cut-off which was actually used is of course based on judgement rather than on any scientific criterion, but it is generally consistent with cut-offs used in previous research, such as ANIS.  In the next stage of selecting the much smaller "common noise areas" used for statistical sampling, additional stratification by Lav and Nav was employed to ensure that areas with higher levels of aircraft noise exposure were adequately sampled notwithstanding the much smaller numbers of population exposed at the higher sound levels.

**2.5     Calculating the distributions of Lav and Nav for all census OAs defined as in scope**

2.5.1   The next step was to obtain detailed air traffic data for each of the twenty commercial airports now defined as in-scope to inform fully detailed INM models of each airport which would then allow any desired sound level indicator to be derived.  For each airport, the aircraft sound level calculations required full information about each aircraft type flying at that airport and the numbers of operations of each of those types using the various routes to and from that airport.  The INM models were used to calculate average LAmax and SEL values for each aircraft type and route combination for each census output area centroid previously defined as in scope.  These interim data were then input into simple EXCEL spreadsheets to derive the various combined indicators required which were aggregated according to the air traffic data provided.

2.5.2   Full details of the calculation procedures are given in Appendix A2.  It should be noted that not all of the twenty airports were able to provide the required air traffic data within the limited timescale for the survey sample selection to be completed, and at some other airports the traffic was found to be so irregular that meaningful assessments of long term average sound level indicators could not be made.  Separate arrangements then had to be made for selecting limited sampling at these other airports, as discussed in Chapter 5.

2.5.3   For the ten airports remaining in scope for the full sample (because of the regular traffic statistics - see Chapter 5) the distributions of Lav and Nav were then plotted out to allow the cutting points for the sampling matrix described in Chapter 5 to be selected.  The cutting points at 71 and 77 Lav and at 380, 1200, and 3800 Nav shown in Table 5.1 were selected based on the observed distributions of these variables to provide the most even coverage in each cell of the matrix across the whole range of each variable, and this led to the final decision to use 36 full SP survey sample sites.

2.5.4   The lower sampling matrix cut-off point of 65 Lav was a direct consequence of the decision

**mva**consultancy

to use 65 LAmax as the cut-off point for the overall sample in scope, since if there were no aircraft exceeding 65 LAmax then Lav must be lower than 65.  It should be noted that within the areas in which the noisiest aircraft types exceed 65 LAmax, there will normally be many quieter aircraft types which do not exceed 65 LAmax.  The lower cut off at 120 Nav represents a minimum of one aircraft event in any daytime and evening 4 hour time period (0700 to 2300 hrs) over the standard 30 day counting period for Nav.

2.5.5    The final step was to re-calculate LAmax contours for the noisiest aircraft types at each airport in 0.5 dB steps across each in-scope sample area to determine variation in LAmax across each census output area selected by the stratified random sampling procedure described in Chapter 5.  The intention here was to identify any selected census output area within which aircraft sound levels varied by more than 3.5 dB to ensure that only census output areas with lesser variation in aircraft sound levels would be included as 'common noise areas'.  The reason is that aircraft sound levels calculated for each census output area centroid were assumed to be representative of all residential addresses within that output area.  Variation in aircraft sound levels in excess of 3.5 dB was judged as violating that assumption.  If any such output areas with excess variation in aircraft sound levels had been selected they might either have needed to have been excluded from the sample (and re-sampled) which was not desirable from the statistical point of view, or special instructions would have had to have been devised to limit interviews to within areas within the output area with variation less than 0.5 dB, which would have created further problems.  Fortunately, no such areas were selected by the random sampling process so the problem did not arise.

## 2.6    Calculating Lav, Nav, SEL, LAeq for the selected survey sample sites

2.6.1    Up until the survey sample sites had actually been selected, the aircraft sound level calculation models had to be run for all census output areas previously defined as in scope, representing a total residential population approaching 3,000,000.  This represented a very large computing load, and explains why it was not feasible to carry out any validation checks against field measurements before the survey sample sites had been finally selected.  As soon as the survey sample sites had been selected it was then possible to go back and run the aircraft sound level models in more detail to produce the more detailed sound level statistics required for each selected sample survey site.

2.6.2    Each INM model was developed to take into account additional details about aircraft operations which might be relevant to actual aircraft sound levels at the selected sample survey sites.  These additional details included more detailed assumptions about aircraft stage length, flight track dispersion to either side of the nominal flight routes as shown in the UK AIP or on airport Noise Preferential Routes and any other more detailed information that became available and was considered likely to have some effect on actual sound levels.  It was found that by assuming heavier take-off weights than for the initial INM defaults assumptions, some of the calculated sound levels (for departures) were marginally increased as compared to the initial set of calculations used for the sample site selection.  Including flight track dispersion in the calculations only had very small effects on the calculated LAeqs, although the effects on Lav and Nav separately were more significant.

2.6.3    Because the first stage of aircraft sound level modelling had to be completed in the first few months of 2005 to inform the sample survey site selection process, it was necessary to base

**mva**consultancy

these calculations on summer 2004 traffic.   Each airport was requested to provide comprehensive details of summer 2004 traffic for a representative busy month, which was either June of July for most airports.   Stansted airport provided data for January 2005. Because the particular months for which the air traffic data was provided might not have been representative of the long term runway direction modal split at each airport, the calculations were then weighted according to the long term runway direction modal split at each airport.   At Southampton, London City, Luton and Leeds, these weightings were based on five years runway direction data; at Birmingham on eight years runway direction data; and at Manchester, Heathrow, Gatwick, and Stansted on ten years runway direction data.

2.6.4    The requirement for the first stage of aircraft sound level modelling to be completed in advance of the sample survey selection process meant that whereas the questionnaire response data applied to 2005 traffic, the corresponding aircraft sound level data applied to 2004 traffic.   While the total amount of air traffic has been increasing year on year over the past few years, the percentage change in overall traffic from 2004 to 2005 was not large enough to have any major effects on long term average sound levels considered overall. Nevertheless there have been changes to the detailed pattern of aircraft types operated and routes flown at particular airports which could have had more significant effects at particular sample survey sites.   Because the Heathrow data represented approximately half of the entire data set, all aircraft sound levels for Heathrow were re-modelled based on summer 2005 traffic, and in this case using the CAA preferred 92 day summer period as the basis period for the air traffic data.   This re-modelling work was carried out at the same time as the various re-calibrations against field measurement data which are further described in section 2.8 below.   The re-modelling work for Heathrow led to some small changes in the detailed statistical analyses of the results but it did not materially affect the overall conclusions of the study.

**2.7    Calibrated recordings representative of outdoor sound levels for use in the SP interview procedures**

2.7.1    The full SP interview procedures required calibrated recordings of representative generic aircraft types suitable for reproduction via portable loudspeakers at sound levels which were representative of actual outdoor sound levels for those aircraft types at each sample survey site.    These recordings were made by visiting the actual sample survey sites wherever possible.   However, existing background sound levels at the actual sample survey sites were in most cases too high to permit sufficiently clean recordings to be obtained and so instead, many of these recordings were obtained at quieter sites (ie away from roads and other sources of background sound) in similar orientations relative to the aircraft flight tracks and then edited to calibrate the reproduced sound levels to the 'correct' levels.   The 'correct' sound levels were initially defined according to the INM models for each sample survey site, but were subsequently adjusted up or down as necessary when informed by actual field measurements as soon as these had been completed at each survey site.   Further details are provided in Appendix A2.

**2.8    Field measurements for testing and re-calibrating the INM models**

2.8.1    Comprehensive field measurements were carried out at 19 of the sample survey sites during the summer and autumn of 2005.   The field measurement sites were chosen on the basis

**mva**consultancy

that they were likely to produce the most useful information for use in first testing and then re-calibrating the models as appropriate.  No field measurements were carried out at sites where limited numbers of aircraft overflights or high background sound levels would have constrained the usefulness of the resulting data.  Similarly, where selected sample survey sites were relatively close together, it was not necessary to carry out field validation measurements at more than one of each cluster of sample survey sites.

2.8.2    The comparisons showed some variations between modelled and measured data which were in all cases attributable to small differences between the input assumptions made to produce the initial set of aircraft sound levels for use in the sample survey selection process and more realistic assumptions based on actual field observations.  For example, it was found that by assuming greater departure stage lengths than the default assumptions applied as standard within INM provided closer correspondences between modelled and measured data in most cases.  It should be noted that data on actual stage lengths as flown during the survey periods were not available, but detailed perusal of scheduled flight destinations provided some support for these changed assumptions.  Regarding flight track dispersion, within the scope of the ANASE study, it was not possible to take into account actual flight tracks recorded for every flight during the survey period.  Comparisons between modelled data with a range of different flight track dispersion assumptions and the field measurement data showed that LAeq calculations were relatively insensitive to the flight track dispersion assumptions made.  For this aspect of the work, we simply used the default assumptions set out in ECAC.CEAC Doc 29, as mentioned earlier.   It should be noted that actual flight track dispersion, while generally have little effect on LAeq sound levels, can have more significant effects on Lav and Nav calculated separately, particularly where a range of different LAmax cut-off values are used.  For arrivals traffic it was found expedient for some aircraft types to model marginally different glide slopes (by up to 0.5°) from the standard $3^0$ glide slope as actually flown in order to obtain the closest possible correspondence between modelled and measured data.

2.8.3    Figure 2.1 shows the high degree of correspondence obtained between modelled and measured SEL values for arrivals with the adjusted glide slope assumptions applied as set out above.  Figure 2.2 shows the similarly high degree of correspondence obtained between modelled and measured SEL values for departures with the adjusted departure stage length assumptions applied as set out above.
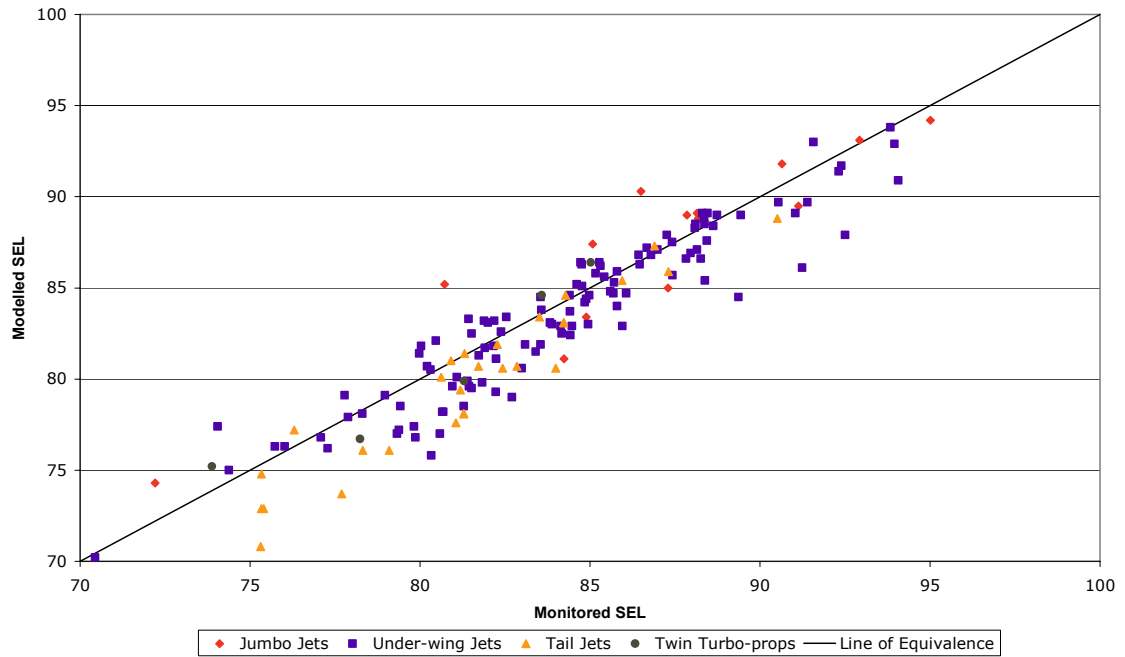
**mva**consultancy

**Figure 2.1  Comparison between Modelled and Measured SEL – Arrivals**
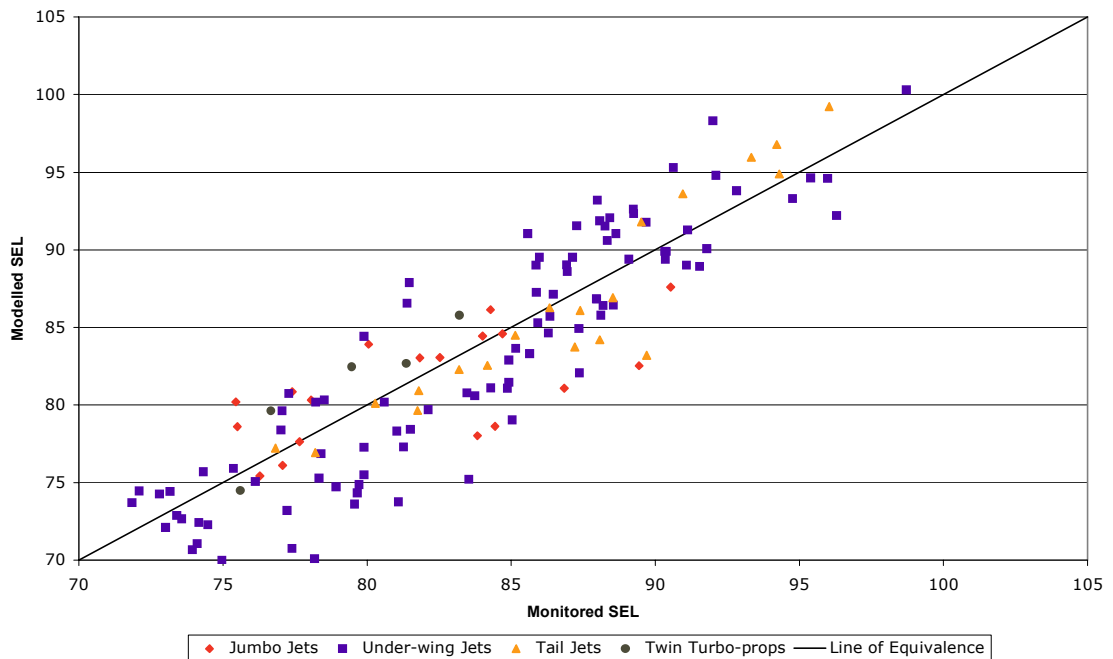


**Figure 2.2  Comparison between Modelled and Measured SEL - Departures**

# 3  Valuation Methodology

## 3.1  Introduction

3.1.1  One of the three objectives of the study – marking a major departure from ANIS and its predecessors – was to "examine (hypothetical) willingness-to-pay in respect of nuisance from [aircraft] noise…". This will inform the value that people give to relief from noise and aid policies designed to ensure that the aviation industry meets its 'external' costs. The methodology for this was prescribed in the study specification: that it should be on "the basis of stated preference survey evidence".

3.1.2  There is some confusion over the terminology. Within the transport modelling field, stated preference has generally denoted a set of procedures in which respondents are offered a set of hypothetical options varying according to a limited set of attributes, whose values are set by rules relating to experimental design: respondents are then asked to choose the 'best' option, or to rank the options presented, or otherwise to rate them in some way. This corresponds to what is often referred to within the field of Market Research as Conjoint Analysis. However, stated preference is also used as a generic term for all kinds of questions based on hypothetical constructs, of which the Contingent Valuation Method (CVM) is an approach widely used within the field of environmental assessment.

3.1.3  We have chosen to use the more restrictive definition of stated preference, and in particular, given the way the field has developed, we use 'stated preference' to refer to the particular form of stated preference often referred to as "Stated Choice" (also referred to as "Choice Modelling" and "Contingent Choice"). Hence, we distinguish explicitly between stated preference – SP throughout this report - on the one hand, and CVM on the other.

3.1.4  SP (in this sense) is becoming a fairly common technique in the field of valuing environmental factors. But noise (from any source) has been relatively little investigated using SP. This is partly because of the sheer difficulty of constructing scenarios in terms of different amounts of noise, and then presenting them to respondents. Certainly, this was seen at the outset of the present study as probably the greatest challenge to the successful application of SP to the valuation of annoyance from aircraft noise. Much of Phase 1, therefore, comprised a sequence of sub-phases that explored (mainly) elements of developing SP into a fit-for-purpose tool for the main study.

3.1.5  The ANASE study was specifically set up to investigate the application of SP methods to the valuation of aircraft noise annoyance in England. Following a piloting exercise during Phase 1, it was agreed that a CVM question should be included in the Phase 2 survey for comparison.

### Willingness-to-Pay

3.1.6  When consumers purchase goods or services ordinarily, the implication is that they prefer what they have bought to the alternative of retaining the money paid in their purses or wallets. Their "willingness-to-pay" for the goods or services concerned is therefore at least as great as the prices charged. If prices were reduced, the willingness-to-pay of the same consumers would not change, but additional consumers would be attracted. These clearly did not have a willingness-to-pay as high as the original prices, but can be inferred to have a willingness-to-pay at least as high as the new, lower, prices. In practice, of course, at any

**mva**consultancy

given level of prices it would not be possible to distinguish between those purchasers with higher or lower willingness-to-pay.

3.1.7    Several important points are illustrated by the previous paragraph.  First, willingness-to-pay is equivalent to stating that the goods or services purchased are worth to consumers at least as much as they pay for them.

3.1.8    Second, at a given price level, the willingness-to-pay of most purchasers will in fact be higher than the price: possibly very much higher.  In effect, only a preference ordering of the purchase/not-purchase options is observed.

3.1.9    Third, willingness-to-pay is clearly not a function of the prices charged.  So there can be a willingness-to-pay for goods or services – or anything else, such as an improved environment – that have a zero price or are received without need for payment.

3.1.10    Fourth, (potential) consumers' willingness-to-pay for a given good or service will vary.  Some will place a value on having it at least as high as its price, and so will purchase it, while those who do not value it as highly as its (current) price will not purchase it.

3.1.11    Fifth, at a given price, actual purchasing behaviour will distinguish only between the two broad groups of (potential) consumers described in the previous paragraph.  For something with a zero price, or received without need for payment, even this distinction becomes meaningless.

3.1.12    Two further inferences can be drawn:

- if willingness-to-pay for something could be established with reasonable reliability and accuracy, this would be a justifiable approach to estimating people's value of it; and

- particularly for something with a zero price or received without need for payment, willingness-to-pay cannot be inferred directly from actual purchasing behaviour; some alternative – unavoidably indirect – approach is needed.

3.1.13    These inferences are particularly germane in the field of environmental appraisal, where benefiting from improvements typically does not require payment by individuals, and they have motivated the development and application of several indirect techniques.

### Valuing Noise Impacts

3.1.14    A number of studies have been carried out in recent years, investigating the value people put on marginal changes in noise from transport sources and, specifically, from aircraft noise.

3.1.15    Three main methodologies are used in the valuation of environmental impacts: hedonic pricing (HP), CVM and SP.  They have been usefully defined as follows, in relation to aircraft noise, in a recent (2006) paper by **Brooker**:

- **Hedonic pricing** investigates the extent to which 'people may be willing to live in an area that is subject to aircraft noise, but only if they receive a discount on the price: the size of the discount measures their aversion to aircraft noise exposure';

- **CVM** studies ask people in a survey 'how much they would be willing to pay for an aircraft noise environment, or the amount of compensation they would be willing to

**mva**consultancy

accept to give it up: thus, willingness to pay[9] is contingent on a specific hypothetical scenario'; and

■   **SP** studies also ask people 'to make valuation choices based on a hypothetical scenario, but it does not ask people to state their values directly: values are inferred from the hypothetical choices or trade-offs.'

3.1.16   CVM's approach is through direct willingness-to-pay and/or willingness-to-accept questions, while the format of SP studies is such that respondents are asked to choose between (hypothetical) options rather than to undertake (hypothetical) transactions. By including money payments as one of the attributes of the options, the indifference curves with respect to money can be derived.

3.1.17   In all cases, the essence of all methods of valuing "non-market commodities" (ie in this context, aspects of the product which cannot be directly traded for money) is firmly based on the concept of "willingness to pay". In considering the valuation of a single commodity, it is natural to work with the implicit indifference curves between the commodity and "money". In other words, we seek the amount of money £X which the respondent would be willing to pay (accept) in return for the gain (loss) of a specified amount Y of the (non-market) commodity, in such a way that they are indifferent between their **current position** and the new "option" (−£X,+Y) [or (+£X,−Y)].

3.1.18   The remainder of this chapter considers the three different methodologies in more detail.  A summary of the use of these methods in previous studies is given in Appendix A3.

### 3.2    Hedonic Price Studies

3.2.1   Hedonic price studies are based on actual behaviour where preferences for quieter environments are revealed by higher prices paid for houses in those areas.   In essence, this presumes that – after allowing so far as possible for intrinsic differences between properties (such as the size and condition of the house, the type of area, and the location with respect to desired facilities) – the difference in prices is a reflection of the willingness-to-pay to be subject to less noise.

3.2.2   Previous hedonic price studies have derived values for the Noise Sensitivity Depreciation Index (NSDI) in residential areas around airports ranging from around 0.5% to 1% per dB, other things being equal.  This implies a 5% to 10% reduction in price for a 10 dB increase in noise levels.

3.2.3   In the absence of a directly-priced market in noise exposure, hedonic pricing is not to be lightly dismissed.  A high-quality study clearly has major issues to address, however.  Such a study ideally requires the actual selling prices of properties at more-or-less the same time (to eliminate market fluctuations) but must often fall back on advertised prices or estimates of property values from estate agents.   There is the difficulty of finding otherwise "comparable" properties in areas of different noise levels, or of seeking to account fully for the effect of intrinsic property differences on price differentials.  In all this, nothing is known of the circumstances or feelings of purchasers or vendors; it seems inevitable, for example,

---

[9] The study objective includes the phrase 'willingness to pay'.  It should be noted that a respondent's 'willingness to accept' monetary compensation is often found to be different from their 'willingness to pay' towards removal of some unwanted feature of their environment.  If the two values are different, then it is not always clear which should be used to inform policy.

**mva**consultancy

that the "true" noise element of property price differences would understate average population willingness-to-pay because more noise-tolerant purchasers are likely to bid for properties in noisier areas and require less "discount" on prices to attract them.

3.2.4    In application also, hedonic pricing has limitations.  By its nature, it relates to the "total" noise experience at properties in different areas and is not able to distinguish components of this "total".  Arising from this, for the present study two particular constraints of hedonic pricing would be that:

- it cannot provide monetary valuations for variations in the temporal distribution of noise, such as changes in night-time exposure; and

- it offers no practical opportunity to elicit separate weights on the numbers of aircraft movements in the vicinity of a property and the noisiness per movement.

3.2.5    The appeal of such an approach, however, is that the monetary value of aircraft noise nuisance is deduced from the differences between property prices in areas of higher and lower noise - that "reveal" the value apparently placed by house-buyers – their willingness-to-pay – upon avoiding aircraft noise.

3.2.6    For residential areas around airports it is unclear to what extent the results might have been influenced by accessibility and associated employment opportunities and this affects the perceived reliability of the results.  In addition, there are theoretical concerns about possible differences in noise sensitivity between persons active in housing markets in noisy as opposed to quiet areas.

3.2.7    A major UK study by **Bateman at al (2004)**[10], subsequently reviewed by **Nellthorpe et al** in 2005, used hedonic pricing methods to estimate how much households in Birmingham are willing to pay to avoid noise pollution from road, rail and aircraft noise respectively.

3.2.8    Data on more than 10,000 house sale prices in the Birmingham area were analysed, together with detailed information on:

- the characteristics of the individual properties;

- their local area;

- the noise environment in terms of exposure to the three transportation noise sources of interest; and

- other environmental indicators, and the socio-demographic composition of the area.

3.2.9    The study successfully estimated models for road and rail noise, but was less successful in the case of aircraft noise.  The authors suggest that this results from a problem observed in previous HP studies, in that aircraft noise does not vary significantly within local sub-markets, where house prices should be most comparable in other respects.

3.2.10   Non-noise factors which were found to influence the valuations placed on changes in noise exposure included ethnicity, age and family composition.

---

[10] "The valuation of transport-related noise in Birmingham: non-technical report to the DfT", Bateman, Day and Lake, Transportation Research Part D 9 (2004)

mvaconsultancy

3.2.11    In their review of this research for the Department for Transport, **Nellthorpe at el** benchmarked the results against other European research, and considered how results from Birmingham might be transferred between locations in the UK and over time.  Bateman et al's analysis was extended to cover the social rented sector, and to sound  levels down to 45 dB.

3.2.12    The review noted that Bateman et al's results confirmed previous evidence that the marginal value of noise is related to the existing sound level, and that valuations appear to vary by time of day, and to be correlated with reported annoyance.

3.2.13    Comparison of noise:annoyance and noise:willingness to pay relationships indicate that the latter is almost linear, while the former rises sharply over part of the sound level  range, and then levels off.

## 3.3    Hypothetical Methods

3.3.1    While hedonic pricing ultimately relies on the vagaries of actual market prices that are presumed to be related in part to differences in noise levels, hypothetical market techniques can seek willingness-to-pay directly for noise, by setting up markets in which survey respondents indicate whether – in effect – they would be prepared to "purchase" defined levels of noise exposure at the prices offered.

### Contingent Valuation

3.3.2    CVM has been widely used, having established a particularly strong following within the environmental research community, for valuing environmental and amenity factors.  This contrasts with transport research where, after early encounters with CVM, what we here describe as SP techniques have dominated, to place values on (for example) time spent in different travel circumstances, and on aspects of improvement to transport services.

3.3.3    CVM could be construed as an ideal method for valuation, since it aims to obtain a **direct** assessment of the survey respondent's valuation of an improvement of interest (such as reduction in noise), in terms of willingness to pay (WtP) or willingness to accept (WtA).  It attempts precisely to obtain the point on the indifference curve corresponding to the specific improvement presented to the respondent.

3.3.4    In CVM, respondents can be asked (for example) to state how much money they would be prepared to pay for removal of all aircraft noise from their place of residence. By using this method, it is possible:

- to obtain equivalent valuations from respondents who do not object to the principle of making a financial contribution towards the removal of an unwanted feature of the environment;

- to investigate any relationship between willingness to pay and disposable income; and

- to identify individuals who might be able to pay but who would refuse to do so as a matter of principle.

3.3.5    The main concern relating to CVM is in connection with the exact formulation of the question and, by implication, the respondents' ability to "locate themselves on the true indifference curve".

**mva**consultancy

3.3.6    A historically-popular means of eliciting the required information has been simply to ask questions of the form:

> *"What is the most you would be willing to pay to gain the [specified] improvement?"*

3.3.7    Such "open-ended" questions have now fallen out of favour, and more refined versions are being widely used. However, there remains a concern that respondents might be tempted by the scenario presented to them to respond tactically.  For example, a recently-reported CVM exercise (Wardman, M and A L Bristow, "Traffic related noise and air quality valuations: evidence from stated preference residential choice models", *Transportation Research Part D*, vol 9, 2004, pp1-27) found that, of about 400 respondents asked how much more they would be prepared to pay for a specified reduction in traffic noise, nearly two-thirds responded "zero".

3.3.8    When asked further why they had responded thus, about half of these stated that they were not prepared to pay more Council Tax (this being the medium in which, hypothetically, the money amounts were couched).  It is not at all clear whether all, some or none of the "zero" responses should be excluded to remove bias in estimating the population or community valuation of noise, yet the impact on the result would be substantial.

3.3.9    In spite of such questions, CVM has become widely accepted as an appropriate means of valuing environmental "goods", and a considerable body of experimental work has been aimed at improving its reliability.

### Stated Preference

3.3.10    The standard approach in SP is to offer a series of pair-wise comparisons between combinations of attributes (which may or may not include price).  The preferred combination is assumed to have the higher utility, and the aim of the **analysis** of the exercise is to devise a formula for utility which explains the choices as far as possible.   Note that the presentational problems associated with the notion of indifference are explicitly avoided: respondents are to choose the preferred alternative.

3.3.11    However, since indifference is ultimately the required concept, the data has to be sufficiently rich to allow the tradeoffs to be estimated with confidence.  This has consequences both for the amount of data required **and** the design of the options to be traded-off.

3.3.12    The use of hypothetical scenarios implicit in SP methods offers considerable benefits from the point of view of controlling for correlations between independent variables, but is also open to criticism because respondents need not be committed in any way to the choices they make.  The use of hypothetical scenarios allows for a wide range of detailed variations to be tested, something that is not possible using hedonic price methods.  In addition, the sample can be drawn from the entire population resident within defined common noise areas, rather than just being limited to the proportion active in the housing market at the time.

**mva**consultancy

3.3.13   Any concerns that stated preferences are essentially hypothetical and hence might not reflect actual behaviours can be mitigated by making sure that:

- respondents are given as much information as possible in order to be able to make informed choices;

- that all hypothetical scenarios as actually presented are essentially 'believable' within the context of what those particular respondents are used to in everyday life; and

- that respondents are qualified by their everyday experience to be able to make meaningful or properly informed choices within the context of the alternative scenarios presented.

3.3.14   The detailed methods adopted in the ANASE study were developed and tested through an extended series of pilot studies during which it was confirmed that statistically consistent and apparently plausible results could be obtained using this method.

**mva**consultancy

# 4 Phase 1 and Survey Design

## 4.1 Overview

4.1.1 A comprehensive review of dose/response relationships from a wide range of research, including the "baseline" **ANIS study,** indicated the following key issues for ANASE:

- re-assessing the relative weights on sound level and number of aircraft movements;

- identification of thresholds and non-linearity within the sound level : annoyance relationship;

- choice of indicators of annoyance: eg mean annoyance scores, % above certain annoyance thresholds;

- existence of confounding factors such as links to the airport, or double glazing; and

- identification of any variations in annoyance by time of day, or related to the mix of aircraft types operating.

4.1.2 To investigate these, and other issues pertinent to the study objectives, the project team sought to develop a quantitative survey approach that combined internationally-recognised questions relating to community annoyance with aircraft noise, with willingness to pay valuations derived from SP questions, in the context of a rigorously designed social survey of residents close to airports. However, to do so successfully required an initial phase of exploratory research.

4.1.3 This chapter describes the elements investigated in this initial phase, Phase 1, and how they affected the main survey design of Phase 2.

## 4.2 Preliminary Investigations

### Preliminary Qualitative Exercise

4.2.1 A **preliminary qualitative** exercise was carried out with a small sample of 28 respondents, all of whom lived close to an airport. This sought to establish the "discriminable factors" that influence community annoyance, through in-depth interviewing of samples of residents close to four airports of varying operational characteristics.

4.2.2 This initial qualitative research identified that the main dimensions contributing to annoyance were: aircraft type, number, and time of day or night; and identified the discriminable differences in levels for each dimension.

4.2.3 A second part of the preliminary qualitative exercise involved the presentation of SP-like trade-off options. These also were well received and understood by almost all respondents, indicating:

- sufficient consistency in respondents' terminology to enable generic 'aircraft types' to be defined that could be generally understood and traded-off;

- respondents' ability to identify with, and respond to, different hypothetical scenarios depicting different patterns of individual aircraft types flying overhead at different times of the day;

**mva**consultancy

- respondents' ability to trade-off between different patterns of different types of aircraft, and hence aircraft noise, provided that the presented alternatives are realistic and have noticeably different levels; and

- respondents' willingness, as well as ability, to understand and respond rationally to the idea of monetary compensation and trade-off against different amounts of aircraft noise.

4.2.4    These cognitive findings confirmed the main dimensions of aircraft noise perceptions, and that the number of different levels of each of these variables, in terms of aircraft type, time period, and money levels, that respondents were able to discriminate between was probably limited to three or four levels in each variable.  Thus, with due attention to possible confounding factors, the preliminary qualitative exercise found sufficient consistency regarding aircraft noise dimensions and levels to enable meaningful hypothetical scenarios to be devised that could form the basis of effective SP surveys and analysis.

### Impact of Noise at Night

4.2.5    Residents in the vicinity of one of the four airports used in the Preliminary Qualitative Research Phase appeared to be much more sensitive than residents near the other airports to aircraft noise at night.  The airport concerned – Nottingham East Midlands – has an unusually high proportion of its movements at night.  This prompted a comparative investigation of night-noise sensitivity against residents close to a fifth airport – Gatwick – that also has substantial night-time activity: more in absolute terms than Nottingham East Midlands, though less on a proportional basis.

4.2.6    Comparisons of views expressed by a sample of residents near Gatwick with those obtained from the Nottingham East Midlands Airport sample suggested that differences in attitudes appeared to be related to specific features of Nottingham East Midlands operations – notably cargo movements by heavier, older and therefore noisier aircraft – that are concentrated at night.  However, such operations are not intrinsically a night-time activity, so it could not be concluded that there is necessarily greater sensitivity to aircraft noise at night compared to during the day, on a like-for-like basis.

4.2.7    This finding did not detract from the merit of seeking to establish whether there is greater sensitivity to aircraft noise at night, though it obviously indicated that such variation in sensitivity could not be assumed *a priori*.  The methods developed in Phase 1 to estimate willingness-to-pay values of aircraft noise continued to allow for time-of-day being a possibly-important source of variation in these values.  Such analysis would obviously depend upon whether respondents could think separately about noise at different periods of the day.  The Nottingham East Midlands and Gatwick comparison indicated that they could.

### Measurement of Personal Aircraft Noise Exposure

4.2.8    In parallel with the preliminary qualitative exercise, an Acoustic Measurement Pilot was carried out at the homes of some respondents.  This measured sound levels outside the properties (as is conventional), indoors, and "at the respondent's ear" using personal dosemeters.  The strength of relationships between the sound level data and standard air noise contours, and also with the passage of aircraft as evidenced from control tower records, was then explored.

**mva**consultancy

4.2.9     For the near-airport sites surveyed, the degree of correspondence found between the outdoor acoustic measurements and published aircraft noise contour information was impressive, and very encouraging for the general approach being contemplated for the design of the Phase 2 main study, since that was anticipated to be based mostly on calculated aircraft sound levels (in the manner used to map aircraft sound level contours) with sample measurements for calibration and validation purposes only.

4.2.10    Semi-automatic methods developed for identifying specific aircraft noise events from the continuous sound level monitoring data were found to be very reliable when compared against appropriate air traffic control tower log data.  This was also a very encouraging finding, and supported the proposed methodology for assigning aggregated aircraft sound level values to each hypothetical scenario tested as part of the anticipated SP exercises.

4.2.11    The study was also successful in using the control tower log data to associate acoustic events with individual aircraft types, from which fairly wide ranges of event sound levels emerged for individual aircraft types, rather than clear-cut divisions into bands of sound level.  This needed to be taken into account in designing the alternative mixes of aircraft types to be presented in the SP exercises.

4.2.12    The close correspondence between aircraft movement data and noise events identified from outdoor monitoring, and between measured and calculated sound levels, meant that outdoor logging could generally be relied upon to provide an accurate record of the external noise burden imposed by aircraft upon properties.  However, the test sites were relatively close to airports: a key issue at greater distances from airports would be how to set the threshold for triggering the semi-automatic event identification system at the appropriate sound level for each measurement site.

4.2.13    There was effectively no correlation between the sound level patterns recorded by the personal dosemeters and those obtained from the outdoor monitors.  Dosemeter logs are likely to be dominated by non-aircraft sound sources, especially from an individual's own activities; and there would of course be no correlation at all when the individual is away from home.

4.2.14    There thus appeared to be no justification for widespread aircraft sound level measurement using either fixed indoor monitors or personal dosemeters in the Phase 2 main study.  On the other hand, some outdoor monitoring was judged to be necessary as a basis against which to provide empirical validation of calculated outdoor sound levels for some lower sound level sites included in the Phase 2 main study.

### Investigating Annoyance Levels in Low Aircraft Noise Areas

4.2.15    Following ANIS in the 1980s, 57 dBA LAeq had been established as the threshold for the onset of significant community annoyance from aircraft noise.  It was felt that attitudes to aircraft noise amongst residents in areas below 57dBA should be investigated within the main Phase 2 of the study and a minimum cut-off (below which all/almost all are unaffected by aircraft noise) would be valuable information to inform the Phase 2 survey design.

4.2.16    Exploratory research was carried out in six localities outside Heathrow's 57 dBA LAeq contour so that an initial assessment of annoyance from aircraft noise in some "low aircraft noise" areas could be made.  The survey areas were overflown at heights ranging from 4000 to 12000 feet, with about 50 residents being interviewed in each.  The sites chosen were well

mvaconsultancy

outside the 57 LAeq threshold and included areas in close proximity to arrivals 'stacks' (where aircraft circle awaiting approval to land). None reported "extreme annoyance" from aircraft noise, and only 4% expressed themselves "very annoyed".

4.2.17    In the two survey localities close to stacks, the degree of annoyance was relatively greater than in other areas where aircraft typically are lower and therefore could be expected to be more noisy. The outcome thus suggested that there may be some additional factors associated with stacks *per se* that are a source of annoyance independently of sound levels. In-depth qualitative interviewing would lead to a better understanding of what is giving rise to annoyance in such locations, but sound level measurements could not be expected to provide any useful insights, given the low levels of aircraft noise experienced. In turn, this would preclude the use of SP techniques in these areas for the purpose of valuing annoyance arising from aircraft noise.

4.2.18    A further consideration when conducting research in low noise areas is that modelling aircraft sound levels below 57 dBA LeAq becomes increasingly inaccurate as the sound level decreases primarily because of variability in atmospheric conditions along the propagation path. Indeed, sound level contours are not routinely produced below this level because of the uncertainties. For Phase 2 it was agreed that reported community annoyance with aircraft noise should be obtained at low noise areas (including some sites below 50 dBA), but that the uncertainty associated with modelled LAeq and other sound level metrics at such sites would need to be borne in mind when drawing conclusions about the relationship between community annoyance and sound levels in low noise areas.

### 4.3    The SP Pilots

4.3.1     While piloting is standard good practice for any SP study, it was accepted that it would need to be much more extensive than usual in this case, because of the particular difficulties of presentation. In essence, the central issue was: how should alternative aircraft-noise scenarios at different times of day be conveyed to survey respondents so that their stated preferences could be construed in unambiguous relation to the noise environment that the scenarios were intended to represent?

4.3.2    Moreover, with a primary objective of the study being to estimate separately the weights put upon the noisiness of individual aircraft movements and the number of movements, it was important that the SP scenarios should be able to convey variations in these two dimensions of "overall aircraft noise" to respondents. As was described above, a target of the Preliminary Qualitative Research Phase was to assist in identifying how ordinary members of the public articulate the features of "overall aircraft noise" that annoy them.

4.3.3    Frequency and time-of-day were two prominent aspects, and variations in these pose no difficulty with presentation to respondents. The real challenge, however, was to present different levels of noisiness per movement, and here the findings of the Preliminary Qualitative Research Phase were especially valuable, as a result of the aircraft-type labels that respondents themselves associated with different levels of noisiness.

4.3.4    This provided the basis for a possible format of SP choice scenarios. It would entail sampling respondents in the noise "catchment" of particular airports, and using aircraft-type labels of the kind used by respondents in the Preliminary Qualitative Research to distinguish between different levels of noisiness per movement. In the SP scenarios, three or four labels would

**mva**consultancy

represent appreciably different levels of noisiness per movement.  The scenarios would then differ through the number of movements assigned to each label for a specified time of day – and hence to the different levels of noisiness per movement implied by the labels.

4.3.5    Respondents could then express their preferences for scenarios presented, for example, in pair-wise comparison.  Statistical analysis of the preferences, using standard principles of Discrete Choice Analysis, would estimate the time-period-specific weights that respondents were implicitly placing on movements of different aircraft types when expressing their preferences.  Each weight would represent the "disutility" to respondents of an additional movement of the aircraft type and in the time period concerned, this "disutility" arising from the different levels of noisiness that the aircraft types represented.

4.3.6    An elaboration to permit the relative disutility weights to be given monetary value would then be to incorporate a money variable in the SP options.  The weight estimated on the money variable would be in the same "utility" units as the aircraft-type weights, enabling the disutility of a movement of each aircraft type to be equivalenced to the amount of money that would generate the same disutility.  Thus the disutility of an additional aircraft movement could be expressed in money terms.

4.3.7    "Vehicles" for describing the money amounts had also been also tested in the Preliminary Qualitative Research Phase, where rebating of Council Tax appeared to be an understood and acceptable format for the money attribute (though see paragraph 4.3.17 below).

4.3.8    Thus each SP scenario might be characterised by the numbers of movements of different aircraft types at different times of day, and the money attribute.  This was seen, however, to be introducing a potentially large number of attributes for respondents to trade between. For example, only three aircraft types for each of, say, six periods, plus the money attribute, would imply 19 attributes for each scenario.  Given the already novel context of the SP application, the risk of overwhelming respondents with this large number of attributes was regarded as too great.

4.3.9    It was therefore decided that separate SP exercises would be conducted for individual time periods.  Respondents would not be required to trade between combinations of aircraft types and time periods, and hence the scenarios for each time period could potentially have more aircraft types and thus be more discriminating between different levels of noisiness per movement.  The relativity of weights on noisiness at different times of day would still be available through the money values estimated for each aircraft type in each time period.

4.3.10   Another development was to present respondents with three-way rather than pair-wise comparisons.  They would be asked for their most and least preferred, thereby obtaining the equivalent of two pair-wise choices more economically than from actually presenting two separate pair-wise comparisons.

4.3.11   A first SP pilot was carried out with respondents living in the vicinity of Heathrow, and a second pilot, building on the promising experience of the first, with residents around Nottingham East Midlands Airport.  Both were exploratory of the proposed SP format, and thus involved samples of only about 30 residents each.

4.3.12   The details of each pilot are provided in Appendix A4, and the main findings from the two pilots were:

**mva**consultancy

- the majority of respondents understood the SP exercises, participated strongly and provided rational responses;

- the difficulties that were encountered appeared to be primarily presentational, and a number of areas for presentational improvements were subsequently adopted; and

- the estimated SP models bore out the rationality of responses and respondents' systematic choice behaviour in trading between the aircraft types, the estimated weights all having the expected negative signs (to reflect disutility of marginal increases in aircraft or council tax), and showing high correlation with the noisiness of different aircraft types.

4.3.13    Crucially, there was no indication that fundamental changes in the SP design were needed, though some modifications were seen to be desirable, notably that the number of movements for each type of aircraft should be more in keeping with the actual numbers of movements in each period.  It was also agreed that a common length of period should also be adopted.  Overall, however, the two pilots indicated that the main thrust of the proposed SP approach was workable and merited more intensive consideration.

### Presentation Pilots

4.3.14    There remained two particular issues, concerning the form of presentation of the SP scenarios to respondents, and their "cognition" of the SP choice tasks that they were being asked to perform.  "Presentation Pilots" were therefore conducted to explore the efficacy and cognition of a number of alternative ways of presenting aircraft noise options within a SP trade-off environment.

4.3.15    Cognitive investigation affirmed a high level of understanding by respondents of the essence of the choices they were being invited to make, namely trade-offs between alternatives made up of different mixes of sound levels and/or money amounts.  As they made choices they were able to explain their rationales, which clearly demonstrated engagement with the trade-off opportunities provided by the presented alternatives.

4.3.16    A small minority of respondents in the Presentation Pilots, however, were concerned that options which offered money awards could be taken to imply that they would in practice accept a higher level of noise than they currently experience.  By including money awards in **all** SP options, including those for which respondents perceived a lower-than-current noise content, any appearance of "buying off" respondents to accept higher sound levels would be avoided.

4.3.17    Different "vehicles" in which the money variable could be set were investigated.  Though the concept of Council Tax rebates used in the first two SP pilots worked well for many respondents, for some there were connotations relating to how the rebates could effectively be negated by local council action.  An alternative vehicle in the form of an annual (or monthly) grant seemed to have no negative connotations, and was adopted in place of the Council Tax rebates.

4.3.18    During the Presentation Pilots, several other presentational approaches were tested and found to be unsatisfactory because they could not be fully grasped or were even disliked by a relatively high proportion of respondents.  These approaches were: showcards with pictograms to represent the number of aircraft movements; depicting the numbers of movements in bar chart form; expressing different aircraft noise alternatives in terms of

**mva**consultancy

decibels; and associating different aircraft noise alternatives with "event characteristics" (such as vibration).

4.3.19   Respondents needed support from visual and, especially, aural stimuli to form an adequate strength of association between aircraft type descriptors and sound levels.  Testing various presentational approaches suggested that the structure of choices used in the earlier SP pilots (generic aircraft type-noise categories, numbers of noise events for a specified period of the day, and money variables) remained viable, but would benefit from clarification of the meaning of the aircraft type descriptors, especially the sound levels to be associated with them.

4.3.20   A major enhancement resulting from the Presentation Pilots was that each respondent should be exposed to customised presentation material in both aural and visual forms. Customisation would be achieved by recording and photographing examples of the relevant aircraft types close to respondents' homes.  At interview, before embarking on the formal SP exercises, respondents would be played and shown this material and asked to confirm that it did reasonably represent what they would hear and see outdoors at their homes.  When this was tested, most respondents had a clear opinion of whether the level of sound levels played to them corresponded to the aircraft being shown to them, and there was strong evidence that the more customised the aural and visual material was to their home localities, the more likely respondents were to feel that it represented their experience.

4.3.21   Tests of the customised material took place in local venues rather than respondents' homes to permit control and adjustment of the presentation of (in particular) the aural playbacks. This also avoided the intrusion imposed upon respondents of installing, calibrating and dismantling the required bulky equipment in their homes, and the risk of jeopardising quality unless this process were overseen by a professional acoustician, potentially increasing costs significantly.

4.3.22    Nevertheless, it was recognised that in the main study, with its much greater emphasis on ascertaining attitudes to noise from a rigorously-selected, probabilistic sample, the advantages of venue-based interviewing would need to be weighed against the risk that recruitment would be more difficult when travel to a local venue was required.  In fact, more extensive venue-based interviewing in a subsequent pilot found that almost all respondents were content to accept the customised playbacks as good representations of the aircraft noise levels they experienced at home.  This allowed playbacks to be adapted "on the spot" to be dispensed with, reducing the extent of equipment required and the need for its supervision by professional acousticians.

4.3.23    A further pilot was undertaken to investigate the merit of a complementary SP format, in which respondents would be invited to choose between different distributions of aircraft movements spread over 24 hours, presented in histogram form.  This format had emerged as potentially promising during the Presentation Pilots.  If successful, it could provide valuations of annoyance by time of day more directly than the SP format discussed hitherto. (The histogram approach would give only **relative** valuations for one time period compared with another; it would not provide **absolute** monetary valuations of aircraft noise.)

4.3.24    Some 185 respondents from ten localities in different aircraft noise environments in a broadly-defined "catchment" of Heathrow were recruited for the pilot.  Coupled with the formal SP exercise there was considerable cognitive exploration to assess respondents'

understanding of the SP histogram approach and to seek to appreciate how respondents were tackling the SP choices and hence arriving at their preferences.

4.3.25   The conclusion of the research team was that the histogram approach in its current form could not be recommended. The SP review group concurred that the design as presented was too complicated to be progressed to a main study, and that extensive development to make the histogram SP choice tasks "accessible" to a much higher proportion of respondents should also not proceed.  Therefore, our planned approach for Phase 2 reverted back to that originally proposed (and described in Section 4.3) to investigate variation in annoyance by time of day indirectly, presenting each SP exercise in the context of a single time period.  Six equal (four-hour) periods were adopted, as shown in Table 4.1, that broadly coincide with the existing periods of the day for deriving the different aircraft metrics, and reflected people's main breakdown of the day (i.e. early morning, late morning, early afternoon, and so on).

### Table 4.1 Time Periods for the SP

| Time Period | Duration |
|---|---|
| Period 1 | 0300-0700 |
| Period 2 | 0700-1100 |
| Period 3 | 1100-1500 |
| Period 4 | 1500-1900 |
| Period 5 | 1900-2300 |
| Period 6 | 2300-0300 |

**"Numbers-Gaming"**

4.3.26   Despite the generally positive outcomes, there remained some lingering doubts that could not be fully explored in the Presentation Pilots: these related to whether the willingness-to-pay valuation of annoyance estimated from SP responses might be influenced by the numbers presented in the SP scenarios, rather than (as should ideally be the case) be independent of them.  This possibility was termed "numbers-gaming".  This might arise if respondents were having difficulty in carrying out the SP tasks and were resorting to some simplifying rule to provide a "choice", though this might not be their "genuine" preference if they could be faced with the options in real life.  Clearly, if this **were** the typical SP response behaviour, it would not be appropriate to rely on the SP results as the required estimates of disutility.

4.3.27   The research team designed and carried out a major SP pilot specifically to test for the "numbers-gaming" phenomenon.  The pilot had to be devised in such a way that there would be a reasonable prospect of distinguishing whether respondents' SP choice behaviour was consistent with "genuine" ("conventional") trading, or with "numbers-gaming", but not both.  This was an extremely demanding requirement, backed up by stringent "acceptance criteria",

**mva**consultancy

and, so far as the research team is aware, no comparable investigation has previously been attempted.

4.3.28   The form of SP was essentially that of the first two pilots, with the enhancements identified during the initial Presentation Pilots, notably venue-based interviews to allow for customised aural and visual presentation of the aircraft types.  Special SP designs were developed that would provided a reliable basis for objectively testing the two different hypotheses.  Details of the design and analysis of the "numbers gaming" pilot is provided in Appendix A4.

4.3.29   The pilot findings provided the evidence[11] to reject the numbers-gaming hypothesis while remaining consistent with the conventional hypothesis.  The SP review group (see Chapter 1) agreed that the conventional hypothesis was more plausible, and that there was no further need to consider the form of SP model, which could therefore be taken forward as the primary SP form for a main Phase 2 study.

4.3.30   It was noted that the weights on the money variable were not as well-defined as those on the aircraft types.  The implication for a main study was that trade-offs between movements of different aircraft types at different times of day would be reasonably estimated, but that it might well be appropriate to support the SP estimates by reference to other money valuations, such as those from past "hedonic pricing" studies.  The already-accepted status of hedonic pricing estimates would in any case almost certainly dictate that the SP results were compared with these, with "best estimates" being obtained by taking both sources of information into account.

## 4.4   CVM Question

4.4.1   As an exercise completely distinct from investigating the merits of a possible time-of-day SP approach, the research team was requested to trial a CVM approach to estimating the benefit in monetary terms of eliminating all aircraft noise.  Respondents were asked to imagine that there was a house exactly like the one that they lived in, including all the same local facilities and access to work, shops, etc. The only difference was that no aircraft noise was audible. This imaginary house however, would cost £10 per week (roughly £500 per year) more to live in than their current homes. Respondents were asked to choose whether they would prefer to:

■   Get rid of all aircraft noise, but therefore have £500 less per year to spend on other things; or

■   Put up with the current level of aircraft noise at home, so that they retained the £500 per year to spend on other things.

4.4.2    At this point, respondents who had indicated preparedness to forgo £500 per year to eliminate aircraft noise were asked how much more they would give up, thus indicating their maximum amount.  Respondents who were not prepared to give up £500 per year were also asked for their maximum below this amount. Their showcard advanced in £50 steps from £0 to £450 per year.

4.4.3   It was not straightforward to interpret the outcome.  The indication was that, once dichotomised by the £500 threshold, many respondents were not prepared to shift from the

---

[11] A more detailed description of the results in given in Appendix A11.2

**mva**consultancy

minimum for their segment: zero or £500.  For those that did offer to forgo more, the distribution was by no means smooth, with a substantial outlying proportion in the "at least £500" segment offering £2500 or more.

4.4.4    A particular question is: to what extent is the incidence of zeros a genuine expression of the nuisance caused by aircraft noise?  Is there a "protest vote" embedded to the effect that peace-and-quiet is a right that one should not have to pay for?  For example, of those "extremely annoyed" still nearly 20% were not prepared to forgo any money to eliminate aircraft noise.

4.4.5    At the other extreme, how should the inability or unwillingness to put an upper bound on the maximum amount be understood?  Did it indicate total exasperation with aircraft noise?  Were these respondents taking the opportunity (when the actual impact on their disposable income would of course be nil!) to put relief from aircraft noise beyond price?  Might "zero" and "no upper bound" responses in fact be registering the same sort of reaction?

4.4.6    For those respondents who gave more "moderate" replies, there remains the issue of how sensitive their stated maxima would be to different aircraft noise scenarios, so long as some appreciable reduction in noise was involved.  There were no "checks and balances" built into the approach to test whether, for example, removing only half of aircraft movements, or just those at night, would have elicited significantly different responses.

4.4.7    A conclusion from the foregoing might have been that achieving a satisfactory degree of confidence in the "no aircraft noise" experiment would require considerable further development.  The SP review group felt, however, that a CVM question could be included in a subsequent main study at minimum risk and effort, on the basis that, if it performed reasonably, useful complementarity to SP results would arise, and little would be lost otherwise.

4.4.8    The review group considered that it would be of particular value to investigate the motivation of interviewees who gave zero or extreme responses to the CVM question.  This would, for example, identify "protest" respondents.  It might then be appropriate to analyse their **SP** responses separately from those of other respondents, to check whether differences in respondents' "outlook" impacted on the inferences that might be drawn from the SP-estimated trade-offs and valuations.

## 4.5    The Standard Reported Annoyance Question

4.5.1     In the same interviews as the "time-of-day" Presentational Pilot, cognitive investigation was undertaken of the standard community annoyance question concerning aircraft noise.  This was the central question in the 1980s ANIS survey, and the modern equivalent International Standard question was to be asked in the Phase 2 main survey.  The objective was to understand better the basis of response to this question so that it could be appropriately employed to (re)assess attitudes towards aircraft noise in Phase 2 in a meaningful and policy-relevant manner: in particular, allowing "backward compatibility" with the ANIS.

4.5.2    The need for cognitive investigation had arisen from the Presentation Pilots.  It was found that the majority of respondents (albeit a small sample) stated, when probed, that they were not confident that they fully understood precisely what was meant by the question.  When respondents were asked to consider particular times of day or seasons, rather than the

**mva**consultancy

whole year as a piece, their reported levels of annoyance often varied, and in reporting their annoyance, respondents might be thinking in "average" terms or – as several explicitly stated – of a "worst case" experience.

4.5.3    The wording of the current ISO standard reported annoyance question is:

*"Thinking about the last 12 months or so, when you are here at home, how much does noise from aircraft bother, disturb, or annoy you?"*

with the following possible responses:

*"Not at all, Slightly, Moderately, Very, Extremely"*

4.5.4    The approach adopted was to ask the standard question and then to probe respondents' views on factors that might cause them to review their initial response to the standard question. The factors covered the sound level of individual flybys, the number of flybys, times of day, weather effects, and runway operation (alternation) at Heathrow.

4.5.5    Of 185 respondents, only one subsequently changed response to the standard question, suggesting that it does indeed provide a robust measure of reported annoyance. For a subsequent main study, there thus appeared to be no reason not to employ the ISO standard annoyance question, and there appeared to be no need to repeat the investigative "challenge" to respondents' initial reports of their levels of annoyance.

4.5.6    In addition, part of our Phase 1 cognitive testing of the standard noise annoyance questions included a set of preceding questions that explored with the respondent the pros and cons of living in the neighbourhood (and excluded any sound level measurement or reproduction equipment), thereby replicating the ANIS-style of context-setting. In piloting, we found similar responses to the standard annoyance rating question with, and without, preceding general neighbourhood questions.

## 4.6    Phase 1 Conclusions

4.6.1    The key strategic outcomes of Phase 1 were as follows.

- SP could be employed to inform on the objectives of the study, namely to:
  - establish the relative weights of "sound level" and "number" in their contribution to annoyance,
  - examine variation in annoyance by time of day; check for varying annoyance levels between different person-types, and
  - estimate money (willingness-to-pay) valuations of annoyance in all these dimensions (possibly with the support of monetary valuations of aircraft noise annoyance from other sources, notably hedonic pricing);
- a CVM test would be included in interviews. This was on the request of the reviewers and would help to identify respondents who might be attempting to "protest vote", so that, if necessary, their SP response data could be separately analysed; and

**mva**consultancy

- the ISO reported annoyance questions would elicit robust qualitative indications of respondents' levels of annoyance and provide some backwards compatibly with the corresponding question in the 1980s study (ANIS).

**mva**consultancy

# 5     Sample Design

## 5.1    Introduction

5.1.1    Following on from Phase 1, it was necessary to finalise and test the Phase 2 sampling strategy, in order to meet the twin objectives of sampling a range of different noise environments, and achieving a proper representation of the population in scope.

5.1.2    The methodology needed to identify aircraft noise exposed areas within England which are suitable for the study, both in terms of identifying the airports themselves and then defining the spatial envelope surrounding them, outside of which it can be assumed that aircraft noise is only faintly audible.

5.1.3    The methodology was finalised and tested in advance of the final Phase 2 surveys, to ensure that the method was both viable and robust. The final method is discussed in this chapter.

5.1.4    In parallel, a full 'dress rehearsal' of sampling procedures was conducted at two sites, including a final assessment of the relative advantages and disadvantages of conducting the SP exercise in respondents' homes or in local venues. Details of the fieldwork procedures adopted for the national survey are given in Chapter 6.

## 5.2    In-scope Population for Phase 2 National Survey

5.2.1    To identify those in scope for the survey, it was a requirement to identify aircraft noise exposed areas within England, and, within these areas, to stratify the population according to the characteristics of the aircraft noise. For this purpose we needed to classify each area within a 'matrix' dimensioned by event sound level (L) and number of movements (N). Then, to ensure that, within each stratum, all residents of every candidate area have the same probability of selection, a stratified random sample of areas was drawn. It was necessary to sample by strata so that the survey was undertaken in all noise environments in scope.

5.2.2    It was agreed with the DfT that the population in scope for the survey should be limited to residents of Census Output Areas (OAs) potentially affected by noise from the 20 largest commercial airports.

5.2.3    As described in Chapter 2, around each airport the 'potentially affected' population was defined as being within the 65 LAmax footprint of the noisiest aircraft operating out of the airport concerned on any flight path. This gave a total of 11,246 OAs.

5.2.4    Airports were identified as having irregular or regular traffic. As the full range of sound metrics are not available for airports with irregular traffic, then for the SP survey to work effectively, the full survey could only be undertaken at sites close to airports with regular traffic.

5.2.5    In addition, we did not think that it was feasible to carry out the SP survey when the Nav level was below 120. However, we did not wish to exclude these sites, nor those with irregular traffic, and so therefore proposed to do a "restricted" survey which would exclude the SP questions. Within the 'noise affected envelope', areas could either be available for

**mva**consultancy

selection for the full (SP) survey or the restricted survey (not including an SP element).  The criteria for this classification were as follows:

■    full survey site: a minimum threshold was defined for SP (and, hence, the full survey) of 120 aircraft above 65dB (on a 30 day month, 16 hour day basis), equating to an average of one movement per four hour period per day; and

■    restricted survey site: restricted surveys were conducted in areas below this threshold, and areas around airports for which sound level data were unavailable, or have irregular air traffic.

5.2.6    There were 6,903 OAs satisfying the "full" criteria, with the remaining 4,343 classified as "restricted".

5.2.7    We aimed to complete 60 interviews in the full survey sites, and 15 interviews in the restricted survey sites.  60 interviews were required to provide enough data for the SP to provide robust parameters (by aircraft type and time of day) at each of 36 sites.  Fifteen responses was the target number at each site for the non-SP questions as this enabled a wide range of different sites to be covered (ie 40 sites of 15 respondents was preferred over 10 sites of 60 respondents).

5.2.8    To classify the in-scope areas into the relevant (L, N) categories required the application of the acoustics methodology described in Chapter 2.

5.2.9    Note that the objective of the procedure used to draw a sample of sites was to obtain a fully representative selection of areas with differing aircraft noise environments.  We were not interested in selecting sites from individual airports per se: it was the characteristics of the aircraft noise environment that was of interest.

### Sample Stratification

5.2.10    Table 5.1 summarises the airports in scope for the survey, along with whether they have regular or irregular traffic, and the number of output areas within the 65 LAmax footprint.

**mva**consultancy

**Table 5.1 Airports in Scope with Regular and Irregular Traffic**

| Airport Name | Regular Traffic (available for Full Survey) | OAs within 65 LAMax Contour | Airport Name | Regular Traffic (available for Full Survey) | OAs within 65 LAMax Contour |
|---|---|---|---|---|---|
| Birmingham | Y | 345 | Leeds Bradford | Y | 469 |
| Bournemouth | N | 141 | Liverpool | N | 333 |
| Bristol | N | 84 | London City | Y | 883 |
| Coventry | N | 131 | Luton | Y | 103 |
| East Midlands | N | 51 | Manchester | Y | 1,469 |
| Exeter | Y | 279 | Newcastle | N | 162 |
| Gatwick | Y | 255 | Norwich | N | 133 |
| Heathrow | Y | 5,440 | Southampton | Y | 438 |
| Humberside | N | 12 | Stansted | Y | 162 |
| Kent International | N | 130 | Teesside | N | 226 |

**Selecting Areas for the Full Survey**

5.2.11    The above table shows that there were ten airports in scope for the full survey.

5.2.12    Modelled data were available for each in-scope OA on both the sound levels of aircraft flying overhead (L), and the number of aircraft noise events (N).  Average sound level values and numbers of events (Lav and Nav respectively) for each OA were estimated for the population-weighted centroid of the OA.  Each OA was then classified according to a sampling matrix dimension by Lav and Nav.

5.2.13    Lav was divided into three categories, as described below:

■    65 - 71 Lav: aircraft audible outdoors and may interfere with conversation outdoors; not usually audible indoors with windows shut;

■    71 - 77 Lav: aircraft can be audible indoors with windows open; and

■    78+ Lav: may interfere with conversation indoors with windows open.

**mva**consultancy

5.2.14    Similarly, Nav was divided into four categories:

- 120 – 380 Nav;

- 380 – 1200 Nav;

- 1200 – 3800 Nav; and

- > 3800 Nav.

5.2.15    The category boundaries were defined so that there was a fairly even coverage of OAs within each category.  However, just 2% of the OAs were located in the high sound level Lav category, and so these were disaggregated by just two Nav categories:

- 120 – 2200 Nav; and

- > 2200 Nav.

5.2.16    Table 5.2 shows the proportion of in-scope OAs within each of the Lav by Nav matrix cells.

**Table 5.2 Proportion of OAs in each Matrix Cell**

| Proportion of in scope OAs per cell | 120 - <380 Nav | 380 - <1200 Nav | 1200 - <3800 Nav | ≥ 3800 Nav |
|---|---|---|---|---|
| | 120 – <2200 Nav | | ≥ 2200 Nav | |
| "High" Noise ≥ 77 Lav | 1% | | 1% | |
| "Moderate" Noise 71 - <77 Lav | 7% | 8% | 2% | 3% |
| "Low" Noise 65 - < 71  Lav | 34% | 30% | 6% | 8% |

5.2.17    For some matrix cells, sampling in proportion to the population would have meant that the vast majority of sampled sites would have been around Heathrow.  This was undesirable, given the study's national focus, so a Heathrow / non-Heathrow stratification was also introduced.

**mva**consultancy

5.2.18    Table 5.3 shows the number of OAs within each matrix cell.

### Table 5.3 Full Survey Sampling Matrix

| Number of OAs in scope for sampling | 120 - <380 Nav | 380 - <1200 Nav | 1200 - <3800 Nav | ≥ 3800 Nav |
|---|---|---|---|---|
| | 120 – <2200 Nav | | ≥ 2200 Nav | |
| **Heathrow** | | | | |
| "High" Noise ≥77 Lav | 0 | | | 84 |
| "Moderate" Noise 71 - <77 Lav | 363 | 262 | 43 | 164 |
| "Low" Noise 65 - < 71 Lav | 1554 | 1150 | 141 | 379 |
| **Other Airports** | | | | |
| "High" Noise ≥77 Lav | 40 | | | 14 |
| "Moderate" Noise 71 - <77 Lav | 120 | 284 | 122 | 39 |
| "Low" Noise 65 - < 71 Lav | 825 | 926 | 251 | 142 |

5.2.19    Once the matrix had been established, sites were randomly selected from each cell.

5.2.20    The sample for the full survey consisted of 36 sampling points with 60 interviews to be conducted at each, giving a total target sample size of 2160 interviews. Two points within each of the 65 – 77 Lav cells, and one point in each of the cells with ≥77 Lav were selected. (The exception to this was the ≥ 77 Lav and 120 – 1200 Nav cell, where there were no OAs close to Heathrow in scope. Two sites at other airports were chosen instead.) Each sampling point was defined such that the noise exposure across the site varied by no more than 3.5 dB (see Chapter 2).

5.2.21    As, in general, there are not sufficient households within an OA to achieve the desired sample size of 60 interviews, each selected OA was paired with another. This second OA was randomly selected from within the same cell of the matrix, from OAs which could be paired without exceeding a maximum variation 3.5 dB LAmax. If it was not possible to pair an OA, this site was rejected and another selected.

### Selecting Areas for the Restricted Survey

5.2.22    There are two categories of areas which are not covered by the full survey:

■    areas outside the 120 Nav threshold ('low traffic' areas) around airports with regular traffic; and

■    areas around airports for which data is not available, or traffic is irregular ('irregular traffic' areas).

**mva**consultancy

5.2.23   For the 'low traffic' areas, the Lav and Nav values are in fact known, which made a Lav/Nav stratification possible.   Due to the low levels of variation in Nav, we restricted the stratification by Lav as follows:

- areas below 71 Lav; and

- areas equal to or above 71 Lav.

5.2.24   For each of these two categories, five sites were selected around Heathrow airport, and five sites from the remaining airports with regular traffic.

5.2.25   For the 'irregular traffic' areas, we did not have Lav and Nav values, and the only possible stratification was by LAmax, as follows:

- areas below 71 LAmax; and

- areas equal to or above 71 LAmax.

5.2.26   Ten sites were selected from each of these two categories.

5.2.27   Because the number of interviews per site was lower in the restricted survey than in the full SP survey, there was no need to conduct any pairing of OAs.

### Distribution of Survey Sites Between Airports

5.2.28   Detailed discussion of the issues involved in the selection of airports and defining the noise affected envelope which surrounds them can be found in MVA's sampling report[12].

5.2.29   Maps showing the distribution of the 76 sites surveyed are given in Appendix A5.  Note that LAeq values could only be calculated for **56** of these sites (the 36 sites for the "full" survey", and 20 out of 40 sites for the "restricted" survey – those described as "low traffic" rather than "irregular traffic").

---

mvaconsultancy

5.2.30    Table 5.4 provides summary details by airport of the number of randomly selected sites in both the full and restricted surveys.

**Table 5.4 Number of Sites Selected by Airport**

| Airport Name | Number of Full Survey Sites Selected | Number of Restricted Survey Sites Selected | Airport Name | Number of Full Survey Sites Selected | Number of Restricted Survey Sites Selected |
|---|---|---|---|---|---|
| Birmingham | 5 | 0 | Leeds Bradford | 2 | 1 |
| Bournemouth | 0 | 3 | Liverpool | 0 | 4 |
| Bristol | 0 | 2 | London City | 1 | 0 |
| Coventry | 0 | 2 | Luton | 1 | 0 |
| East Midlands | 0 | 0 | Manchester | 6 | 7 |
| Exeter | 0 | 0 | Newcastle | 0 | 6 |
| Gatwick | 0 | 1 | Norwich | 0 | 1 |
| Heathrow | 17 | 10 | Southampton | 4 | 0 |
| Humberside | 0 | 0 | Stansted | 0 | 1 |
| Kent International | 0 | 0 | Teesside | 0 | 2 |

**mva**consultancy

# 6 Fieldwork

## 6.1 Introduction

6.1.1 Following Phase 1 of this research and the sampling of sites, the questionnaire and interview procedure for Phase 2 was finalised. This chapter details the selection of individuals within each site, the questionnaire design and interview procedure, and details of the response rates and respondents.

## 6.2 Selection of Respondents

6.2.1 Within each site, a sample of residential addresses was randomly selected for interviewing:

- 120 addresses were selected within each full survey site from which 60 interviews were sought; and

- 30 addresses were selected within each restricted survey sampling point from which 15 interviews were sought.

6.2.2 Booster addresses were provided in the case of high refusal rates or non-contacts. Non-contacts were only recorded after five or more callbacks had been carried out (on different days of the week and times of day) at the the selected addresses. The outcome of each call-back at each household was recorded by interviewers.

6.2.3 Once contact was made with the household, a Kish Grid system was used to ensure that the household member to be interviewed was systematically selected with all household members having an equal chance of being interviewed.

6.2.4 The method avoids the possible bias that can be caused by interviewers interviewing only the most accessible household members. If the selected member of the household was not available at the time that the initial contact was made with the household, the interviewer made call-backs until a successful contact was made with the selected individual. No substitutions of household members were allowed in the case of the selected individual being unavailable or unwilling to participate.

6.2.5 The Kish Grid system was also used in cases where the listed address consists of more than one household. Once a household is selected, the Kish Grid procedure is used again to select a household member.

6.2.6 All interviewers were given a detailed set of interview procedures, and were in regular contact with supervisors to answer queries. In the full SP survey, all interviewers additionally attended a one day face-to-face briefing, which covered all interview procedures as well as the set up and use of the audio presentation equipment.

6.2.7 The full survey was undertaken during the period August 2005 to January 2006, whilst the restricted survey was undertaken between November 2005 and February 2006.

**mva**consultancy

### 6.3    Questionnaire structure

6.3.1    The questionnaire and related survey material drew upon the questions and approaches tested in earlier (Phase 1) pilots, and tested in its entirety in the Comparative Performance Trial.

6.3.2    A fundamental component of the questionnaire was the way in which we asked respondents to provide the attitude to aircraft noise.  As explained in Section 4.5, we chose to adopt the current ISO standard reported question, and include it within the context of eight other possible neighbourhood noise sources that could lead to annoyance.

> *"Thinking about the last 12 months or so, when you are at home, how much does noise from **[LIST SOURCES BELOW]** bother, disturb or annoy you: Not at all, Slightly, Moderately, Very, Extremely?"*

6.3.3    The noise sources were (in order that they were presented to each respondent): road traffic, trains, alarms/sirens, aircraft, animals, neighbours/children, pubs/night clubs, factories/industry, other.

6.3.4    This question was supplemented immediately after with another based on an 11-point numbered scale, as follows.

> *"Next I'd like to ask you how much aircraft noise bothers you when you are at home, on a scale of zero to ten.  If you are not at all annoyed choose zero, if you are extremely annoyed choose ten, if you are somewhere in between, choose a number between zero and ten.*
>
> *So, thinking about the last 12 months or so, what number from zero to ten best shows how much you are bothered, disturbed, or annoyed by aircraft noise?"*

6.3.5    With the exception of the SP and CVM questions, that are explained in detail in Chapter 10, the remainder of the questionnaire collected information that could usefully explore possible factors that might confound or influence the reported annoyance/aircraft noise relationship. These include:

- attitudinal factors (such as whether they think the airport has a positive or negative effect on the local community, awareness of any recent media comments concerning their local airport)
- periods of the day and night when the respondent and other members of the household are at home;
- household factors (such as members of the household employed within the aviation industry, household income, double glazing); and
- personal factors (such as being employed within the aviation industry, gender, age, working status, and whether they are generally around the home and exposed to noise from aircraft).

6.3.6    The CVM question asked respondents to compare their current house with an imaginary house that is exactly like their house in every respect except that it has zero aircraft noise and it would cost more to live there, in order to identify the amount of money respondents would be willing to pay to have no aircraft noise.

**mva**consultancy

6.3.7    For respondents participating in the full survey, after a suitable introduction, they were presented with the SP exercise, trading off between different numbers of different types of aircraft and money (in the form of household grants).

6.3.8    As explained in 4.3.26, each respondent was presented with SP options relating to some time periods.  Therefore, different respondents were presented with exercises for a mix of different time periods in such a way that when combining results across each sub-sample, the whole 24-hour day was covered evenly within each site.  The time periods presented were also systematically varied across the sample to ensure each was considered by around half the sample.

6.3.9    Examples of survey material are provided in Appendix A6.1.


## 6.4    Interview Procedure for SP Questions

6.4.1    Each interviewer in the full SP survey was equipped with a loudspeaker and CD player to play the recordings to each respondent.  In order to ensure that each respondent experienced the aircraft noise recordings in a comparable environment, the following instructions were followed:

   ■    set up the equipment and interview in a quiet room free, in so far as was practicable, from other distractions (TVs and radios to be turned off, no distractions from other household members);

   ■    set up in a room large enough to allow the respondent to be seated 1.5 metres from the loudspeaker, with no obstacles in between (eg tables or other furniture);

   ■    both the loudspeaker and the respondent should be positioned as far away as possible from reflective surfaces such as walls;

   ■    carpeting should be used to cover the floor immediately in front of the loudspeaker in the case of bare floors and wooden floorboards etc, to prevent sound reflection; and

   ■    Loudspeaker should be positioned a measured 1.5 metres from the respondents ear.

6.4.2    Each respondent was played a high quality recording of the sound made by each type of aircraft prior to the SP exercise.  The sound levels for the audio presentations were pre-set on the CDs issued to interviewers according to the calculated LAmax levels for the relevant aircraft types at each sample site area.

6.4.3    The audio presentations were accompanied by an appropriate photo of the aircraft type in question (showing aircraft of the type, size and visual angle of that which would be experienced in each site).

**mva**consultancy

**6.5    Response Rates**

6.5.1    The contact outcomes are summarised in Table 6.1, in the order in which each potential obstacle to a successful interview can occur.

**Table 6.1 ANASE Response Rates**

| Outcomes | Percentage (of total sampled addresses) |
|---|---|
| Total sampled addresses, of which: | 100% |
| Invalid | 1% |
| Refusal | 23% |
| Person unavailable/recruited but failed to be interviewed | 5% |
| Non-Contacts (genuinely exhausted addresses) | 13% |
| Non-Contacts (non-exhausted addresses[1]) | 10% |
| **Completed Interviews (Gross Response Rate)** | **49%** |

[1]    the breakdown by exhausted/non-exhausted is based on a manual analysis of a sub-sample only of all non-contacts

6.5.2    The 'non-exhausted addresses' are addresses which were contacted less than the requisite 5 times before interviewing in the area ceased.  (Typically, there were three interviewers working in each (Full Survey) site, each responsible for achieving 20 interviews from 40 addresses, so when 60 interviews were achieved there were potentially three subsets of addresses not fully exhausted).

6.5.3    At one site near Southampton airport (O6D), one of the interviewers was threatened after 21 interviews were achieved, and so no further interviews took place at this site on the grounds of interviewer safety.

**mva**consultancy

### 6.6    Achieved Sample

6.6.1    Table 6.2 summarises the number of interviews achieved.

**Table 6.2  Interviews Achieved**

| Survey Type | Sites | Target Number of Interviews | Interviews Achieved |
|---|---|---|---|
| Full | 36 | 2160 | 2132 |
| Restricted | 40 | 600 | 601 |
| Total | 76 | 2760 | 2733 |

6.6.2    Across the whole sample:

- 47% were male;

- 32% were aged 18 – 34 and 18% were aged 65 or over;

- 84% of respondents gave their ethnic origin as White, and 8% as Asian;

- 47% of respondents were employed full-time, and 22% were retired;

- of those that worked, 12% of respondents usually worked from home and 17% often worked shifts; and

- for those in multi-person households, 10% of other household members usually worked from home and 14% of other household members often worked shifts.

- 17% of respondents were in socio-economic groups A or B; and 27% were in socio-economic groups D or E.

6.6.3    A full breakdown of the profile of our sample, by site, is provided in Appendix A6.2.  Tables showing how the respondents answered each question are given in Appendix A6.3.

### 6.7    Weighting for Sample Bias

6.7.1    Procedures were needed to deal with two main sources of potential bias arising from the multi-stage sample design:

- the sampling unit was the household but the information obtained is from a single individual; and

- the profile of our samples of respondents may differ from that of the adult populations in the surveyed areas, due to differences in survey response rates (e.g. older residents being more or less likely to participate).

6.7.2    Correction for household size was achieved by post-weighting each individual's responses in proportion to the number of adults living at the same address.  Without this correction, the

**mva**consultancy

views of people in larger households would be under-represented, as they each have a smaller chance of participating.

6.7.3    Post-weighting the data to correct for sampling and non-response bias was based upon the results of two sets of analysis:

■        identification of those profile characteristics that most significantly influence reported annoyance to aircraft noise; and

■        comparisons between our adjusted sample, at the level of the individual, and the 2001 Census.

6.7.4    To make the bias adjustment robust, adjustment factors were calculated separately for the full and restricted datasets.  The sample sizes were too small to calculate factors for each individual site.

6.7.5    We examined the effect of a number of individual variables on the response to the ISO annoyance question about aircraft noise (see Chapter 8).  In terms of variables relating to the individual respondent, age was shown to be an influential non-noise factor on annoyance, and there were significant differences at the 1% level in the age distribution of the sample compared with the population.  It was therefore decided to apply a separate bias adjustment factor to each age group.  The age adjustment factors used for the full survey and restricted survey datasets are shown in Table 6.3.

**Table 6.3 Weighting Factors by Age Group**

| Age Group | Full Survey | Restricted Survey |
|-----------|-------------|-------------------|
| 18-24     | 0.95        | 1.15              |
| 25-34     | 1.23        | 1.17              |
| 35-44     | 1.01        | 1.00              |
| 45-54     | 0.90        | 1.05              |
| 55-64     | 0.91        | 0.93              |
| 65+       | 0.96        | 0.83              |

6.7.6    These factors are used when aggregate measures are derived from the data, for instance for an individual site or for aggregations of sites.  We also carried out selected analysis using unweighted data.  The use of weighted or unweighted data made very little difference to the results.  Details of the weighting process are given in Appendix A6.2.

**mva**consultancy

# 7    Presentation of Reported Annoyance

## 7.1    Introduction

7.1.1    The first two objectives of this study are to re-assess attitudes to aircraft noise in England and their correlation with the LAeq noise index. In order to do this, it is necessary to have two measures for each site surveyed: an annoyance measure and a sound level measure.

7.1.2    The calculation of the noise measures is described in Chapter 2. We generally present the annoyance data against LAeq at the site level. However, there are 20 sites where LAeq is not available, and for these sites we present the annoyance data against LAmax.

7.1.3    This chapter considers the different measures of annoyance that can be used at a site level and how these annoyance measures relate to LAeq. The main measures at site level – LAeq and mean annoyance – are shown overall, and, for the three airports with the most sites, in map form. The objective of this chapter is to present and comment on the survey results. More detailed analysis of the data is discussed in Chapter 8.

## 7.2    Measures of Annoyance

7.2.1    Respondents were asked both versions of the ISO noise annoyance question, directly relating to their annoyance with aircraft noise:

- *"Thinking about the last 12 months or so, when you are at home, how much does noise from aircraft bother, disturb or annoy you: Not at all, Slightly, Moderately, Very, Extremely?"*

- *"Thinking about the last 12 months or so, what number from zero to ten best shows how much you are bothered, disturbed, or annoyed by aircraft noise?"*

7.2.2    Analysis of the responses between these two questions showed that there was a high correlation of 0.89 between these two responses (see Appendix A7), and the subsequent analysis focused on the first question.

7.2.3    In order to obtain a single annoyance score for each site, it was necessary to combine the responses, and this was carried out in two ways:

- Calculating a mean annoyance score; and

- Calculating the percentage of respondents who were annoyed to a given degree.

**mva**consultancy

7.2.4   The calculation of the mean annoyance was carried out in line with research undertaken by Miedema and Oudshoorn[13], who transformed all annoyance scales to run from 0 to 100.  The distribution used in the ANASE survey, matched the definition used by Van Kempen and Kamp[14] which scores the standardised 5-point noise annoyance scale as 10, 30, 50, 70 and 90 points on the Miedema and Oudshoorn scale:

- Not at all annoyed          10

- Slightly annoyed          30

- Moderately annoyed          50

- Very annoyed          70

- Extremely annoyed          90

7.2.5   Each individual response was weighted to remove sample bias, as discussed in Chapter 6.

## 7.3    Annoyance Responses by Site

7.3.1   Table 7.1 shows the percentage of respondents at least slightly annoyed (slightly, moderately, very or extremely), the percentage of respondents at least very annoyed (very or extremely) and the mean annoyance by site.  Each of the 36 full survey sites and the 20 restricted survey sites for which LAeq data are available are shown, in ascending order of LAeq.  Those site names which begin with an "R" are restricted survey sites, and those starting with an "O" or "H" are full survey sites.

### Table 7.1 Reported annoyance and LAeq by Site

| Site | Location | Airport | 16 Hour LAeq | Mean Annoyance | % of Respondents at Least | | | |
|------|----------|---------|------|------|------|------|------|------|
| | | | | | Slightly annoyed | Moderately annoyed | Very annoyed | Extremely annoyed |
| R01 | Tooting | Heathrow | 40.9 | 19 | 35% | 11% | 0% | 0% |
| R02 | Colliers Wood | Heathrow | 41.6 | 17 | 35% | 0% | 0% | 0% |
| R18 | Fallowfield | Manchester | 41.9 | 15 | 22% | 5% | 0% | 0% |
| R11 | Beeston Hill | Leeds | 42.2 | 10 | 0% | 0% | 0% | 0% |
| R20 | Withington | Manchester | 42.9 | 12 | 10% | 0% | 0% | 0% |
| R03 | S Wimbledon | Heathrow | 43.0 | 25 | 36% | 31% | 5% | 5% |
| R15 | Stockport | Manchester | 43.1 | 14 | 21% | 0% | 0% | 0% |
| R13 | Dukinfield | Manchester | 43.3 | 19 | 34% | 6% | 6% | 0% |
| O6F | Hillbrook | Southampton | 43.5 | 44 | 82% | 57% | 26% | 5% |
| R12 | Harlford | Manchester | 44.0 | 11 | 4% | 0% | 0% | 0% |

---

[13] '*Position Paper on Dose Response Relationships between Transportation Noise and Annoyance*' (and Appendix); European Commission Working Group 2 (appendix by H.M.E. Miedema and C.G.M Oudshoorn); ISBN 92-894-3894-0 (Appendix TNO Report PG/VGZ/00.052); 2002 (Appendix July 2000)

[14] 'Annoyance from Air Traffic Noise: Possible Trends in Exposure:Response Relationships'; EEMM van Kempen, I van Kamp; RIVM Report 01/2005

**mva**consultancy

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| R19 | Heaton Moor | Manchester | 44.0 | 25 | 56% | 13% | 7% | 0% |
| R14 | Gatley | Manchester | 45.3 | 17 | 23% | 14% | 0% | 0% |
| O6E | Dibden | Southampton | 45.8 | 18 | 27% | 13% | 3% | 0% |
| H3C | South Ealing | Heathrow | 46.0 | 41 | 70% | 54% | 24% | 7% |
| R06 | Sunbury | Heathrow | 46.5 | 22 | 38% | 14% | 6% | 0% |
| R09 | Teddington | Heathrow | 47.2 | 42 | 78% | 49% | 33% | 0% |
| R05 | WaltonThames | Heathrow | 47.5 | 24 | 38% | 31% | 0% | 0% |
| O6D | Midanbury | Southampton | 47.5 | 28 | 50% | 26% | 13% | 3% |
| R04 | Hersham | Heathrow | 47.6 | 39 | 77% | 54% | 14% | 3% |
| R17 | Harlow | Stansted | 47.7 | 63 | 88% | 84% | 64% | 28% |
| R16 | Smallfield | Gatwick | 48.5 | 25 | 47% | 18% | 6% | 6% |
| R08 | Kempton Park | Heathrow | 48.9 | 27 | 56% | 17% | 12% | 0% |
| O2C | Headingly | Leeds | 49.2 | 38 | 82% | 45% | 11% | 1% |
| H5E | Virginia Water | Heathrow | 49.6 | 30 | 64% | 28% | 5% | 0% |
| O2A | Central Luton | Luton | 49.6 | 33 | 63% | 35% | 12% | 4% |
| O4D | Sutt'nColdfield | Birmingham | 49.7 | 46 | 72% | 53% | 34% | 22% |
| H3A | West Ealing | Heathrow | 50.4 | 45 | 77% | 56% | 32% | 12% |
| R10 | Hanworth | Heathrow | 50.4 | 30 | 62% | 36% | 3% | 0% |
| H3B | North Ealing | Heathrow | 50.5 | 50 | 88% | 62% | 36% | 12% |
| H5A | Windsor For'st | Heathrow | 50.9 | 29 | 55% | 23% | 9% | 8% |
| O4G | Eastleigh | Southampton | 51.0 | 31 | 70% | 30% | 7% | 0% |
| H3D | Chiswick | Heathrow | 52.7 | 48 | 81% | 63% | 34% | 12% |
| O4A | Horsforth | Leeds | 52.7 | 42 | 73% | 47% | 30% | 13% |
| H3E | W Brompton | Heathrow | 53.0 | 35 | 72% | 38% | 14% | 3% |
| O2E | Bredbury | Manchester | 53.9 | 55 | 97% | 73% | 44% | 11% |
| H1P | East Sheen | Heathrow | 54.7 | 64 | 96% | 83% | 57% | 34% |
| R07 | Kneller Hall | Heathrow | 55.2 | 42 | 93% | 49% | 13% | 6% |
| O2D | Knutsford | Manchester | 55.6 | 53 | 94% | 68% | 41% | 12% |
| O4E | C'tle Bromwich | Birmingham | 55.6 | 39 | 69% | 47% | 21% | 8% |
| H5B | Eton | Heathrow | 56.1 | 52 | 82% | 71% | 40% | 20% |
| H5F | Dedworth | Heathrow | 56.2 | 54 | 82% | 72% | 44% | 22% |
| O2F | N Stockport | Manchester | 56.6 | 61 | 95% | 77% | 55% | 30% |
| O4B | Cheadle Heath | Manchester | 58.4 | 53 | 86% | 65% | 42% | 24% |
| H5D | Old Windsor | Heathrow | 58.7 | 59 | 94% | 78% | 50% | 25% |
| H5C | South Windsor | Heathrow | 59.3 | 73 | 99% | 93% | 73% | 48% |
| H1L | Hounslow H'th | Heathrow | 59.7 | 74 | 100% | 96% | 73% | 49% |
| H1M | S Hounslow | Heathrow | 59.8 | 66 | 96% | 80% | 61% | 42% |
| O6B | Glebe Farm | Birmingham | 59.9 | 48 | 84% | 57% | 27% | 20% |
| H1K | Isleworth | Heathrow | 60.3 | 62 | 96% | 86% | 49% | 31% |
| O5A | Thamesmead | London City | 60.6 | 52 | 83% | 72% | 41% | 14% |
| O6C | Colehall | Birmingham | 61.0 | 57 | 84% | 73% | 52% | 24% |
| O4C | Cheadle | Manchester | 61.6 | 57 | 93% | 87% | 46% | 11% |
| H1J | W Hounslow | Heathrow | 61.7 | 64 | 94% | 79% | 61% | 37% |
| O4F | CheadleHulme | Manchester | 62.8 | 59 | 86% | 82% | 48% | 27% |
| H1H | Osterley | Heathrow | 63.1 | 76 | 97% | 96% | 80% | 58% |
| O6A | Kitts Green | Birmingham | 64.2 | 68 | 88% | 77% | 68% | 56% |

**mva**consultancy

**Degrees of Annoyance**

7.3.2    Figure 7.1 shows the percentage of respondents at least slightly annoyed in each of the sites against LAeq:

■     as LAeq increases, the percentage of respondents who were at least slightly annoyed increases up to a LAeq value of around 52;

■     for sites with a LAeq value less than about 46, all but two sites have fewer than 40% of respondents at least slightly annoyed;

■     for areas with a LAeq greater than 48 on the chart, at least 45% of respondents are at least slightly annoyed; and

■     for areas with a LAeq greater than 54, apart from one site, at least 80% are at least slightly annoyed.
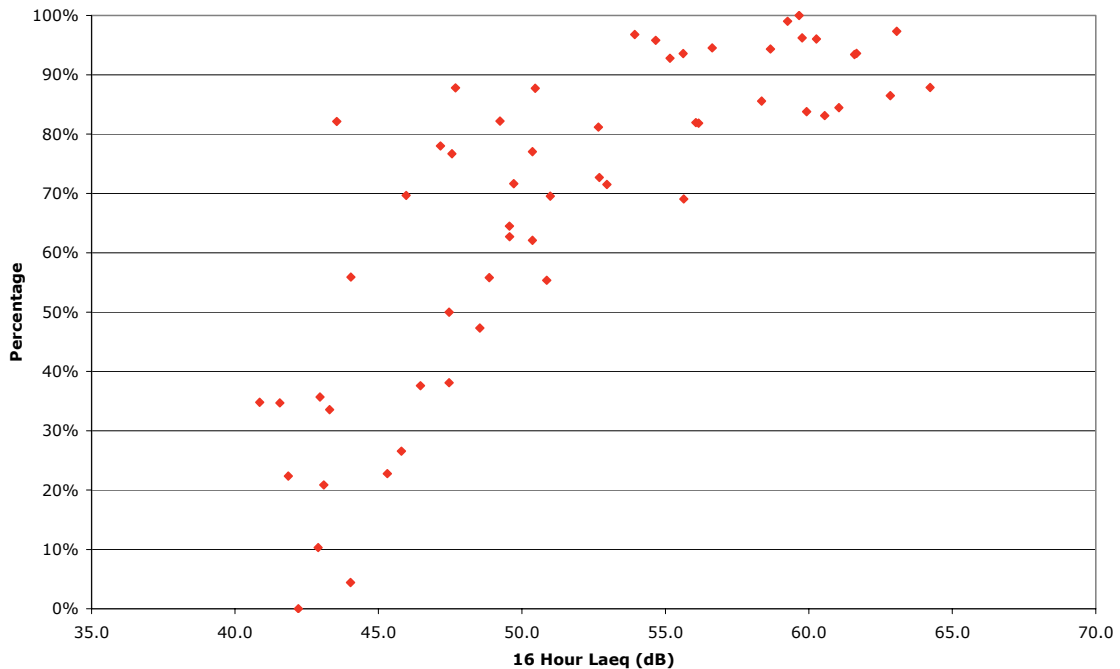


**Figure 7.1 Percentage of Respondents at Least Slightly Annoyed with Aircraft Noise**

7.3.3    There is a greater spread in the percentage of respondents who are at least slightly annoyed at lower LAeq values than at higher LAeq values, since the majority of people are "at least slightly annoyed" at higher LAeq values.  For instance, for sites with a LAeq value of around 44, between 4% and 82% of respondents are at least slightly annoyed.

**mva**consultancy

7.3.4    Figure 7.2 shows the percentage of respondents who were at least very annoyed in each of the sites by LAeq:

■    on this chart, for areas with LAeq less than around 43, the proportion of respondents who are at least very annoyed is less than 12%;

■    the proportion of respondents at least very annoyed generally increases with LAeq for values of LAeq over 43, although there is a relatively large spread in percentages for most LAeq values; and

■    for all except one of the areas with LAeq greater than 57, more than 60% of respondents were at least very annoyed.
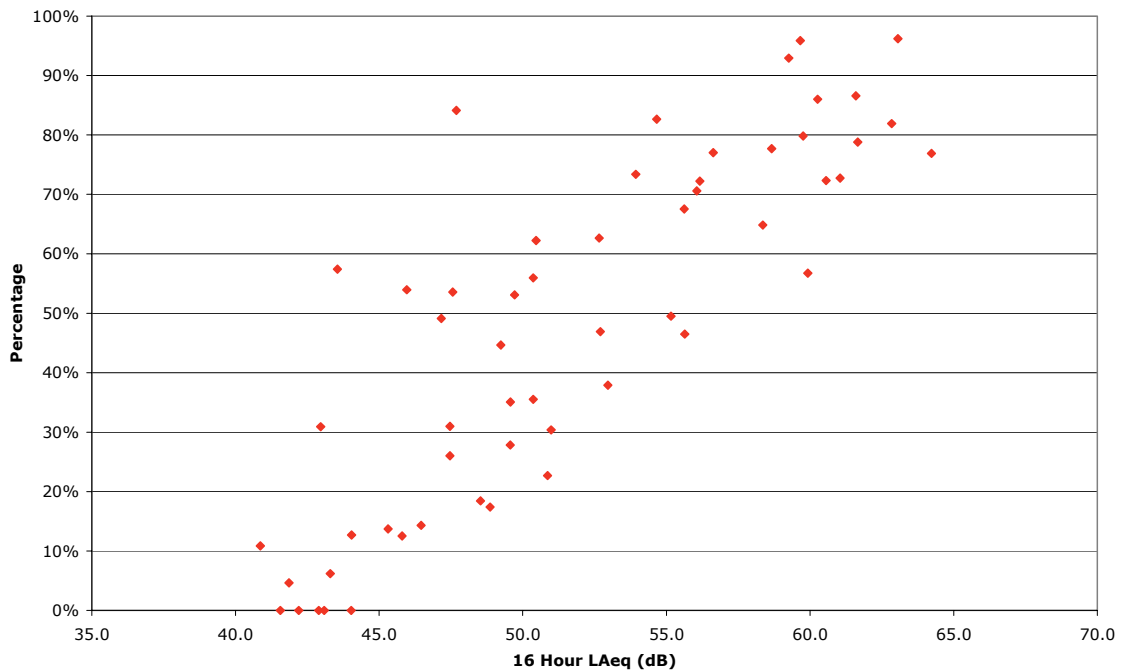


**Figure 7.2 Percentage of Respondents at Least Very Annoyed with Aircraft Noise**

7.3.5    Figure 7.3 shows the proportions of respondents that are at least slightly annoyed, at least moderately annoyed and at least very annoyed plotted against the mean site annoyance score.  The scatter points show the expected shape (ie "at least slightly annoyed" rising quickly and then flattening, "at least moderately annoyed" rising steadily across the range of mean annoyance, and "at least very annoyed" rising slowly at first and then increasing steadily).  Given the strong relationships between each of the annoyance categories and mean site annoyance, and the fact the mean score contains potentially more information, we continue our presentation of the data using LAeq and mean site annoyance.
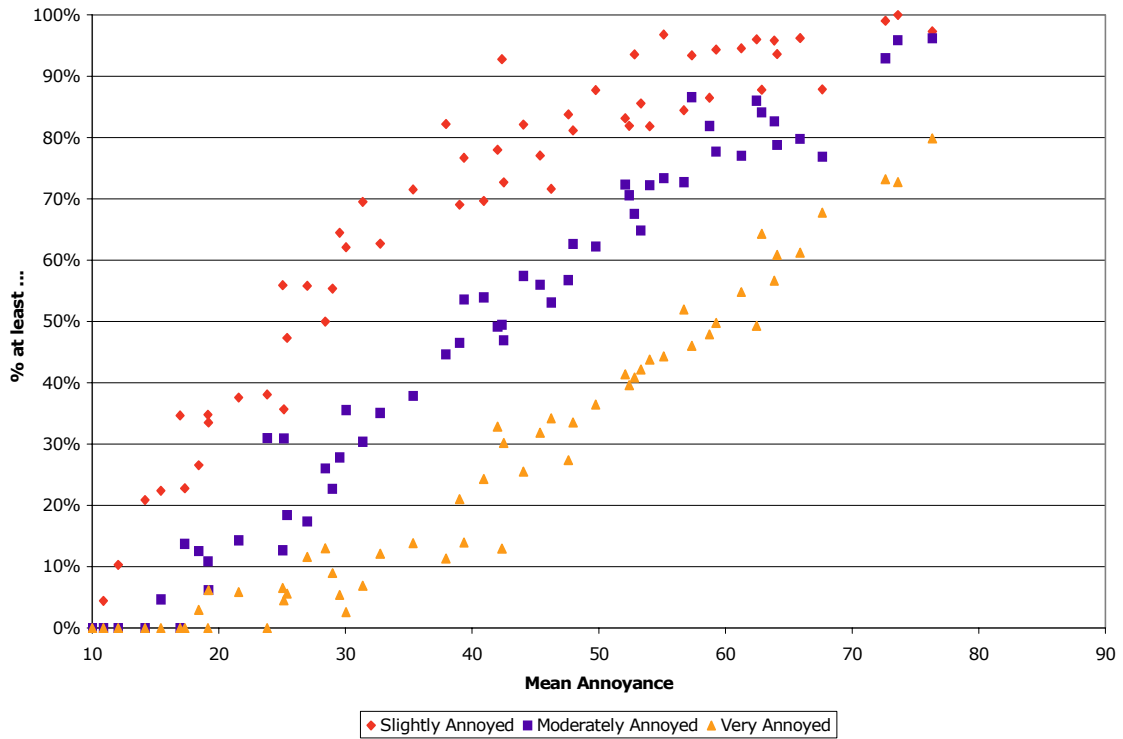


**Figure 7.3  Comparison of Annoyance Metrics**

**Mean Annoyance**

7.3.6    Figure 7.4 shows for each site, identified by the site code, the mean annoyance against the LAeq data, from which it can be seen that:

■    with one exception, the mean annoyance is below 30 ('slightly' annoyed) for all sites with values of LAeq lower than 45;

■    the mean annoyance generally increases from around 20 to 50 ('moderately annoyed') for values of LAeq between 45 and 54; and

■    the mean annoyance is generally at least 50 ('moderately' annoyed) for values of LAeq above 56.
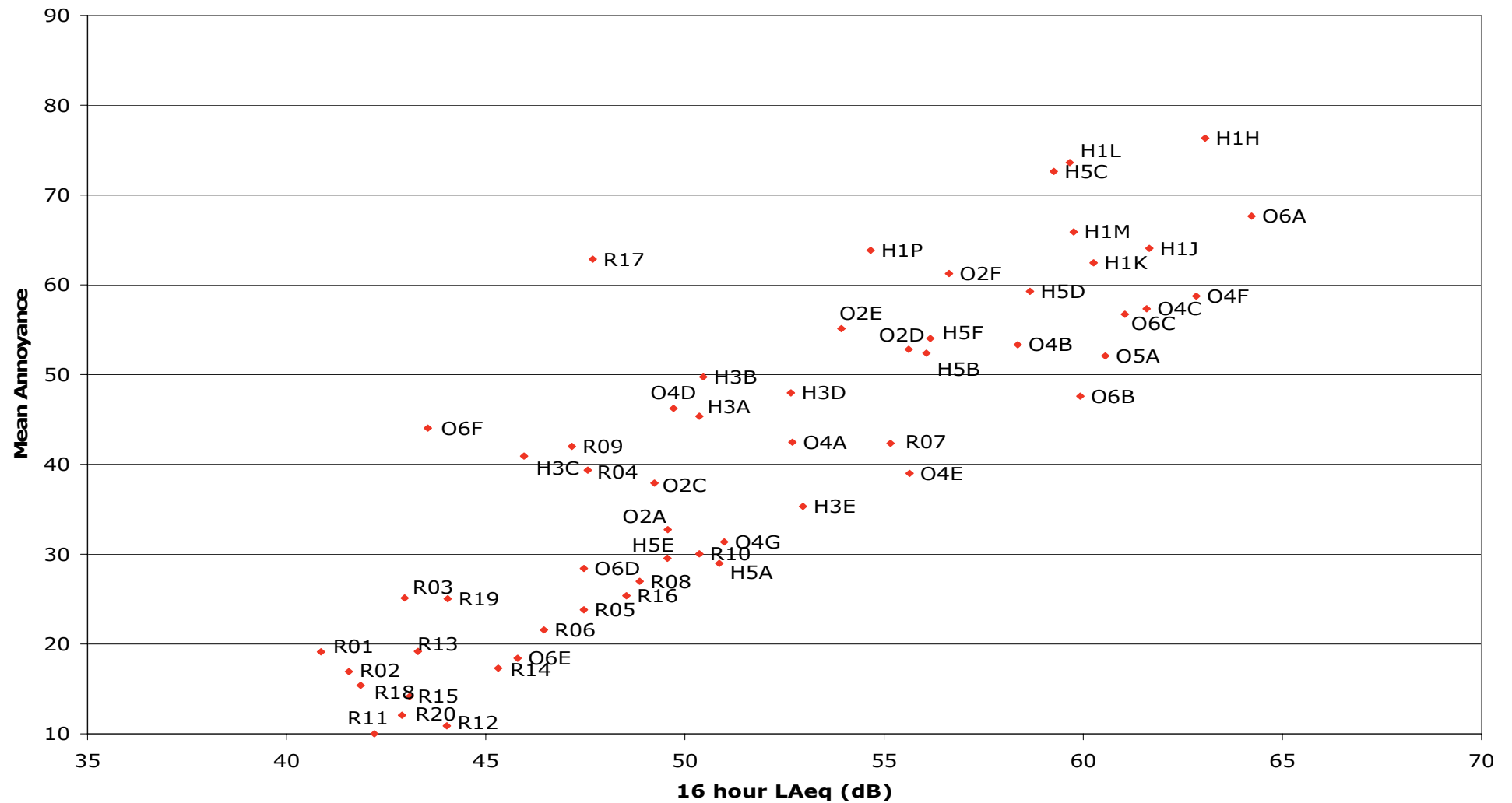
**mva**consultancy

**Figure 7.4 Mean Annoyance against LAeq**

7.3.7   There are two sites, R17 Harlow near Stansted airport and O6F Hillbrook near Southampton airport, where the mean annoyance is high for their LAeq value, compared to the other sites.

7.3.8   In Figures 7.5 to 7.7, we have plotted on maps the mean annoyance of each site[15] close to the three airports with the largest number of sites: Heathrow, Manchester and Birmingham. On all three maps, the colour of each dot indicates which LAeq band the site is within (e.g. dark red = LAeq 61-65 dB) and the higher the corresponding score the greater the mean annoyance expressed by residents.  Seven LAeq bands have been used in the presentation, covering the range of estimated LAeq in 4 dB intervals.

7.3.9   Sites where the full survey took place, and around 60 respondents were interviewed, are shown as circles; and sites where the restricted survey took place, and around 15 respondents were interviewed, are shown as diamonds.  The LAeq 57 contour[16] is also shown on the Heathrow map: the area within this contour has a LAeq value of 57 or more.
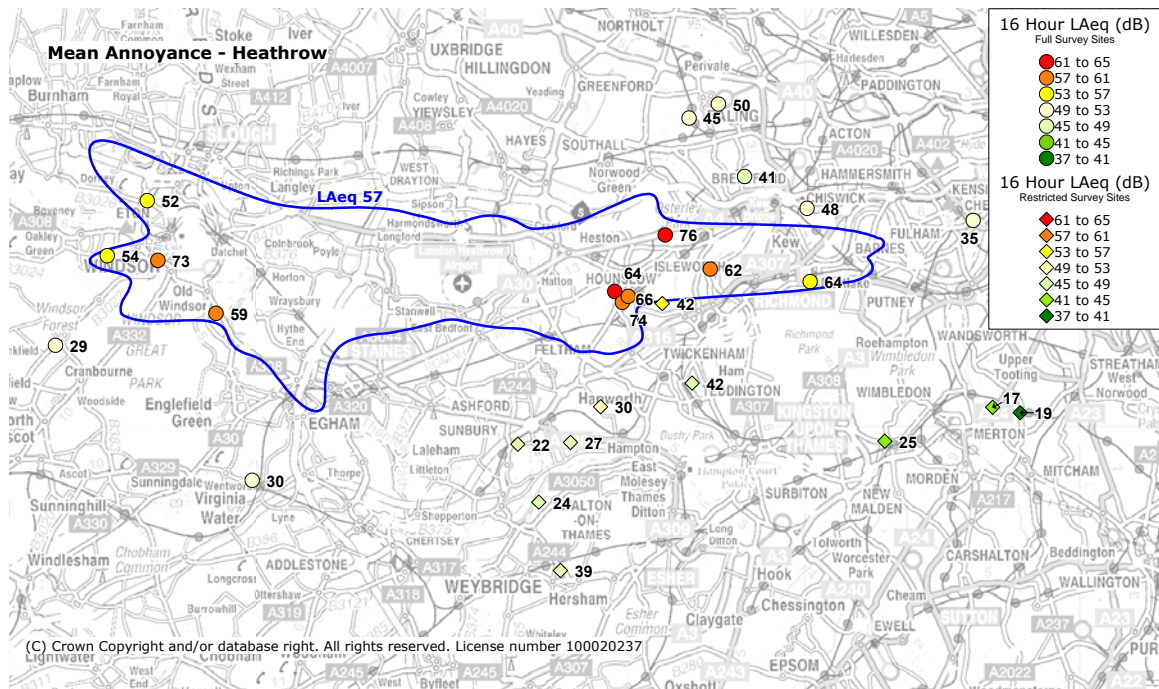


**Figure 7.5 Mapped Reported Annoyance with Aircraft Noise – Heathrow Airport**

---

[15] Using the grid reference of the population-weighted centroid

[16] 92 day long term modal average contour, 2005

7.3.10    Table 7.2 shows for Heathrow the range of mean annoyance scores in each LAeq band.

**Table 7.2 Range of Annoyance Scores in LAeq Band for Heathrow Airport**

| LAeq Band | No of Sites | Annoyance Range |
|---|---|---|
| 37-41 | 1 | 19 |
| 41-45 | 2 | 17-25 |
| 45-49 | 6 | 22-42 |
| 49-53 | 7 | 29-50 |
| 53-57 | 4 | 42-64 |
| 57-61 | 5 | 59-74 |
| 61-65 | 2 | 64-76 |

7.3.11    At levels of LAeq below 45, the annoyance range is small.  However, above LAeq 45, for each LAeq band, the mean annoyance varies between sites by around 20 points in the mean annoyance score, which is equivalent to one category (for example between slightly and moderately annoyed).  For higher LAeq bands, the range of annoyance scores reduces, as all sites in these bands have a high mean annoyance score.  Within the mapped 57 LAeq contour, the mean annoyance score is at least 52, just above moderately annoyed.  As would be expected, the mean annoyance score is lower for sites further away from the airport.

7.3.12    The mean annoyance is generally highest to the immediate east and west of the airport. There is a big change in the mean annoyance between east Windsor (very annoyed - 73, site H5C) and west Windsor (moderately annoyed - 54, site H5F), although the sites differ by just 3 LAeq.  Similarly, just south of Hounslow, there is a site with a mean annoyance score of very annoyed (74, H1L), close to another site with a mean annoyance score of half-way between slightly and moderately annoyed (42, R07), which differ by 4 LAeq.
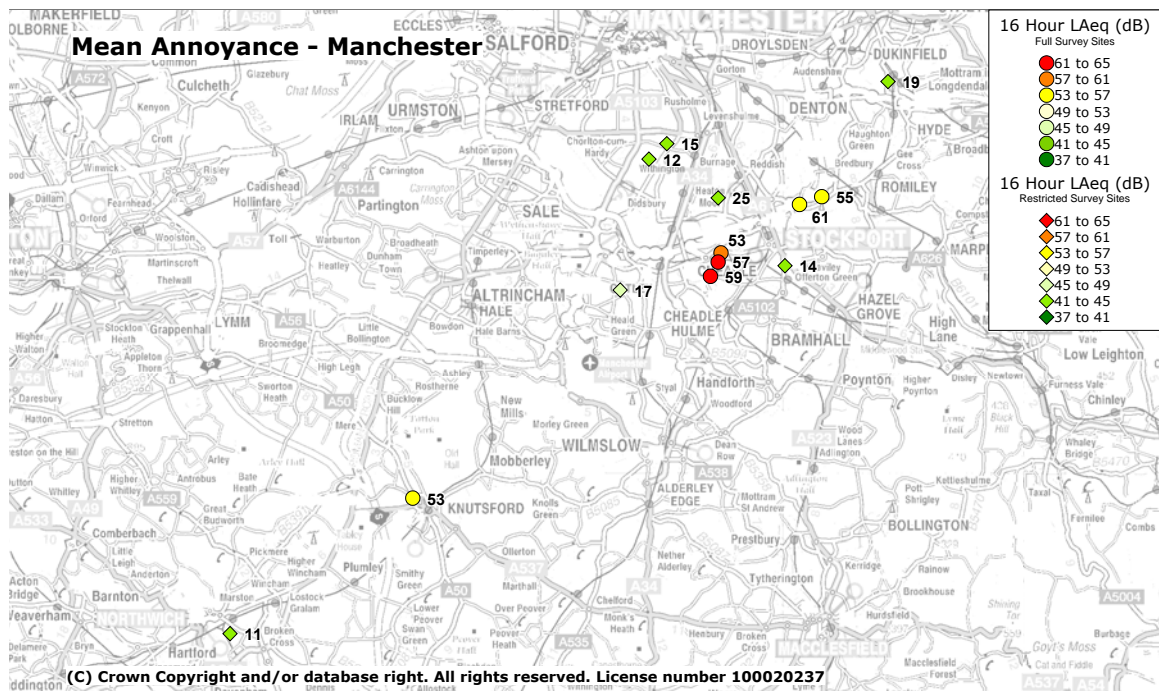
**Figure 7.6 Mapped Reported Annoyance with Aircraft Noise – Manchester Airport**

mvaconsultancy

7.3.13    Table 7.3 summarises the range of annoyance scores in each LAeq band for Manchester airport.

**Table 7.3 Range of Annoyance Scores in LAeq Band for Manchester Airport**

| LAeq Band | No of Sites | Annoyance Range |
|-----------|-------------|-----------------|
| 37-41 | 0 | |
| 41-45 | 6 | 11-25 |
| 45-49 | 1 | 17 |
| 49-53 | 0 | |
| 53-57 | 3 | 53-61 |
| 57-61 | 1 | 53 |
| 61-65 | 2 | 57-59 |

7.3.14    For sites with a LAeq less than 49, the mean annoyance was very low - at less than slightly annoyed (25) for each site.  Above LAeq 53, the mean annoyance was much higher, but within a small range, with the mean annoyance between moderately and very annoyed (53 to 61).

7.3.15    Birmingham airport has fewer sites covered in the research (Fig 7.7) and the range of mean annoyance scores in LAeq bands is shown in Table 7.4.
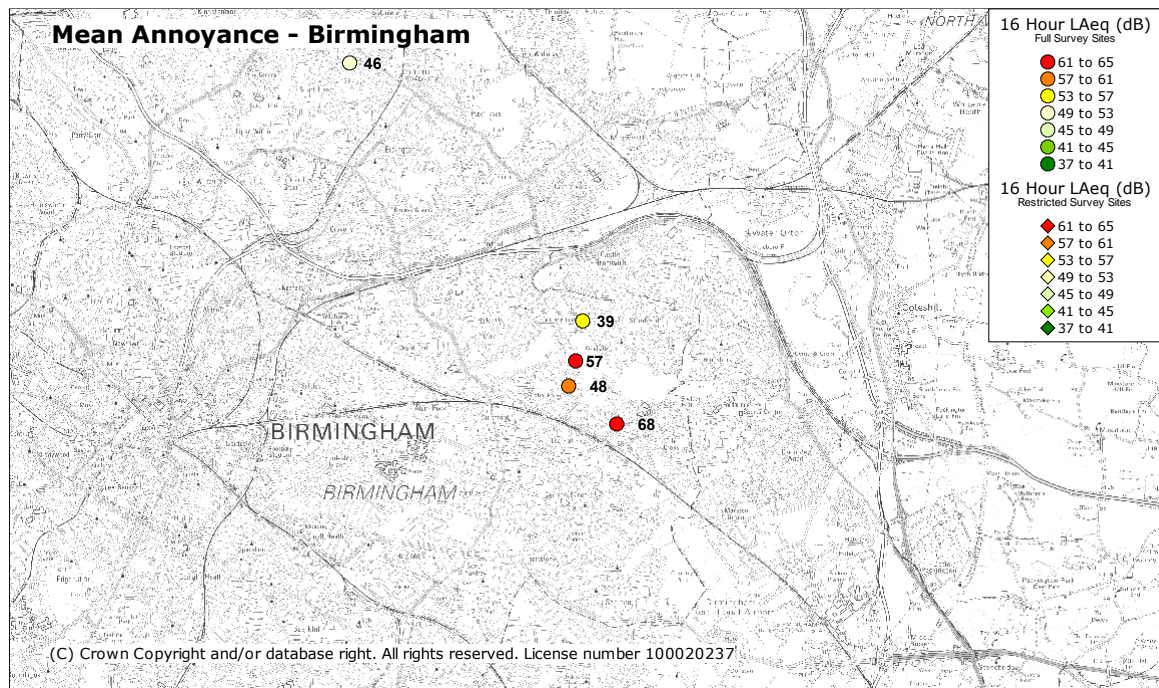


**Figure 7.7 Mapped Reported Annoyance with Aircraft Noise – Birmingham**

**Table 7.4 Range of Annoyance Scores in LAeq Band for Birmingham Airport**

| LAeq Band | No of Sites | Annoyance Range |
|---|---|---|
| 37-41 | 0 | |
| 41-45 | 0 | |
| 45-49 | 0 | |
| 49-53 | 1 | 46 |
| 53-57 | 1 | 39 |
| 57-61 | 1 | 48 |
| 61-65 | 2 | 57-68 |

7.3.16    At Birmingham, the highest mean annoyance scores were in the 61-65 LAeq band.  The three other Birmingham sites had similar mean annoyance scores, between slightly and moderately annoyed (39 to 48),  The site with the lowest LAeq (O4D) was much further away from the airport.

7.3.17    Taken overall, there appears to be a marked increase for Manchester in mean annoyance at a certain LAeq band (between 49-53dB).  For Manchester, all sites with a LAeq above 53 have high annoyance, whereas for Birmingham, sites above 61 have high annoyance.  For Heathrow however, the pattern is slightly different, with low annoyance observed for sites with a LAeq below 45, high annoyance for sites with a LAeq above 57, but a large range of annoyance values for sites with a LAeq between 45 and 57.

7.3.18    Generally, for like-for-like LAeq bands, mean reported annoyance at Heathrow is higher than at Birmingham and Manchester.

**Annoyance at Irregular Airport Sites**

7.3.19    For the 20 irregular traffic airports, the only sound level variable available for analysis is the maximum sound level, LAmax (as defined in para 2.2).  Figure 7.8 shows that there is a relationship of increasing annoyance as LAmax increases, and this is true for both the regular traffic airports (already shown plotted against LAeq in Figure 7.4) and the 20 irregular traffic airports.
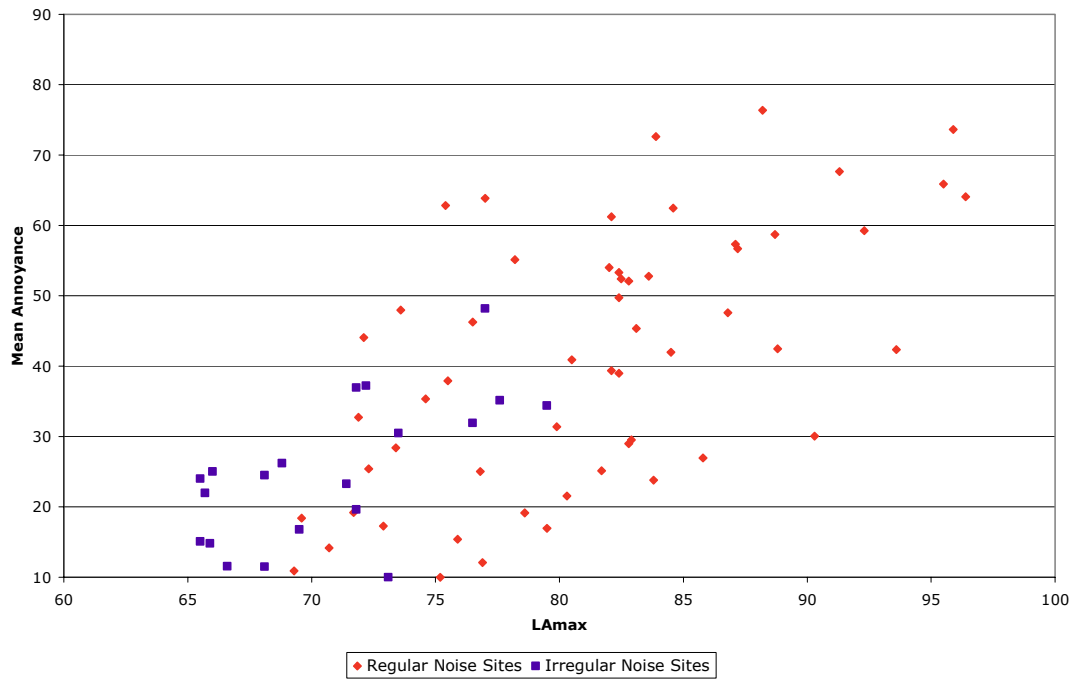
**mva**consultancy

**Figure 7.8 Mean Reported Annoyance with Aircraft Noise by LAmax**

7.3.20    There is no indication from these findings that irregular aircraft lead to greater reported annoyance than regular aircraft, for a given LAmax.


## 7.4    Summary of Main Points

- As LAeq increases, the mean annoyance, percentage of respondents at least slightly annoyed and the percentage of respondents at least very annoyed increases;

- Mean annoyance at sites near Heathrow is generally higher than mean annoyance at other airports for a given LAeq;

- As LAeq increases, the percentage of respondents who were at least slightly annoyed increases up to a LAeq value of around 52;

- For sites with a LAeq value less than 46, generally fewer than 40% of respondents are at least slightly annoyed;

- For areas with a LAeq greater than 48, at least 45% of respondents are at least slightly annoyed;

- For areas with a LAeq greater than 54, generally at least 80% are at least slightly annoyed;

- The proportion of respondents who are at least very annoyed is less than 10% for areas with LAeq less than 43;

- The proportion of respondents at least very annoyed generally increases with LAeq for values of LAeq over 43, although there is a relatively large spread in percentages for most LAeq values;

- At least 40% of respondents were generally at least very annoyed with LAeq greater than 57;

**mva**consultancy

- For a given LAeq range, there is a range of mean annoyance scores indicating that annoyance is not determined solely by the amount of aircraft noise as measured by LAeq;

- Mean annoyance at sites near Heathrow is generally higher than mean annoyance at other airports for a given LAeq;

- The mean annoyance for sites close to irregular traffic airports is similar to the mean annoyance for regular sites against LAmax.

**mva**consultancy

# 8 Relationship between Annoyance and LAeq

## 8.1 Introduction

8.1.1 A principal objective of the study was to consider the relationship between annoyance and LAeq. As LAeq is only available at site level, the majority of the analysis was carried out at a site level.

8.1.2 This chapter considers models, developed using regression analysis, that attempt to relate the reported annoyance, defined by their mean annoyance score (as described in Chapter 7), with the sound level metric LAeq (as discussed in Chapter 2).

8.1.3 Criteria for a satisfactory model include:

- the explanatory power of the model (indicated by a high $R^2$ value);

- the plausibility of the mechanisms suggested by the model (especially the signs of the relevant coefficients);

- the significance of each independent variable (indicated by high t-ratios);

- economy in terms of the numbers of variables used;

- (ideally) the inclusion of variables which are both relevant and predictable in the policy context; and

- random distribution of the residuals.

## 8.2 Type of Model

### Basic Linear Model

8.2.1 The simplest model form was:

Mean annoyance = a + b x LAeq

using the weighted mean annoyance score and the 16-hour LAeq value applying to the site.

8.2.2 The estimated model is:

### Model 8.1      Mean Annoyance = -80.0 + 2.3 x LAeq

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq |
|-------|----------------|-----|-----------|------|
| 0.739 | 0.734 | 56 | -79.95 | 2.34 |
| | | | (-8.09) | (12.35) |

8.2.3 In this and all later models, the bracketed figures are the t-ratios corresponding to the coefficients; values greater than about 2.0 (for a sample of this size) indicate coefficients that are significantly different from zero at the 5% confidence level, meaning that there is less than a 5% likelihood that such differences could have arisen by chance. T-ratios above

**mva**consultancy

about 2.6 indicate values with a 1% confidence level.   Both the constant and the LAeq coefficient are significant at the 1% level in this model.

8.2.4    The $R^2$ value expresses the proportion of the overall variation in mean annoyance that is explained by the model; Model 8.1 suggests that just under three-quarters of the variation in average reported annoyance between sites can be explained by LAeq alone.  The adjusted $R^2$ value takes account of the number of variables used; this allows models with different numbers of variables (degrees of freedom) to be compared on a like-for-like basis.

8.2.5    The coefficient on LAeq indicates the change in the mean annoyance score which results (on average) from a difference of 1 dB in the LAeq index.

8.2.6    Model 8.1 explains a high proportion of the variation, using just one behaviourally plausible and predictable independent variable of the correct sign (annoyance increasing with LAeq) at a high confidence level.

8.2.7    The data for mean annoyance are plotted in Figure 8.1, showing the relationship with LAeq in Model 8.1.  This figure shows the same data as that shown in Figure 7.4.  Curves have also been plotted as a means of identifying sites that appear as outliers from the modelled relationship[17].
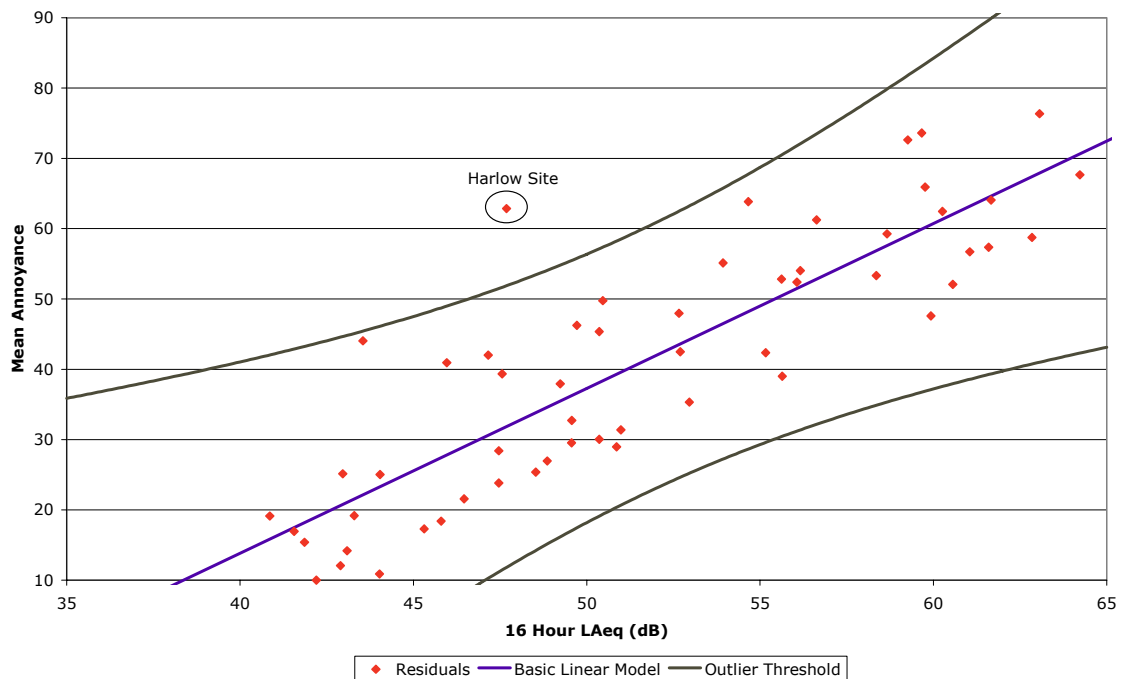


**Figure 8.1 Basic Linear Model of Mean Annoyance against LAeq**

8.2.8    Only one site has a much higher annoyance levels than expected: a site in Harlow, R17, about 19km from Stansted.  Of course, with 56 sites in the model, such a result is not unexpected (we would expect 5% of sites – ie about 3 – to lie outside the range of ± 2 standard errors).

---

[17] The curves identify the area within two standard errors of the modelled relationship at the mean values of LAeq, rising to three standard errors at the extremities.

**mva**consultancy

8.2.9    It should be remembered that the mean annoyance scores are based on (typically) 60 respondents at the full survey sites, and 15 respondents at the restricted survey sites.  As the restricted survey sites are based on less data, we can have less confidence in the mean annoyance scores.   Regressions were carried out with the full survey sites being given greater weights than the restricted survey sites, but these models were very similar to the unweighted models, with a smaller $R^2$ value, indicating a poorer fit to the data.

8.2.10   A characteristic of the estimated model is that for values of LAeq less than 38, the mean annoyance will be less than 10, which is not possible as the 'not at all annoyed' score was given a value of 10.  Similarly, for high values of LAeq (above 73), the model will predict that the mean annoyance is greater than 90 (extremely annoyed).    This is a drawback of the basic linear model.  It is therefore useful to consider different types of models which are constrained to low and high levels of mean annoyance.

### Logistic Model

8.2.11   One such type of model is the logistic model, which can be adapted to cater for general upper and lower bounds, in this case between 10 and 90.
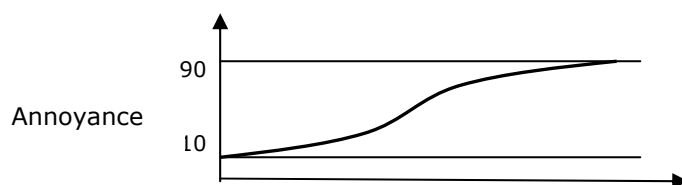


8.2.12   Non-linear least squares regression has been used to estimate the logistic model:

### Model 8.2          A = 10 + 80 / (1 + exp (7.32 − 0.13 x LAeq))

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq |
|---|---|---|---|---|
| 0.729 | 0.733 | 56 | 7.32 | -0.13 |
|  |  |  | (10.34) | (-10.01) |

8.2.13   The adjusted $R^2$ value for the logistic regression is only very slightly lower than the adjusted $R^2$ value of the basic linear model (a value of 0.733 compared to 0.734), indicating that the two models fit the annoyance data equally well.  Both models are shown in Figure 8.2.

mvaconsultancy
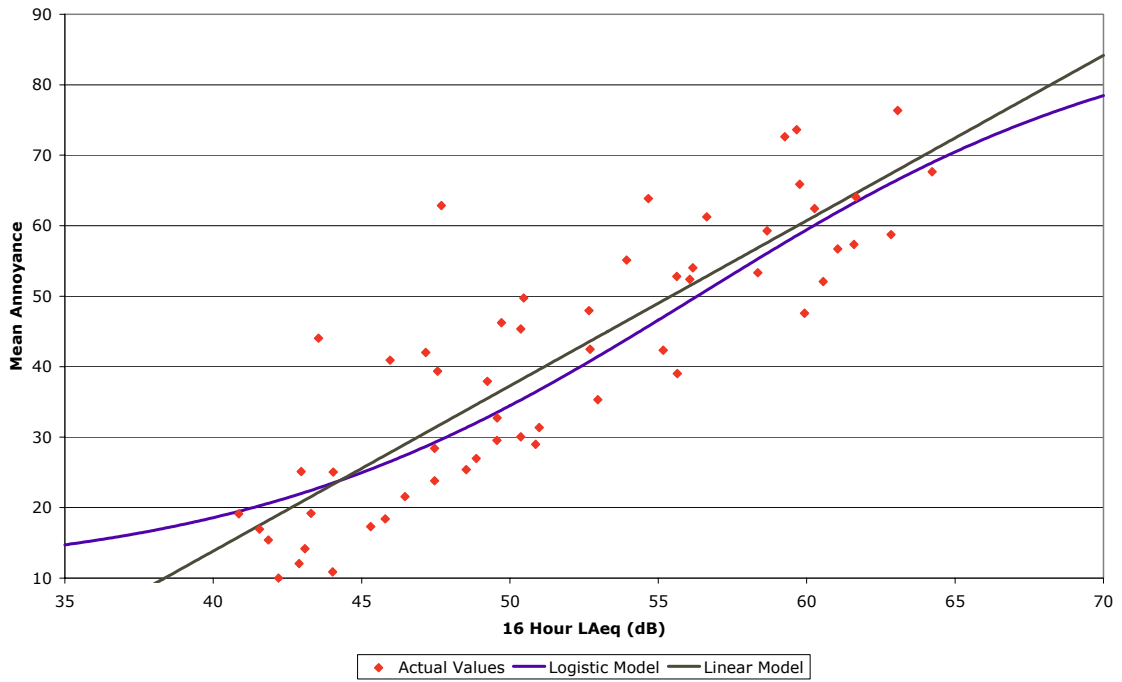
**Figure 8.2 Models for Linear and Logistic Regressions on LAeq**


### Piecewise Models

8.2.14    Piecewise models consist of continuous sections of linear models.  The first piecewise model considered was a relationship where annoyance is flat up to a certain level of LAeq, followed by a linear relationship as LAeq increases.  This form of relationship is shown below.
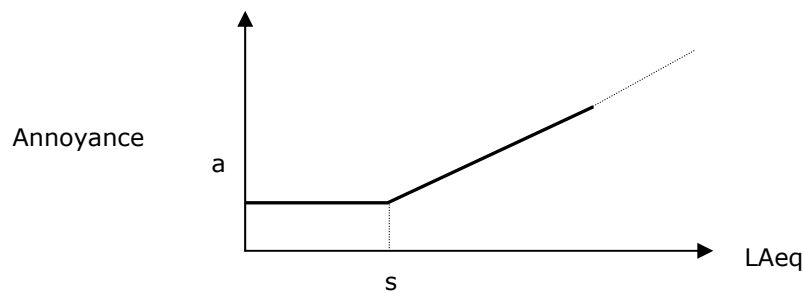


8.2.15    To determine the point where the regression line changes from a zero slope to a positive slope, a series of models was tested using different changeover points; the model corresponding to a change of slope at LAeq 42 produced the highest $R^2$ value.    The corresponding predictive model was Model 8.3.

**Model 8.3**      **Mean annoyance = 18.3**                      **for LAeq <= 42**

                     **Mean annoyance = 18.3 + 2.4 x (LAeq − 42)**      **for LAeq > 42**

| | R$^2$ | Adjusted R$^2$ | N | Intercept | LAeq - 42 |
|---|---|---|---|---|---|
| | 0.739 | 0.734 | 56 | 18.28 | 2.36 |
| | | | | (8.16) | (12.36) |

8.2.16   This model implies that the proportion of people reporting annoyance at very low levels of aircraft noise is insensitive to LAeq below LAeq 42.  The R$^2$ and slope of the LAeq term for this model is similar the basic linear model, Model 8.1.

8.2.17   This model was further expanded to include a cut-off at high values of LAeq.  Again, various changeover points were investigated, and a LAeq of 59 was found to be optimum, producing the highest R$^2$ value, as shown in Model 8.4.

**Model 8.4**      **Mean annoyance = 16.8**                      **for LAeq <= 42**

                     **Mean annoyance = 16.8 + 44.9/17 x (LAeq − 42)** **for 42<LAeq<59**

                     **Mean annoyance = 61.7**                      **for LAeq >= 59**

| | R$^2$ | Adjusted R$^2$ | N | Intercept | LAeq - 42 |
|---|---|---|---|---|---|
| | 0.713 | 0.707 | 56 | 16.80 | 44.85 |
| | | | | (6.35) | (11.24) |

8.2.18   The fit of this model, with the annoyance bounded at both higher and lower values of LAeq, is not as good as the fit of Model 8.3, where the annoyance is bounded for low values of LAeq only, with an adjusted R$^2$ of 0.707 compared to 0.734.  Therefore, although this model implies that people are insensitive to changes in sound level above LAeq 59, there is still a relationship between annoyance and sound level at higher LAeq values, although the relationship is weaker than the relationship below LAeq 59.

**mva**consultancy

8.2.19    Figure 8.3 shows both these piecewise models.
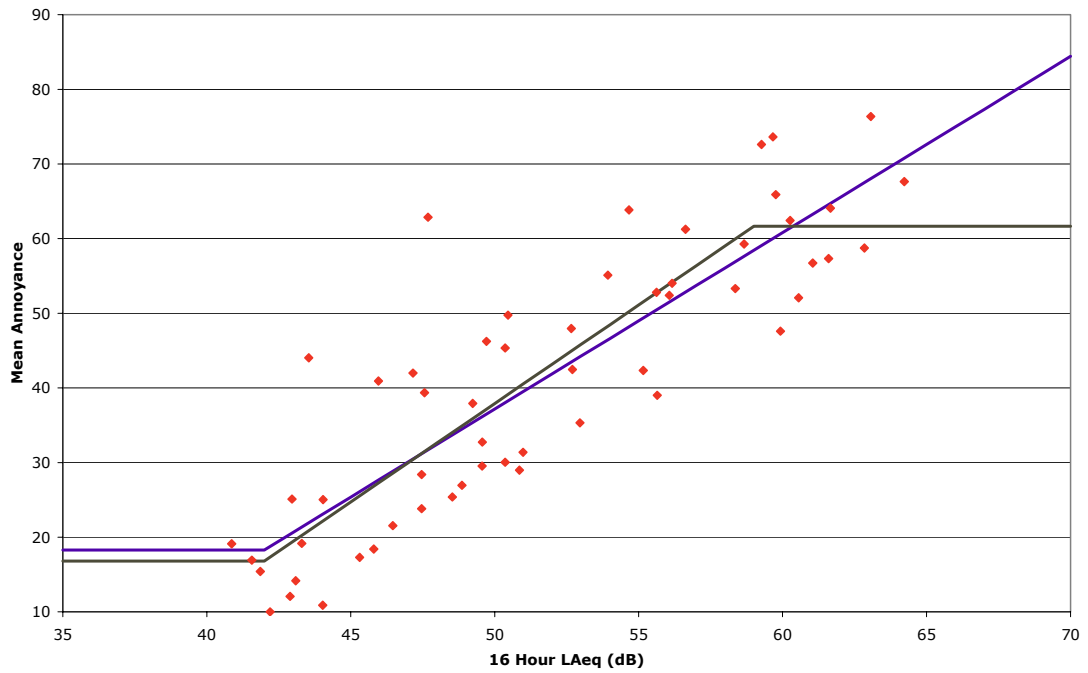


**Figure 8.3 Piecewise Linear Models**

8.2.20    The final piecewise model considered was a two-slope model, where a change in the gradient of the slope was allowed.  Several different values of LAeq were considered for the point of inflection, and the value which produced the best fit with the mean annoyance was 59 LAeq:

**Model 8.5**          **Mean annoyance = 60.2 - 2.5 x (59 - LAeq)          for LAeq <= 59**

**Mean annoyance = 60.2 + 1.1 x (LAeq − 59)          for LAeq > 59**

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq < 55 | LAeq > 55 |
|-------|----------------|----|-----------|-----------|-----------|
| 0.743 | 0.733 | 56 | 60.15 | -2.50 | 1.10 |
| | | | (22.59) | (-9.97) | (0.83) |

8.2.21    This model is illustrated in Figure 8.4.
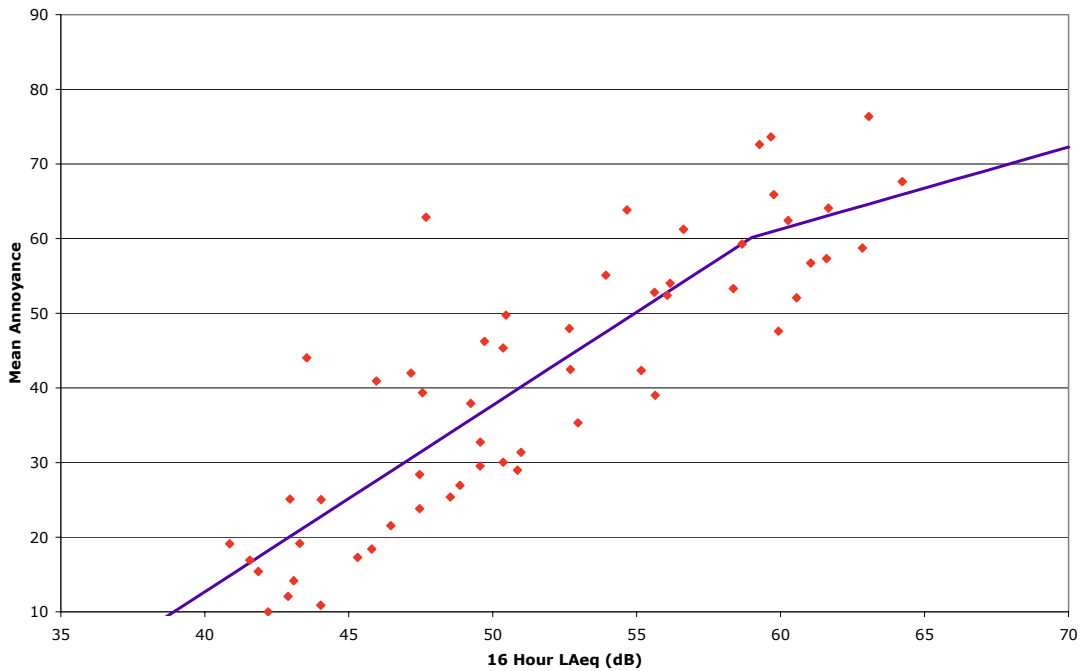


**Figure 8.4 Two-slope Model**

8.2.22    This model shows that for a given change in LAeq, the increase in annoyance is greater for values of LAeq less than 59 than for values of LAeq greater than 59.  This result supports the finding from the comparison of Model 8.3 with Model 8.4, that there is a positive relationship between annoyance and LAeq at higher LAeq values.

**Step Models**

8.2.23    In the ANIS study, step function models were considered as a way of investigating the possibility of a discontinuity in the slope of the relationship, which could be used to justify a sound level threshold for policy purposes.  In a step model, a strong relationship between reported annoyance and LAeq in the middle of the range is simplified into a single 'step' at one point, as illustrated below.
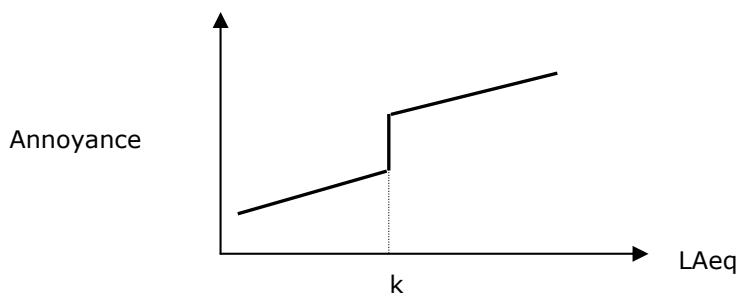


8.2.24    For the ANASE study, we tested step values at intervals of 1 dB for values of LAeq between 41 and 64 dB using model 8.6:

**Model 8.6**        **Mean Annoyance = a + b LAeq + c Step**

**Where Step =        0 for Leq < Step value**

**1 for Leq ≥ Step value**

8.2.25    To illustrate the results obtained, Table 8.1 shows the coefficients for every odd step value between 47 and 63.  There are no values where the step coefficient is statistically significant.  The mixture of positive and negative step values produced confirms the poor fit of these models.  The model statistics are comparable to all the step values tested.

**Table 8.1 Selected step model coefficients**

| Step    Value (LAeq) | $R^2$ | Adjusted $R^2$ | Intercept | LAeq | Step |
|---|---|---|---|---|---|
| 47 | 0.745 | 0.735 | -71.591 (-5.78) | 2.117 (7.59) | 4.664 (1.11) |
| 49 | 0.739 | 0.729 | -79.949 (-5.35) | 2.345 (7.12) | 0.002 (0) |
| 51 | 0.739 | 0.729 | -74.425 (-4.06) | 2.221 (5.65) | 1.889 (0.36) |
| 53 | 0.741 | 0.731 | -69.349 (-3.79) | 2.112 (5.45) | 3.643 (0.69) |
| 55 | 0.741 | 0.731 | -90.385 (-5.12) | 2.572 (6.93) | -3.677 (-0.71) |
| 57 | 0.739 | 0.730 | -84.850 (-5.54) | 2.449 (7.82) | -2.027 (-0.42) |
| 59 | 0.739 | 0.729 | -83.097 (-5.86) | 2.411 (8.39) | -1.452 (-0.31) |
| 61 | 0.743 | 0.733 | -85.789 (-7.38) | 2.467 (10.76) | -4.710 (-0.96) |
| 63 | 0.740 | 0.730 | -78.458 (-7.46) | 2.313 (11.35) | 3.219 (0.44) |

8.2.26    We can conclude that there is no statistical evidence for any discontinuity in the relationship between mean annoyance and LAeq.

**Summary of Models using LAeq alone**

8.2.27    In summary, all of the linear model formulations account for a similar variation in reported annoyance, although none improve on the adjusted $R^2$ (0.734) of the basic linear model (Model 8.1).

8.2.28    The basic linear model (Model 8.1) accounts for around three-quarters of the variation in reported annoyance between sites, with a coefficient on LAeq of around 2.3 - i.e. an increase in LAeq of 1 accounts for a 2.3 scale point increase in the mean annoyance score (with the mean annoyance score ranging between 10 and 90).

**mva**consultancy

8.2.29   A logistic model (Model 8.2) can be fitted to the data so that the model is constrained to the bounds of annoyance in the data, and this produces an almost identical fit to the basic linear model (in terms of adjusted $R^2$).

8.2.30   The piecewise models reveal features that are of interest, notably:

- At the lower end of the range of aircraft sound levels measured in LAeq, there is a region up to 42 LAeq where there is no apparent increase in reported annoyance as LAeq increases; and

- For a given change in LAeq, the increase in annoyance is greater for values of LAeq less than 59 than for values of LAeq greater than 59.

- These implied changes of slope at the lower and upper ends of the LAeq range support the use of the logistic form.

8.2.31   No threshold, or discontinuity, in the relationship between mean annoyance and LAeq was identified.

## 8.3   Adapting the Model to Take Account of Other Influences

8.3.1   There could be many reasons for the variation in reported annoyance between individuals at a single site facing a common aircraft noise environment, and between sites with similar aircraft sound levels as measured by LAeq.  These include:

- individual sensitivity to noise (expected to be random);

- lifestyle (including time spent at home);

- relationship with the airport (especially aviation-related employment);

- awareness with issues related to the airport (including knowledge of possible airport expansion);

- socio-demographic factors (age, gender, income, SEG – partly related to lifestyle);

- local environmental factors (especially other noise sources); and

- other aspects of aircraft noise (time of day pattern, aircraft type, sound level  versus number).

8.3.2   To explore the potential importance of these influences, we analysed the correlation between annoyance and other information obtained from individual survey respondents.

### Individual Level Analysis

8.3.3   Analysis using the statistical tool CHAID (see Appendix A8) showed that the annoyance level was most influenced by the sound level (LAeq) that respondents were exposed to.  Once differences in sound level had been taken into account, the following factors influenced individual respondents, such that those with the characteristics tended to show greater annoyance than those without the characteristics:

- living near Heathrow (with those living near Heathrow being generally more annoyed than those living near other airports);

- working from home (with those respondents who usually work from home being

mvaconsultancy

generally more annoyed);

■        gross household income (those with a greater household income being generally more annoyed than those with a lower household income);

■        being at home between 11am and 3pm on a typical weekday (those at home during this time being generally more annoyed than those not at home);

■        age-group (those aged less than 35 and over 65 being generally less annoyed with aircraft noise than those aged 35 to 64); and

■        SEG (those in SEG A and B generally expressing the greatest annoyance).

8.3.4    It is reasonable to expect that the non-noise factors indicated by the analysis at the individual level are likely also to influence mean annoyance at the site level.

8.3.5    Model 8.1 was the starting point for examination of additional variables at the site level, initially using indicators of the socio-demographic characteristics of the sites.  The aim was to improve the overall fit of the model with significant coefficients for the additional variables, and to obtain models that represented realistic behavioural relationships.  Further variables might also help account for the outliers.

8.3.6    Model 8.1 can be modified in two basic ways to introduce further variables:

■        by allowing these variables to adjust the intercept term, so that the model becomes

Mean Annoyance = a + b x LAeq + c x V

Where $V$ is the site average value of the new variable, and $c$ is its coefficient; and

■        by allowing these variables to adjust the slope of the curve, ie the coefficient of LAeq, so that the model becomes

Mean Annoyance = a + (b + c x V) x LAeq

Where $V$ and $c$ are as before.

8.3.7    Each variable was tested separately on the intercept and the slope (see Appendix A8).  In all cases, the coefficient of LAeq remained significant, but only three of the socio-demographic variables had significant coefficients:

■        working from home (those who work from home generally have a greater level of annoyance);

■        income (those who have a higher household income are generally more annoyed – measured either as average household income or as the proportion with an income more than £40,000); and

■        SEG (those in a higher SEG category are generally more annoyed – measured as either in SEG A or B, or in SEG A, B or C1).

8.3.8    Note that some of the variables (living near Heathrow, home in the afternoon and age) identified in the analysis at an individual level do not appear to be significant at the inter-site level, ie when site averages are used.  In the case of age, this arises because the average ages of the different sites were broadly similar, whereas the spread of ages amongst individual respondents was much greater.  Working at home and home in the afternoon can

mvaconsultancy

be expected to be strongly related.  We return to reconsider the Heathrow variable later in this section.

8.3.9   Using the variables identified above, a stepwise regression estimation was carried out, to identify the most significant combination of variables.  It is not possible to support both income and socio-economic group in the same model because of their high correlation.  The best fit was produced using the percentage working from home and income (Model 8.7).

**Model 8.7        Mean Annoyance =   -76.4 + 2.1 x LAeq**

**+ 43.8 x % Respondents who work from home**

**+ 0.2 x % Income greater than £40,000 x LAeq**

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq | % Work from home | % Income > £40k x LAeq |
|---|---|---|---|---|---|---|
| 0.818 | 0.808 | 56 | -76.36 (-9.02) | 2.15 (12.79) | 43.76 (3.08) | 0.25 (2.28) |

8.3.10   Examination of the outlier sites showed one notable difference from the other 55 sites in the survey: 43% of working respondents at the Harlow site worked from home.  The other sites range from 0% to 30% of working respondents working from home, with an average of 8%.  It therefore appears that the Harlow site, which consists of 15 respondents, has had an impact on the regression, causing the variable "% respondents who work from home" to enter the equation.

8.3.11   The stepwise procedure was therefore repeated but excluding the outlier site at Harlow.  This resulted in Model 8.8 with income being the most significant additional variable:

**Model 8.8      Mean Annoyance = -79.5 + (2.24 + 0.33 x % income>£40k) x LAeq**

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq | % income > £40k x LAeq |
|---|---|---|---|---|---|
| 0.825 | 0.818 | 55 | -79.47 (-9.64) | 2.24 (13.77) | 0.33 (3.39) |

**mva**consultancy

8.3.12   However, since income greater than £40,000 does not take the full range of incomes into account, and cannot be readily examined for changes over time (see Chapter 9), this model was re-estimated using mean household income.  This produced Model 8.9, which has a slightly lower, but comparable, adjusted $R^2$ value:

**Model 8.9     Mean Annoyance = -79.4 + (2.14 + 6.61x10$^{-6}$ x mean income) x LAeq**

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq | Mean household income x LAeq |
|---|---|---|---|---|---|
| 0.787 | 0.782 | 55 | -79.44 (-9.53) | 2.14 (12.07) | 6.61 x 10$^{-6}$ (3.19) |

8.3.13   The $R^2$ values cannot be compared directly with the earlier models because one outlier site has been taken out; however the coefficient of LAeq is similar to the basic linear model (2.1 against 2.3).  The model is intuitively plausible and fits the remaining sites reasonably well.  Figure 8.5 shows the plot of residuals for Model 8.9 (a positive residual indicates that the observed annoyance is higher than the modelled annoyance and vice versa).  These show an acceptably random distribution, although there is one site near Southampton airport with a comparatively large positive residual, Hillbrook, a full survey site with a high proportion of respondents in SEG A/B/C1.
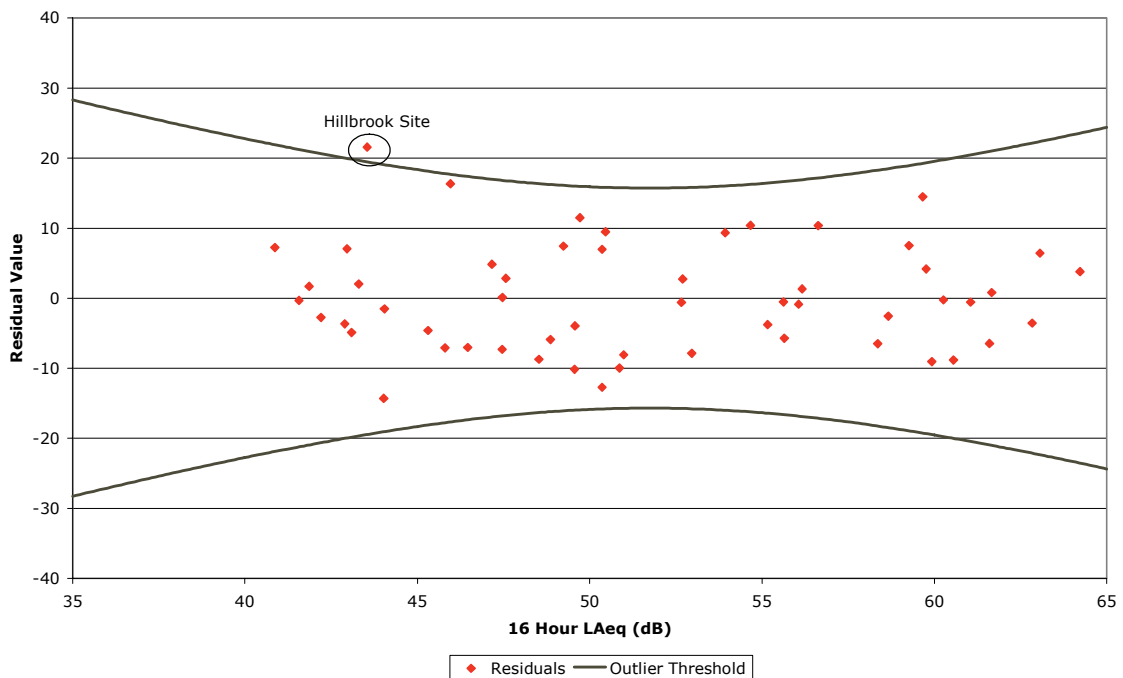


**Figure 8.5          Model 8.9 Residuals Plotted Against LAeq**

8.3.14   A variable distinguishing Heathrow was not significant when all sites were included in the model, but with the exclusion of the site at Harlow, Heathrow becomes significant, although its significance level is lower than both mean income and SEG. A model including income or SEG does not support the inclusion of Heathrow as a further significant variable. The model including LAeq and Heathrow only is shown in Model 8.10.

**mva**consultancy

**Model 8.10    Mean Annoyance = -84.3 + 2.36 x LAeq + 5.55 x Heathrow**

Where Heathrow =     1 if site is close to Heathrow airport
                     0 otherwise

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq | Sites close to Heathrow |
|---|---|---|---|---|---|
| 0.810 | 0.803 | 55 | -84.27 (-9.88) | 2.36 (14.44) | 5.55 (2.53) |

8.3.15    Model 8.10 shows that sites close to Heathrow airport have an annoyance score approximately 5 points, ie a quarter of an annoyance scale band, higher than sites which are close to other airports.  This could be for a number of reasons, including both the socio-demographic make-up and income levels of people living close to Heathrow airport, and the type of air traffic, for example the much larger number of aircraft using Heathrow airport. The effect of the number of aircraft on the level of annoyance is discussed in Chapter 9.

## 8.4    Summary of Main Points

- The LAeq metric is effective at explaining much of the reported variation in annoyance;

- At the site level, comparison of a basic linear model with other models show that the goodness of fit is similar for different types of model;

- Based on comparing values of adjusted $R^2$, of the linear forms tested, the basic linear model is to be preferred;

- The basic linear model indicates a coefficient on LAeq of around 2.3, ie an increase in LAeq of 1 accounts for a 2.3 scale point increase in the mean annoyance score (with the mean annoyance score ranging between 10 and 90);

- A logistic model – constrained to the bounds of annoyance in the data – produces an almost identical fit to the basic linear model.  Given its asymptotic properties, this form of model is superior to the basic linear model;

- No threshold, or discontinuity, in the relationship between mean annoyance and LAeq was identified;

- Other variables, not related to noise, improve the fit of the model: working from home, income, and SEG;

- Excluding a single outlier site, the best model for mean annoyance contains the variables: LAeq and income; and

- Once LAeq and income had entered the model, no further location effect, eg living near Heathrow, was significant.

**mva**consultancy

# 9    Change in Attitudes to Aircraft Noise

### 9.1    Introduction

9.1.1    In order to assess changes in attitudes to aircraft noise over time, we need information from comparable studies which have recorded the relevant sound level (dose) and response (annoyance) quantities at different points in time.  ANASE provides information for 2005: there is only one comparable study – ANIS conducted in 1982.

9.1.2    This chapter examines differences in responses between ANIS and ANASE, through comparing the results of regression analysis on data from the two surveys.  The chapter begins by considering the LAeq metric and its relationship with reported annoyance in both the ANIS and ANASE studies.  The chapter goes on to consider the possibility of other measures for relating the effect of sound level and number of aircraft to mean annoyance.

### 9.2    ANIS Characteristics

9.2.1    In comparing results from the two surveys, we need to bear in mind that the surveys differed in some important aspects.  ANASE was never intended to be an exact replica of ANIS, and the ANASE terms of reference differ from ANIS in significant respects.  In particular, the ANASE survey design benefited from the general development of survey methodology over time, from research carried out elsewhere in the intervening years, and from an extensive piloting phase, much of which was concerned with the conduct of the SP aspects of ANASE, an element not covered by ANIS.  Some important differences between the two studies are shown in Table 9.1.

**Table 9.1 Differences between for ANIS and ANASE**

| ANIS - 1982 | ANASE - 2005 |
|---|---|
| Sites pre-selected based on sound level and number of aircraft | Sites randomly selected, stratified by sound level and number of aircraft |
| 23 sites – all but 4 around Heathrow | 76 sites around 16 airports |
| Leq range: 51 to 72 dB | Leq range: 40 TO 64 dB |
| Households selected at random | Households and individual within household selected at random |
| Main annoyance question used four point scale plus "not heard" | Main annoyance question used five point ISO scale |

9.2.2    Like ANASE, the ANIS sites were selected to include a full range of aircraft noise environments in terms of both 'noise' and 'number'.  However, unlike ANASE, individual sites were pre-selected.

**mva**consultancy

9.2.3   The most important difference between ANIS and ANASE was in the recording of reported annoyance.


9.3    **Reconciling Reported Annoyance between ANIS and ANASE**

9.3.1   In terms of annoyance, ANIS respondents were asked " What are the different kinds of noise you hear around here?", with a prompt question if aircraft noise was not mentioned spontaneously – "Do you ever hear aircraft fly by here?".  If respondents reported hearing no noise, no annoyance questions were asked.  Respondents who did hear aircraft were asked "…how much the noise of aircraft here bothers or annoys you?", with responses recorded on a scale of very much, moderately, a little, not at all annoyed and not heard.  A "don't know" category was also offered, but only two respondents in the whole survey gave this response.

9.3.2   Table 9.2 compares the two response scales for the noise annoyance question.

### Table 9.2 Annoyance Scales for ANIS and ANASE

| ANIS Scale | ANASE Scale |
|---|---|
| Very much annoyed | Extremely annoyed |
| Moderately annoyed | Very annoyed |
| A little annoyed | Moderately annoyed |
| Not at all annoyed | Slightly annoyed |
|  | Not at all annoyed |

9.3.3   To compare responses to the annoyance questions shown in Table 9.2 we needed to devise a comparable scale.  It would have been possible to match the ANIS and ANASE scales by giving the same score at the points where the same semantic descriptions are used in both surveys ("moderately annoyed" and "not at all annoyed").  However, we concluded this would not provide a comparable result as the "same" descriptions are at different points on the two scales: "moderately annoyed" is the second highest category in the ANIS scale, but the third highest category in the ANASE scale.

9.3.4   To reconcile the two scales, the latest advice has been taken, following international best practice[18], which scores the categories in equally spaced intervals between 0 and 100.  This results in some verbal descriptions being scored differently in the two surveys (for example "moderately annoyed" is scored as 62.5 in ANIS and 50 in ANASE), but does take account of the different numbers of categories used (the "not heard" category in ANIS has been grouped with the "not at all annoyed" category).  The weighting also uses all the response data in each survey.  Table 9.3 shows the scoring system we have adopted.

---

[18] This scoring system follows analysis carried out by Miedema and Oudshoorrn: 'a set of annoyance categories divides the range from 0-100 in equally spaced intervals'

**mva**consultancy

**Table 9.3 Weights used to calculate mean annoyance for ANIS and ANASE**

| ANIS Scale | | ANASE Scale | |
|---|---|---|---|
| Very much annoyed | 87.5 | Extremely annoyed | 90 |
| Moderately annoyed | 62.5 | Very annoyed | 70 |
| A little annoyed | 37.5 | Moderately annoyed | 50 |
| Not at all annoyed | 12.5 | Slightly annoyed | 30 |
| | | Not at all annoyed | 10 |

9.3.5   Table 9.3 gives what we believe to be the most appropriate set of weightings to reconcile the two datasets.  If it looks "precise", that is because we need to give a specific weight to each category to enable a calculation of mean annoyance.  There is a degree of uncertainty attached to the weights and, although there is no single best way to reconcile the datasets, we believe we have followed international best practice recommendations.  Sensitivity analysis of the weighting is described in Section 9.5.

9.3.6   Figure 9.1 shows the values of mean annoyance calculated from ANIS and ANASE plotted against LAeq.  The ANASE points are, of course, identical to those shown in Figure 8.4.   For a given value of LAeq the ANIS points are generally below those from ANASE.
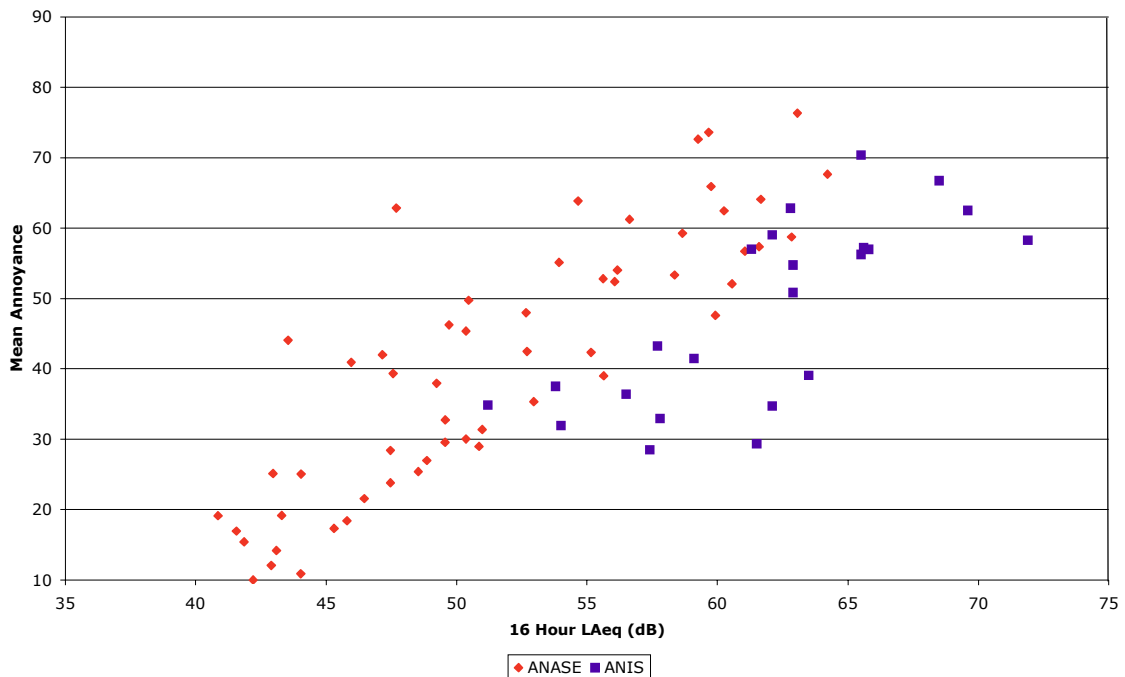


**Figure 9.1  Mean Annoyance against LAeq for ANASE and ANIS**

### 9.4     Relationship between Annoyance and LAeq

9.4.1     Using the scoring system described in the preceding section, a linear regression for the relationship between mean annoyance and LAeq for the 23 sites included in the ANIS survey gives Model 9.1.

**Model 9.1**       **Mean Annoyance = -68.9 + 1.9 x LAeq**

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq |
|---|---|---|---|---|
| 0.559 | 0.538 | 23 | -68.93 | 1.89 |
|  |  |  | (-3.03) | (5.16) |

(The equivalent formula for ANASE, as discussed in Chapter 8, is:

Model 8.1       Mean Annoyance = -80.0 + 2.3 x LAeq       $R^2 = 0.739$)

9.4.2     As the ANIS data is based on a different number of points, it is not possible to compare the $R^2$ values directly for the two studies. 56% of the variation in the ANIS annoyance data can be explained by LAeq alone, compared to 74% for ANASE. The slope of Model 9.1, at 1.9, is less than that in ANASE (2.3). Although the differences are not statistically significant, they do suggest that for a given change in LAeq, respondents in ANASE reported a bigger change in their level of annoyance than respondents in ANIS.

9.4.3     The ANASE model is fitted to more data points over a greater range of LAeq values than ANIS, and shows a better fit.

9.4.4     Figure 9.2 shows the mean annoyance against LAeq for both ANIS and ANASE, along with the basic linear model lines.
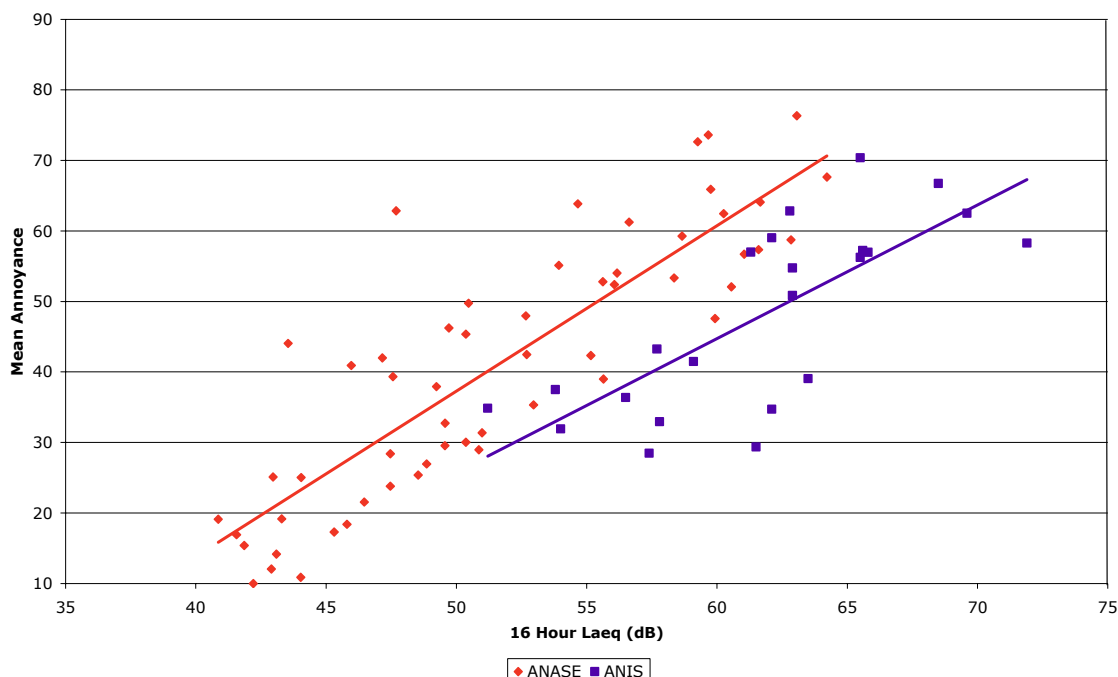


**Figure 9.2   Models of Mean Annoyance against LAeq for ANASE and ANIS**

**mva**consultancy

9.4.5   It is clear that the mean annoyance values for ANASE are generally higher for a given LAeq than they are for ANIS.  For a LAeq of 57 (identified in the DORA report as the onset of significant annoyance), the modelled value of annoyance for ANIS is 39 (slightly higher than "a little annoyed" on the ANIS scale), whereas for ANASE it is 53 (somewhat higher than "moderately annoyed" on the ANASE scale).

9.4.6   To examine the possibility that the difference may be due to the lower sound level sites that are not in ANIS, the basic linear model for ANASE has been estimated for sites which have a LAeq value greater than 50 only.  This produces the following model:

**Model 9.2          Mean Annoyance = -74.6 + 2.2 x LAeq**

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq |
|---|---|---|---|---|
| 0.564 | 0.548 | 30 | -74.63 | 2.25 |
|  |  |  | *(-3.50)* | *(6.02)* |

9.4.7   Figure 9.3 illustrates Model 8.1 – based on the full set of ANASE data – and models 9.1 and 9.2, the ANIS model and the ANASE model fitted over the same range of LAeq values.  The ANASE lines are virtually identical.  For the ANASE sites in areas with a LAeq greater than 50, half of the variation in mean annoyance can be explained by LAeq alone – almost the same as the ANIS level of explanation (Model 9.1 has an adjusted $R^2$ value of 0.538).
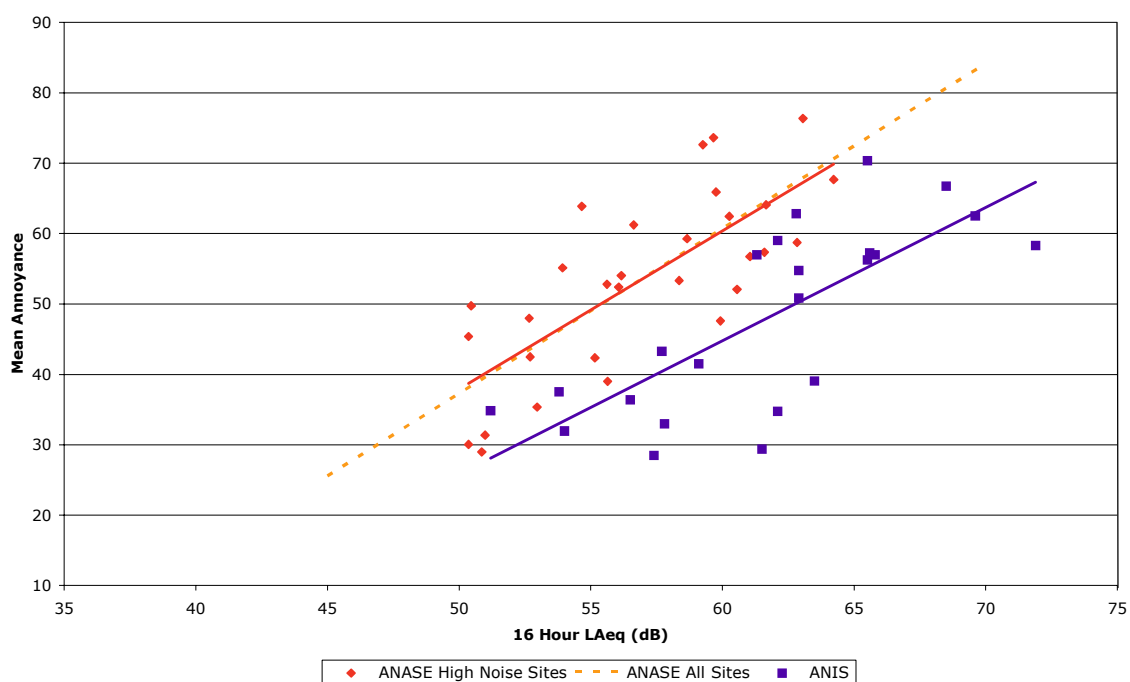


**Figure 9.3 Mean Annoyance against LAeq for ANIS and ANASE for Higher sound level Sites**

9.4.8   This therefore reinforces the conclusions.  The linear models shown in Figure 9.3 are reasonably parallel (the ANASE model has a slope of 2.25 and the ANIS model a slope of 1.89), suggesting that when confining the data to the higher sound level range, there is a similar relationship of increasing annoyance as sound level increases, but the level of annoyance is consistently about 14 points greater in ANASE than it was 23 years ago.

**mva**consultancy

9.4.9    For a given mean annoyance score, the 16-hour LAeq value in the ANASE survey is lower by between 3 dB (for a mean annoyance score of 10 or not at all annoyed) and 11 dB (for a mean annoyance score of 90 or extremely annoyed).  A modelled mean annoyance of 50 is at 63 dB in ANIS and at 55 dB in ANASE, a different of 8 dB.

9.4.10   In support of the differences revealed by the ANASE/ANIS comparison, recent preliminary results reported by the EU-funded HYENA study[19] are of interest.  The study collected data from 4816 respondents who had been living near one of six major European airports (Heathrow was the UK airport surveyed) for at least 5 years.  The sample areas were chosen to maximise the range of exposure to aircraft and road noise and to exclude areas exposed to other noise sources.  Fieldwork took place in the period 2003 – 2005.

9.4.11   Generally, the Integrated Noise Model (INM) was used to estimate aircraft noise exposure, based on radar tracks.  In the UK, the ANCON model was used.  Noise annoyance was assessed using the 10 point ICBEN scale.

9.4.12   The study estimated relationships between noise level and annoyance from aircraft and road noise and compared the results with standard EU curves derived by Miedema et al.  Although the study results for aircraft noise are influenced by high levels of reported annoyance from respondents living around Athens and Malpensa airports, when results from these two countries are excluded, annoyance ratings due to aircraft noise are higher than predicted from the EU curves.

9.4.13   The HYENA study concludes "The data supports other findings suggesting that people's attitude towards aircraft noise has changed over the years".


9.5    Sensitivity Analysis

9.5.1    The analysis reported above relates mean annoyance to LAeq.  There is uncertainty associated with the measurement of both quantities and to test the robustness of the difference in the ANIS and ANASE outcomes we conducted a series of sensitivity tests.

9.5.2    In these tests, we took no account of uncertainties associated with the ANIS data and instead, accepted the ANIS results as given.  We then concentrated on examining the sensitivity of the ANASE models to a range of assumptions.  The details of these tests are shown in Appendix A9.  We summarise the findings in the following sections.

Sensitivity to uncertainty in LAeq estimates

9.5.3    Subsequent to the ANASE analysis, alternative estimates of LAeq for sites around Heathrow were made available, produced by the CAA's model ANCON, based on Summer 2005 data.

9.5.4    The largest differences between the two sets of LAeq values were predominantly at lower noise sites, where it is acknowledged that measurement becomes progressively less accurate below 57dBA.

9.5.5    We estimated a regression model where the Heathrow data from ANASE was replaced with

---

[19] Associations between Traffic Noise, Aircraft Noise and Noise Annoyance.  Preliminary Results of the HYENA Study.  Paper presented to the 19th[th] International Congress on Acoustics, Madrid, 2-7September 2007

mvaconsultancy

the CAA values and compared the coefficients of that model with Model 8.1.  The differences in the model coefficients were not statistically significant.

### Sensitivity to uncertainty in Mean Annoyance

9.5.6   We stated above (para 9.3.4) that there is no single best way to reconcile the annoyance scales used in the two surveys. We examined how much of the difference between the models derived from the two datasets resulted from the weightings applied to the different annoyance responses (see Table 9.2).   When different weights were investigated, for example giving "Moderately annoyed" the same score in each survey, the difference between the ANASE and ANIS regression lines increased.

9.5.7   Concern had been expressed that responses to the annoyance question in ANASE may have been subject to response bias, whereby respondents gave an exaggerated response to their level of reported annoyance.  Such a systematic exaggeration may have resulted in ANASE showing a greater difference in reported annoyance for a given level of LAeq when compared with ANIS.

9.5.8   We explored the issue of response bias in the Phase One piloting work and were confident that it was **not** having a seriously distorting effect on responses.  The justification for this view is given in Appendix A9.  However, to investigate the possible presence and scale of an effect, we conducted sensitivity analysis by estimating regression models from different groups of respondents who, it could be argued, might have been subject to different and possibly biasing influences. The subgroups examined were:

- Those reporting awareness of any recent comments or articles in the newspapers or on TV concerning the local airport;

- Those who said (in response to the contingent valuation questions, discussed in more detail in Chapter 10), that they would not be prepared to pay anything at all to remove aircraft - "protest" voters; and

- Those who were surveyed at restricted sites.[20]

### Sensitivity Analysis Results

9.5.9   Table 9.4 summarises the regression models produced through various adjustments to the ANASE data, along with the modelled mean annoyance value with a LAeq of 57 dB, and the difference in mean annoyance points implied from the ANIS regression model.

---

[20] The ANASE survey was carried out at 56 sites where LAeq can be derived.  36 of these sites were full survey sites, where 60 responses were obtained, including responses to the SP questions, which required the use of the noise playback equipment.  The remaining 20 restricted sites had no SP and therefore no noise playback equipment was used. It is therefore possible to compare the full and restricted sites to seek to identify a possible "equipment" effect in the responses

**mva**consultancy

**Table 9.4  Regression Models Produced by Variations of ANASE Datasets**

| Model No[21]. | Description of Sensitivity Test | Regression Model | Modelled Annoyance at Leq 57 | Diff. from original ANASE | Diff. from ANIS |
|---|---|---|---|---|---|
| 9.1 | ANIS | A = -68.9 + 1.89 x LAeq | 38.83 | | |
| 8.1 | ANASE (central estimates ) | A = -80.0 + 2.34 x LAeq | 53.38 | | 14.55 |
| A9.1 | ANIS with revised weights | A = -65.8+ 1.72 x LAeq | 32.36 | 21.02 | -6.47 |
| A9.2 | ANASE with CAA LAeq estimates at Heathrow | A = -87.4 + 2.45 x LAeq | 52.25 | -1.13 | 13.42 |
| A9.3 | ANASE for those unaware of local airport issues | A = -79.2 + 2.30 x LAeq | 51.9 | -1.48 | 13.07 |
| A9.4 | ANASE for non-protest voters | A = -78.5 + 2.29 x LAeq | 52.03 | -1.35 | 13.20 |
| A9.5 | ANASE with restricted sites adjustment | A = -69.4 + 2.07 x LAeq | 48.59 | -4.79 | 9.76 |
| A9.6 | ANASE – with CAA LAEq estimates at Heathrow and restricted sites adjustment | A = -73.3 + 2.09 x LAeq | 45.83 | -7.55 | 7 |

9.5.10   Using the ANASE annoyance and LAeq values indicate that the mean annoyance is about 14.5 points higher in ANASE than ANIS for a LAeq of 57.  When "Moderately annoyed" and "Not at all annoyed" are given the same score in each survey, with the remaining ANIS scores adjusted appropriately, the difference between the ANASE and ANIS regression lines increases, and ANASE is 21 points higher than ANIS for an LAeq of 57.

9.5.11   If the CAA LAeq estimates are used for the Heathrow sites, then this difference is reduced slightly to 13.4 mean annoyance points (though the differences in the coefficients in models 8.1 and A9.1 are not statistically significant).

---

[21] Details of the models with an A prefix can be found in Appendix A9

**mva**consultancy

9.5.12   The effect from respondents aware of local airport issues and those who gave a "protest" vote in the contingent valuation question is small, with a difference in mean annoyance of 1.1 and 1.5 respectively from the core ANASE model, and the differences are not statistically significant.

9.5.13   With an assumed "equipment" effect causing a response bias of 5.6 mean annoyance points for the full survey sites in ANASE, then the difference between ANASE and ANIS reduces to 10 mean annoyance points at a LAeq of 57. The model coefficient representing the "equipment effect" is not statistically significant.

9.5.14   If the ANASE data are adjusted to take account of two of these effects (using the CAA LAeq estimates for Heathrow sites, and modelling the restricted sites with a dummy variable for an "equipment" effect), the modelled mean annoyance values for ANASE are reduced further so that for a LAeq of 57, the mean annoyance derived from the adjusted ANASE data is 7 points greater than the modelled ANIS mean annoyance score.

9.5.15   Here the coefficient of the variable representing the "equipment" effect is significant at the 95% level. This is largely due to the effect of the low noise sites, where the ANASE LAeq estimates have a tendency to be marginally lower than the CAA estimates.

9.5.16   The results from the sensitivity tests confirm the conclusion that at similar levels of aircraft noise measured in LAeq, people are more annoyed in 2005 than they were in 1982, though the size of the difference is affected by assumptions made.

9.5.17   A particular issue affecting the size of the difference is whether the introduction of sound event playback equipment into the respondent's home generated an exaggerated response to the annoyance rating question. There is no statistical support for that effect on its own from the ANASE data; substitution of the CAA LAeq estimates at Heathrow which are not significant on their own produces a significant combined effect.

9.5.18   The range of difference in the ANIS and ANASE scores at 57 LAeq is from 7 to 21. The difference produced from the comparison of the central ANASE model and ANIS is 14.55, very close to the mid-point of the range.


### 9.6    Exploring the differences between ANASE and ANIS

9.6.1   We examined how much of the difference between the models derived from the two datasets resulted from the weightings applied to the different annoyance responses (see Table 9.2). When different weights were investigated, for example giving "Moderately annoyed" the same score in each survey, the difference between the ANASE and ANIS regression lines increased further.

9.6.2   In order to seek to explain the differences between ANASE and ANIS, we can conjecture two extreme hypotheses:

■    H1: LAeq is the appropriate measure, and people really are more annoyed by a given sound level now than in the early 1980s;

■    H2: LAeq is not the appropriate measure, and annoyance in both studies would correlate better with another measure of aircraft sound levels.

mvaconsultancy

9.6.3   Of course, either or both of these hypotheses may be only partly valid, but it is useful to begin by exploring their implications.

## 9.7     H1: LAeq is the appropriate measure

9.7.1   Taking the first hypothesis, we saw from the models described in Chapter 9 that annoyance was a function of both LAeq and mean income (Model 9.8).   Real incomes have grown substantially in the 23 years between the surveys (households' disposable income has increased by 58% in constant prices) and this income effect could account for some of the difference in annoyance seen in the comparison.

9.7.2   The income effect in Model 9.8 was derived from the sample taken at a single point in time - that is, it shows the **cross-sectional effect** of income on reported annoyance.  The value of the elasticity of annoyance with respect to income for the ANASE sites is 0.25.  There may be considerable differences between cross-sectional and temporal elasticities.

9.7.3   If we want to look at changes **over time,** we need evidence of a time series effect of income on reported annoyance.  Whilst the literature provides material showing time series results for noise **valuation,** we are not aware of any evidence available for changing sensitivity to sound levels over time as a function of income[22].

9.7.4   In a study to develop WebTAG guidance on the valuation of transport-related noise, Nelthorpe, Bristow and Mackie[23], in agreement with DfT, set the time-series elasticity equal to the cross-sectional elasticity as an interim measure.  If we make the same assumption, we can see what effect rising incomes have on the ANASE relationship.

9.7.5   Figure 9.4 shows the regression lines with mean income of £28,000 (the mean household income value for respondents in the ANASE study), and with a mean income of £17,700, which is the equivalent (in real terms) of £28,000 in 2005 in 1982, allowing for the 58% growth discussed above.

---

[22] There is no income data collected in the ANIS study

[23] Developing Guidance on the Valuation of Transport-related Noise for Inclusion in WebTAG.  Paper presented to the Seminar "Valuing Transport related Noise" held at the Department of Transport, May 2005
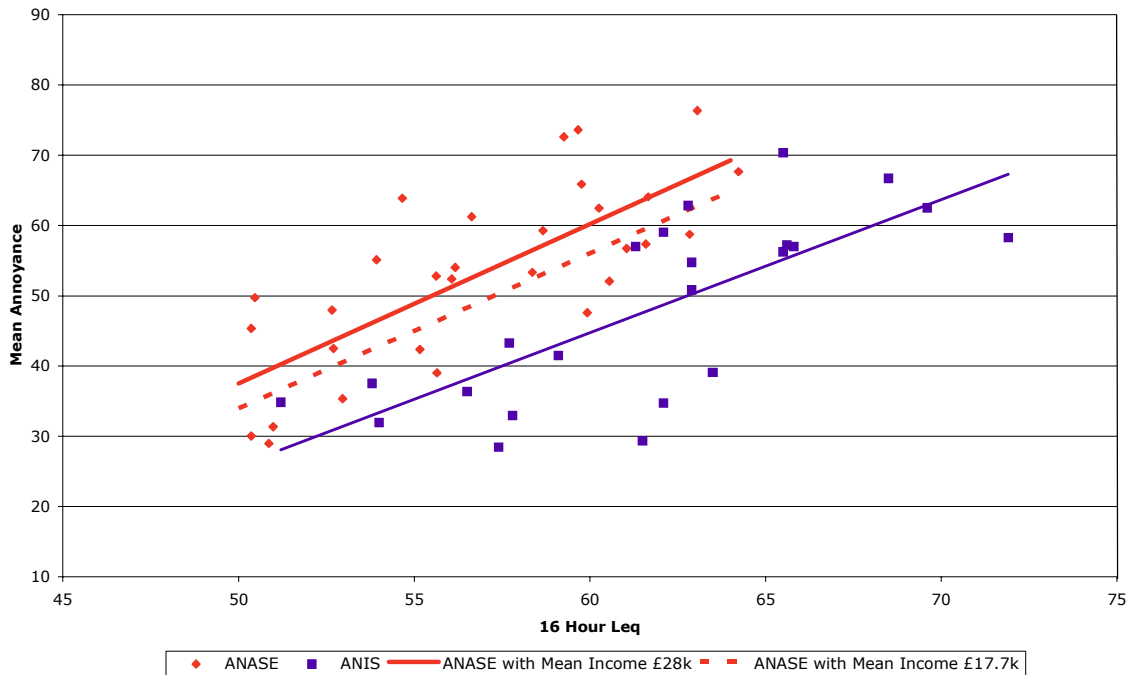
**mva**consultancy

**Figure 9.4  Effect of Mean Income**

9.7.6    The difference between ANASE and ANIS, having made allowance for rising income, is still large.

9.7.7    The remaining difference could be considered to be the result of changing tastes.  Society's level of tolerance to, and expectation of, acceptable living conditions is likely to have changed over time.  Likewise, people's willingness to be more openly critical of officialdom and government policy has increased.  Although such changes are difficult to quantify, it could be conjectured that the differences between the ANIS and ANASE results are the result of a combination of increasing income and a "taste effect", whereby people have become less tolerant of intrusive noise over time[24].


**9.8    H2: LAeq is not the appropriate sound level metric**

9.8.1    The measure used to relate annoyance to aircraft sound levels in the analysis so far is LAeq.  As can be seen from the description of LAeq in Chapter 2, it is a function of two components: number of aircraft and their sound exposure level – SEL.  The relative weight given each component in the LAeq formulation is fixed.  In order to explore whether LAeq is appropriate, we investigated the relationship between reported annoyance and the separate components of LAeq.

9.8.2    In addition to LAeq, the average sound level, Lav, and average number, Nav, of aircraft at a site have been calculated for values of LAmax over 65 dB in ANASE and over 67 dB in ANIS.  Figure 9.5 shows the site values of Lav and Log Nav for ANIS and ANASE.

---

[24] Social trends data indicate that there has been a five-fold increase in the number of complaints about noise received by Environmental Health Officers between 1981 and 1996/97.  Table 11.6, Social Trends 30, Office for National Statistics

**mva**consultancy
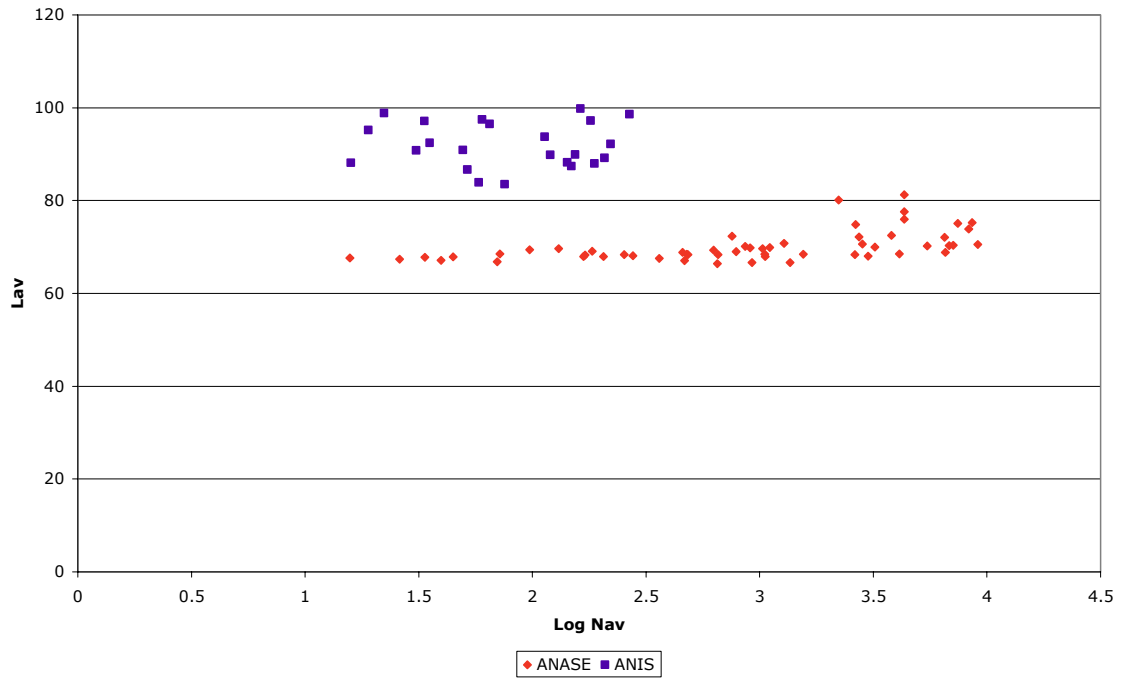
**Figure 9.5  Lav against Log Nav for ANIS and ANASE**

9.8.3    It can be seen that **there is no overlap whatsoever between the data points from the two surveys**[25]**.**   For the ANIS study, undertaken in the early 1980s, there were generally fewer aircraft, but the average sound levels were higher.  In the more recent ANASE study, there is a greater number of aircraft, but average sound levels were lower.

9.8.4    The difference in the mean annoyance against LAeq between the studies could therefore be related to the differences in the patterns of aircraft sound levels that are currently experienced.

9.8.5    Figure 9.6 shows the relationship between mean annoyance and Lav for both the ANIS and ANASE survey.

---

[25] It should be noted that this is not a fault of the sample design.  Because of changes that have taken place between 1982 and 2005, it was not possible to test the original ANIS findings under exactly similar aircraft noise exposure conditions.

**mva**consultancy
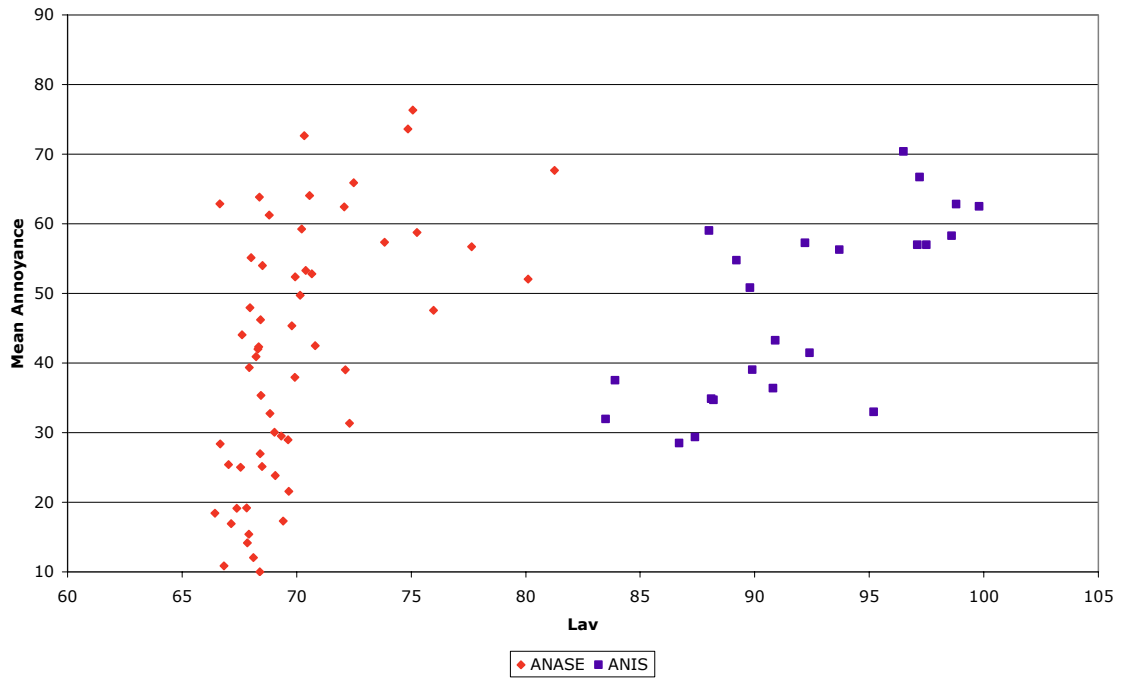
**Figure 9.6  Mean Annoyance against Lav for ANIS and ANASE**

9.8.6    While, for ANIS, there is a fairly clear relationship of increasing annoyance as Lav increases, for ANASE the relationship is less well defined: a large number of sites with lower values of Lav display a wide range of mean annoyance.

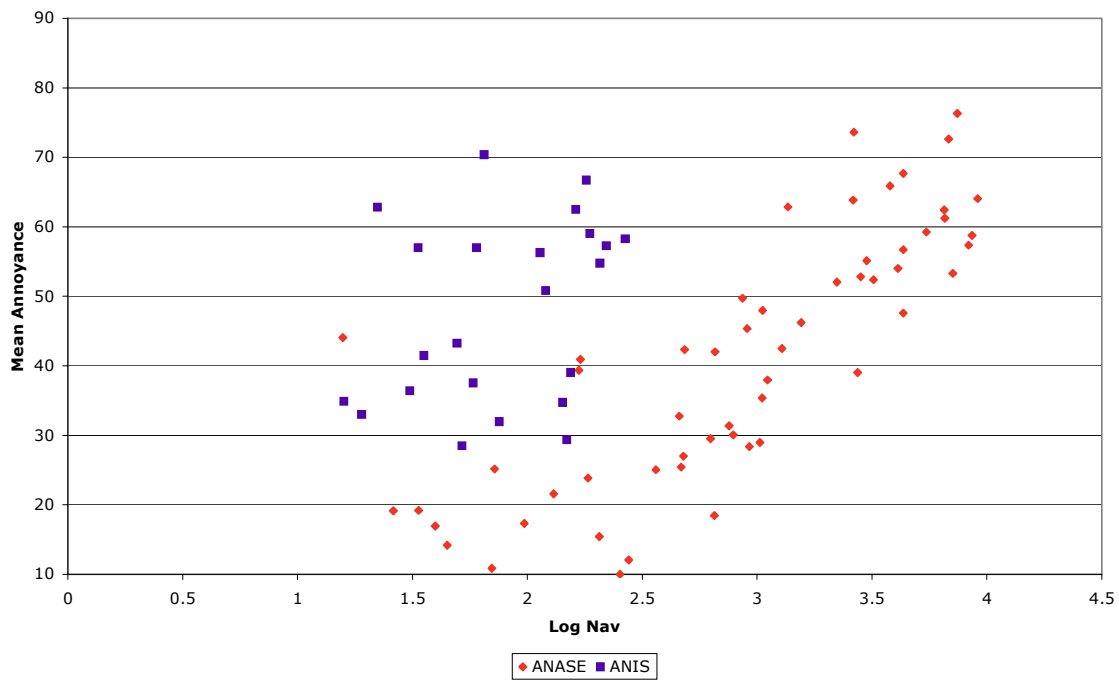9.8.7    Figure 9.7 shows a similar plot of mean annoyance against log Nav.



**Figure 9.7  Mean Annoyance against Log Nav for ANIS and ANASE**

mvaconsultancy

9.8.8   Here the opposite can be seen.  For this comparison, while ANASE shows a clear relationship of increasing annoyance as the number of aircraft (expressed as log Nav) increases, for ANIS this relationship is not apparent; instead a wide range of annoyance is reported for the range of log Nav recorded.

9.8.9   Therefore, between the ANIS and ANASE surveys, it appears that there has been a shift in the relative importance of the two components of annoyance: the sound level of the aircraft and the number of aircraft.

9.8.10  To explore the relationships further, regression analysis was undertaken of the form

Mean Annoyance = a + b x Lav + c x log Nav

9.8.11  For this regression on the ANASE data (Model 9.3), a ratio, c/b, of 21 produced the best match with the mean annoyance.  A similar analysis on the ANIS data (Model 9.4) gives a very similar fit to the data overall but with a much smaller ratio, c/b, of 6.  This confirms the increase in importance of the number of aircraft (relative to average sound level) on the reported annoyance between the two surveys.

**Model 9.3          Mean Annoyance = -71.6 + 0.86 x Lav + 17.9 x log Nav**

| $R^2$ | Adjusted $R^2$ | N | Intercept | Lav | Log Nav |
|---|---|---|---|---|---|
| 0.656 | 0.643 | 56 | -71.58 | 0.86 | 17.87 |
|  |  |  | (-2.06) | (1.58) | (7.40) |

**Model 9.4          Mean Annoyance = -158.3 + 1.99 x Lav + 12.5 x log Nav**

| $R^2$ | Adjusted $R^2$ | N | Intercept | Lav | Log Nav |
|---|---|---|---|---|---|
| 0.646 | 0.611 | 23 | -158.25 | 1.99 | 12.45 |
|  |  |  | (-4.57) | (5.50) | (2.63) |

9.8.12  Further analysis was done to see how goodness of fit changed with variations in the ratio.  Figure 9.8 shows the $R^2$ values for different values of the c/b ratio from regression analysis of the ANASE and ANIS data.

**mva**consultancy
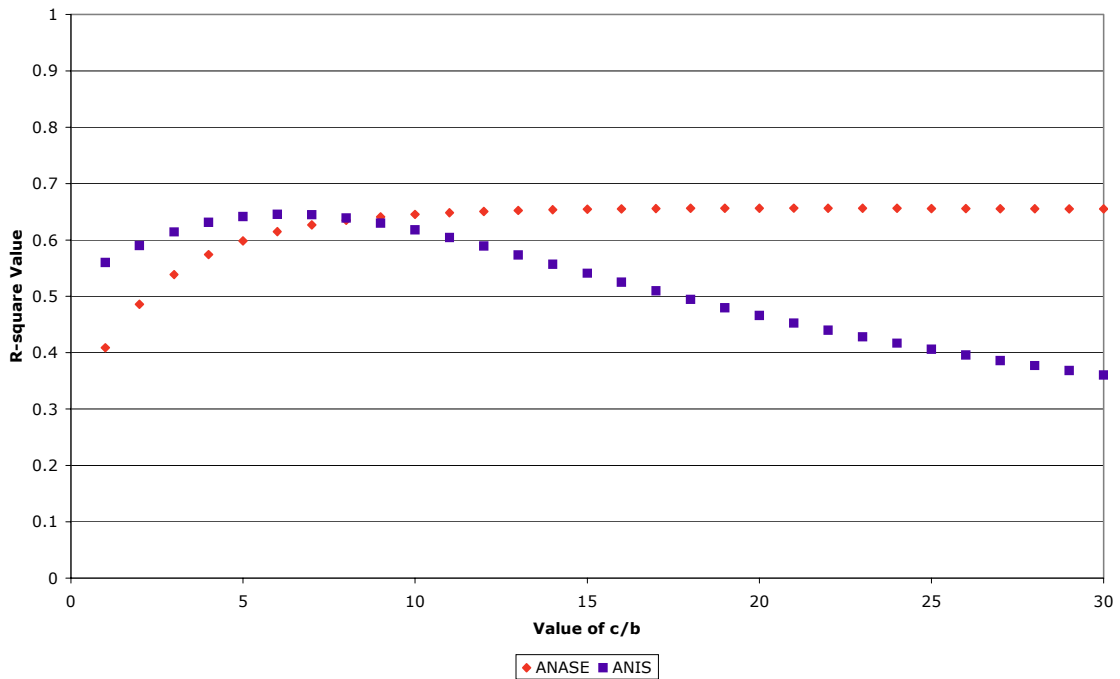
**Figure 9.8   R² values for Regressions of Mean Annoyance against Lav + c/b log Nav (ANASE and ANIS data)**

9.8.13    For ANASE, although the optimum value of c/b is at 21, it is clear that for values above 15, the size of R$^2$ does not change very much.  For ANIS, the optimum occurs at 6, and drops off quickly for higher values.

9.8.14    To provide an alternative metric for comparing ANIS and ANASE (ie. with the same relative weighting for sound level and number), we have used a weight of 15 for log Nav, in between the optimum values for ANIS and ANASE.  A value of 15 also corresponds with the Noise and Number Index.

9.8.15    The Noise and Number Index was defined as:

NNI = L + 15 log N – 80

Where L is the logarithmic average sound level (measured in PNdB[26]), and N is the number of aircraft noise events, excluding noise events below 80 PNdBs.

---

[26] PNdB - Perceived Noise Level in decibels.  PNdB is a complex measure based on one third octave band sound levels.  In the ANIS study, PNdB was not measured directly, but was instead estimated by measuring LAmax and adding 13 dB.

**mva**consultancy

9.8.16   Model 9.5 shows the coefficients produced for the ANASE data, and Model 9.6 shows the coefficients produced for the ANIS data using the Lav + 15 log Nav metric.

**Model 9.5          Mean Annoyance (ANASE) = -87.9 + 1.1 (Lav + 15 log Nav)**

| $R^2$ | Adjusted $R^2$ | N | Intercept | Lav + 15 log Nav |
|---|---|---|---|---|
| 0.655 | 0.648 | 56 | -87.93 _(-6.85)_ | 1.13 _(10.12)_ |

**Model 9.6          Mean Annoyance (ANIS) = -112.1 + 1.3 (Lav + 15 log Nav)**

| $R^2$ | Adjusted $R^2$ | N | Intercept | Lav + 15 log Nav |
|---|---|---|---|---|
| 0.541 | 0.519 | 23 | -112.06 _(-3.48)_ | 1.33 _(4.97)_ |

9.8.17   The coefficients for ANIS and ANASE are similar, and not statistically significant, with the mean annoyance increasing by 1.1 for ANASE and 1.3 for ANIS for a unit increase of Lav + 15 log Nav.

9.8.18   Figure 9.9 plots mean annoyance against Lav + 15 log Nav for ANIS and ANASE and shows that the relationship for ANIS and ANASE is very similar.  The overlap between the 1982 ANIS and the 2005 ANASE data suggests that an NNI type metric could provide a better fit than LAeq to the combined data set.
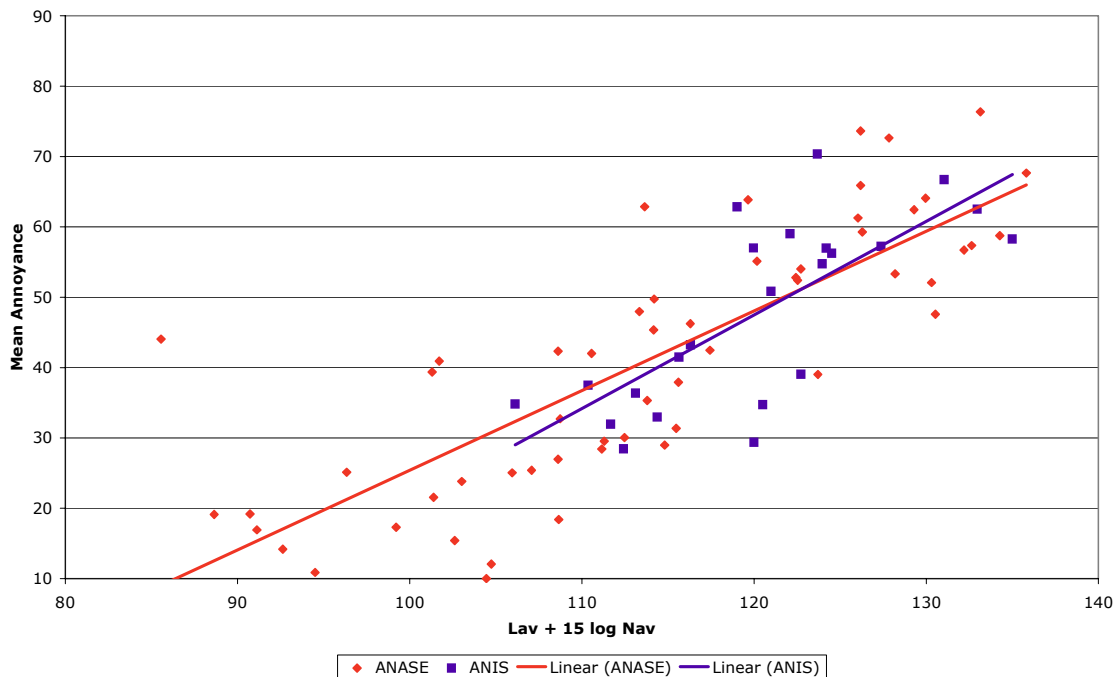


**Figure 9.9          Mean Annoyance against Lav + 15 log Nav**

9.8.19    The above analysis was carried out using a cut-off of 65 LAmax for the ANASE data, and a cut-off value of 80 PNdB (equivalent to 67 LAmax) for the ANIS data.  To investigate the sensitivity of the cut-off value, Figure 9.10 shows the mean annoyance against Lav + 15 log Nav for the two surveys using a cut-off of 55 LAmax for ANASE, and 70 PNdB (equivalent to 57 LAmax) for ANIS.
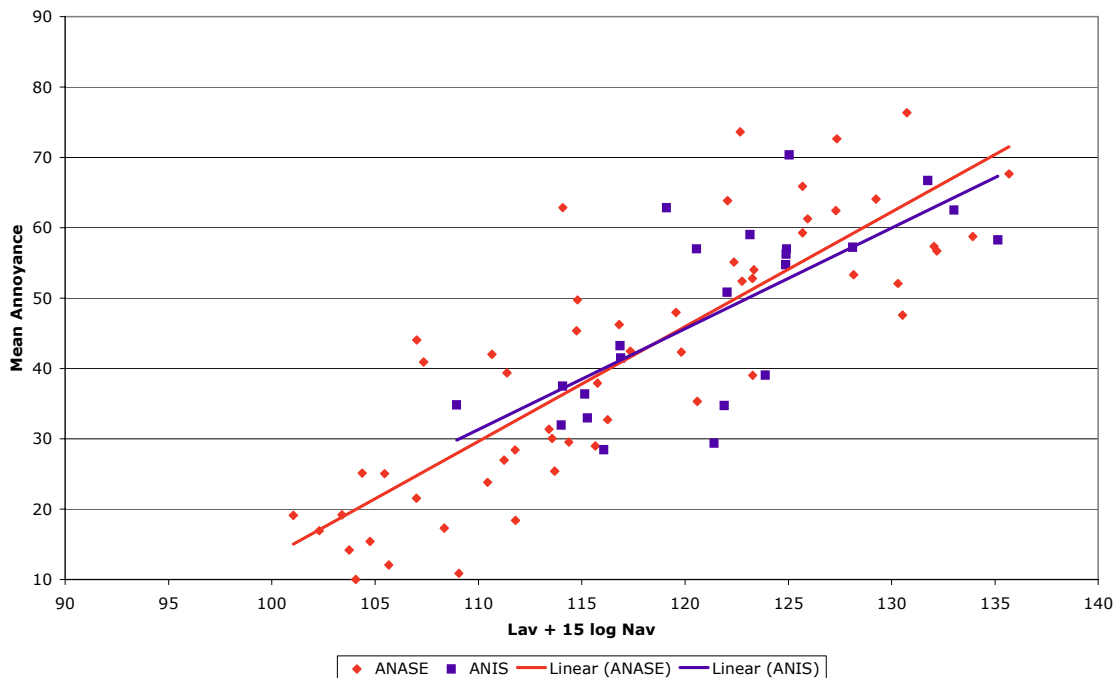


**Figure 9.10       Mean Annoyance against Lav + 15 log Nav using a cut-off of 55 LAmax**

9.8.20    This change in cut-off has resulted in the Lav + 15 log Nav values increasing, as would be expected from the increased number of aircraft in scope for the metric, but the ANIS and ANASE values still overlap, suggesting that the comparison is not sensitive to cut-off values.

9.8.21    Hence we see that, whereas the best fit ratio of the coefficients on log Nav and Lav was 6 in ANIS, it now appears to be over 20, indicating that the relative importance of the number of aircraft has increased over time. By taking a compromise value of 15, it is possible to obtain a close match between the points from the two datasets, as shown in Figures 9.9 and 9.10.


**9.9    A brief discussion of H1 and H2 analysis**

9.9.1    Both hypotheses receive reasonable support from the data.  In portraying them as extremes, neither hypothesis was expected to provide a full explanation for the difference between ANASE and ANIS.  It is not so much a question of choosing between competing hypotheses, as of using the analysis of each to provide some evidence to help account for the differences in attitudes to noise between the ANIS and ANASE surveys.

9.9.2    The main points from the analysis are included in the following summary.

**mva**consultancy

### 9.10    Summary of Main Points

- The LAeq metric is effective at explaining much of the variation in respondent's annoyance in both the ANASE and ANIS studies;

- Using the LAeq metric, the mean annoyance is higher in ANASE (2005) than ANIS (1982);

- Measured on the scale of mean annoyance, the level of annoyance is consistently about 14 points greater in ANASE than it was 23 years ago (where a difference of one category on the ANASE annoyance scale is allocated 20 points);

- For a given mean annoyance score, the 16-hour LAeq value in the ANASE survey is higher by between 3 dB (for a mean annoyance score of 10 or not at all annoyed) and 11 dB (for a mean annoyance score of 90 or extremely annoyed);

- A modelled mean annoyance of 50 (moderately annoyed) is at 63 dB in ANIS and at 55 dB in ANASE, a difference of 8 dB;

- Sensitivity tests confirm the conclusion that for the same amount of aircraft noise, measured in LAeq, people are more annoyed in 2005 than they were in 1982, though the size of the difference is affected by assumptions made;

- A particular issue affecting the size of the difference is whether the introduction of noise playback equipment into the respondent's home generated an exaggerated response to the annoyance rating question.  There is no statistical support for that effect from the ANASE data; substitution of the CAA LAeq estimates at Heathrow produces a significant effect;

- Taking the worst case assumptions about the "equipment effect", the level of difference in annoyance reduces from our central case of 14 points to 7 points. Conversely, alternative assumptions about "aligning" the ANASE and ANIS annoyance ratings increases the level of difference in annoyance to 21 points;

- The sensitivity analyses demonstrate that the relationship between LAeq and reported annoyance is **not** stable over time, and that this is a robust result;

- Based on the cross-sectional evidence derived from ANASE, rising incomes between 1982 and 2005 explain some of the increased annoyance between ANASE and ANIS;

- One possible conjecture is that the differences between the ANIS and ANASE results is the effect of a combination of increasing income (as identified in Chapter 9) over the 23 years between the surveys, and a general "taste effect", whereby people in general have become less tolerant of intrusive noise over time;

- There has been a substantial change in the make-up of aircraft between the two surveys, with many more aircraft in ANASE but which are quieter than they were in ANIS;

- The relationship between annoyance and sound level is stronger for ANIS, but the relationship between annoyance and number is stronger for ANASE;

- The relationship between reported annoyance, sound level and the number of aircraft has not been stable over time.  The weight on aircraft numbers (relative to sound level) has risen from 6 in ANIS to over 20 in ANASE, so the contribution of aircraft numbers to annoyance has increased quite markedly;

- An index based on the Noise and Number Index, as used before the 1982 ANIS study,

**mva**consultancy

where the number of aircraft is given a weight of 15 relative to the sound level of the aircraft, correlates well with reported annoyance in both ANIS and ANASE;

- This correlation is not sensitive to changing assumptions about the LAmax cut-off;

- Because of its instability over time, use of the LAeq measure to predict future levels of annoyance may be misleading.

- The NNI-type index is also not stable over time, with the later ANASE result giving greater weight to aircraft numbers.  However, the ANASE result is relatively insensitive to a weight greater than 20, so an NNI type measure may provide a better tool for predicting annoyance from aircraft noise.

**mva**consultancy

# 10   SP and CVM Analysis

## 10.1   Introduction

10.1.1   This chapter reports the results of our investigation into respondents' trade-off preferences in relation to sensitivity to aircraft noise by time of day, and willingness to pay to reduce or avoid aircraft noise.  The latter was one of the three primary objectives of the study, as identified in the introductory chapter, and the prescribed method was SP.

10.1.2   We have included two types of question, SP and CVM.  In the following sections of this chapter we report the results of our analyses of the data.

## 10.2   ANASE SP Research – the task presented to respondents

10.2.1   Building on the work in Phase 1, we divided the day up into six four-hour time periods (see Table 4.1).  To investigate the variation in annoyance throughout the day, the SP experiment was presented for three selected time periods for each respondent.  For a given time period, respondents were presented with four choice sets each comprising three options and, for each choice set, were asked to identify which option they preferred most and which one they preferred least, thereby ranking the options.

10.2.2   We defined each option in terms of two or three aircraft types with varying numbers of each aircraft flying overhead in the given time period, based on the most common aircraft to fly over each site.  In addition, a money variable was included, which, again based on the Phase 1 research, was defined as a household grant, presented as an annual figure on each SP showcard[27].

10.2.3   Prior to the SP exercises respondents were exposed to samples of locally recorded aircraft noise for the aircraft types that were included in the SP choice-sets: this ensured that respondents approached the trade-off exercises in a consistently informed manner. In addition, the use of line drawings to represent the aircraft types provided a cognitive link for the respondent between the recordings heard earlier in the interview and the aircraft types presented on the cards.

10.2.4   Respondents were also presented with information about the current level of aircraft activity in their area, which they could use as a baseline reference during the SP exercises. When considering the different options in each choice set, respondents were instructed to assume that other aircraft types (not included in the SP) remained constant.  Where such aircraft were of significant number, the showcards depicting each SP option included a constant number of aircraft of this type in the total aircraft figure for the period.

10.2.5   Fieldwork practices were designed so that different respondents were presented with exercises for a mix of different time periods in such a way that when combining results across each sub-sample, the whole 24-hour day was covered evenly within each site.  The time periods presented were also systematically varied across the sample to ensure each was considered by around half the sample.

---

[27] In the analysis models, the money variable has been input as the equivalent monthly amount and in negative form, hence, the reported money coefficients represent a response to a **monthly** negative cost

**mva**consultancy

10.2.6    The context to the SP questions was as follows:

*"Please think about what it would be like for you and the other member(s) of your household if there were different numbers of certain aircraft flying over YOUR HOME. I also want you to imagine that households near the airport qualify for an annual grant. <u>This household grant can be spent on anything your household wants.</u> So, you could spend it on improvements in insulation or double-glazing, or put it towards something like a new car or a holiday.*
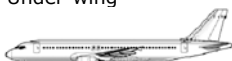
*The Next questions are about the time of day between …… and …….*

*On a typical day there was this number of each type of aircraft flying over your area [on their way to land at/taking off from] the airport between …… and …….*

*I now want you to think about three different situations.* **[PRESENT COLOURCARD 2A]** *Have a look at the situations described in each of the three boxes A, B and C, and tell me which you think would be the best situation for you and your household.*

*Please assume that there are no other differences in the numbers of the other types of aircraft, and no changes in the numbers of aircraft outside the hours …… and ……. We want you to only think about the effect of the differences shown on this card. Everything else remains the same.*

*Which do you think would be the best situation for you and your household?.*



| BOX A<br>…..the following planes fly overhead: | BOX B<br>…..the following planes fly overhead: | BOX C<br>…..the following planes fly overhead: |
|---|---|---|
| 0 Jumbo | 1 Jumbo | 1 Jumbo |
| 7 Under-wing | 3 Under-wing | 7 Under-wing |
| 0 Tail-jet | 0 Tail-jet | 6 Tail-jet |
| 7 Total aircraft | 4 Total aircraft | 14 Total aircraft |
| No grant | No grant | £200 grant every year |

*And which would be worst for you and your household?"*

10.2.7    Most features of the SP designs at each site were identical. Each site SP exercise required the development of six designs, one for each time period. Each of these six designs was based upon a common "skeleton" design, that is, a structure in which options are expressed purely in terms of levels rather than actual values.

10.2.8    Both skeleton designs comprised 8 choice sets, each with 3 options, with 3 levels for each variable.

**mva**consultancy

10.2.9    For each site and time period, customised values were assigned to the skeleton design to ensure that the presented SP options were realistic, with regard to the existing aircraft noise environment.    All designs used in the surveys were tested using software that simulates respondent choices.    This enabled the designs to be tested for their ability to recover coefficients for a range of different assumed valuations of the disutility caused by aircraft noise.    Correlation between variables can also be identified through this approach.    Full technical details of the designs are provided in Appendix A10.1.

### 10.3    ANASE SP Research – preliminary analysis

10.3.1    Analysis of responses to a number of choice sets for a particular time period across a sample of respondents allowed the estimation of coefficients β of a "utility function" model (as is conventional in standard discrete choice analysis) of the basic form:

$$U = \beta_1.N_1 + \beta_2.N_2 + \ldots + \beta_N.N_N + \beta_M.M + \varepsilon \qquad (10.1)$$

where $N_a$ represents the number of aircraft of type a, and M is the grant (expressed, as noted earlier, as a negative monthly cost)

10.3.2    The β coefficients represent the utility associated with an increase of one unit in the variable with which they are associated.    Hence $\beta_1$ represents the utility of one additional aircraft of type 1 flying over the respondent's area, and is expected to be negative.    The reported t-statistics that corresponds to each β takes into account the fact that multiple responses are recorded for each respondent.    The term $\varepsilon$ is the error component in the "Random Utility" model of Discrete Choice.

10.3.3    Presenting respondents with separate trade-off exercises related to different time periods allowed a set of coefficients to be estimated for each time period.    However, there is no reason to expect the coefficient on money $\beta_M$ to vary by time period, and this allowed money to be used as a numeraire to compare the results for different periods to assess the variation in noise annoyance across the day.    Hence for *A* aircraft types and *T* time periods, a utility function involving (*A* x *T*)+1 β coefficients can be formulated, of the form:

$$U = \sum_{t=1}^{T} \sum_{a=1}^{A} \beta_{at}.N_{at} + \beta_m.M + \varepsilon_0 \qquad (10.2)$$

where

|        |   |                                                        |
|--------|---|--------------------------------------------------------|
| *A*    | = | the number of aircraft types presented                 |
| *T*    | = | the number of time periods                             |
| $N_{at}$ | = | the number of aircraft of type *a* presented in time period t |
| M      | = | the amount of monetary compensation presented          |
| $\beta_{at}, \beta_m$ | = | coefficients to be estimated               |

10.3.4    According to standard principles of Discrete Choice analysis, models can be judged both on the basis of individual coefficients (t- ratios) and overall goodness of fit ("log-likelihood") with respect to the number of coefficients estimated (more discussion is provided in

**mva**consultancy

Appendix A10.2).

10.3.5    The SP data analysis reported throughout this report is unweighted.  The analysis was also carried out with the data weighted to represent the population in scope for the full survey, with only marginal differences in results.

10.3.6    The SP analysis was initially undertaken at the level of the 36 individual sites, before moving systematically on to models that combined data across sites.  Details of each stage of SP modelling, and comparisons of different goodness-of-fits of the different models, are provided in Appendix A10.2.

10.3.7    The 'Basic' models produced a relative coefficient ($\beta_{at}^{s}$) for each aircraft type $a$ in each time period $t$, and a relative coefficient for money ($\beta_{m}^{s}$), for each of the 36 sites $s$.  Each Basic SP Model, therefore, derived 13 or 19 relative coefficients, depending upon whether the site-specific SP exercise comprised 2 or 3 aircrafts with varying levels.  Apart from one site (H1H) which had eight aircraft coefficients with low t-ratios, of the remaining 629 coefficients only 7 had unexpected signs, and 82% of the 622 coefficients with the expected sign were statistically significant.  Given that only 60 respondents contribute to the estimation of each (site-specific) Basic SP model, the fact that the vast majority of estimated coefficients were statistically significant was encouraging for the robustness of the design and respondents' willingness to trade between aircraft types and money.

10.3.8    Further tests of the basic models were carried out to see whether the implied variation in coefficients by a) time period and b) aircraft type could be justified[28]. Generally, this restriction resulted in considerable loss of explanatory power at site level.  This finding indicated that both the time of day and type of aircraft contributed significantly to respondents' relative valuations.

10.3.9    In order to test whether the effects of different aircraft types and time of day were **independent**, site-level "multiplicative" models were developed to separate out these two effects.  The site-level multiplicative (non-linear-in-parameters) model form is:

$$U = \sum_{t=1}^{T} \ \sum_{a=1}^{A} \beta_{a}^{\ s}.\beta_{t}^{\ s}.N_{ats} + \beta_{m}^{\ s}.M_{s} + \varepsilon_{0} \qquad (10.3)$$

10.3.10   With this formulation, the models derived relative values $\beta_{a}$ for each aircraft type for a given "base" time period (arbitrarily defined to be Time Period 1) plus modifying factors $\beta_{t}$ for each of the other 5 time periods relative to the base period, as well as a money coefficient. Therefore, the site-level multiplicative models derived 8 or 9 relative coefficients (2 or 3 aircraft types + 5 relative time periods + money) plus the base time period (set at unity).

10.3.11   Though there was some loss of explanatory power in moving to the site-specific multiplicative models (as would be expected from the reduced number of explanatory variables), the loss in goodness of fit was relatively small, indicating that the "interaction" between aircraft type and time period was not very strong.

10.3.12   At this point we explored whether a better explanation of the SP data was provided by assuming that respondents' valuations were dependent upon proportional differences in aircraft, rather than absolute differences.  This would be the case **if** the 'value' of one

---

[28] This can be done by restricting the coefficients $\beta_{at}$ in Eq (10.2) a) to have the same value $\beta_{a*}$ for all time periods t, and b) to have the same value $\beta_{*t}$ for all aircraft types a

**mva**consultancy

more/less jumbo was more strongly correlated with the percentage change in jumbos between two or more SP options than the absolute difference. However, this 'logarithmic' model provided a worse explanation of the data than the SP model that assumed values according to absolute differences in number.

10.3.13    While up to this point all the analysis had been carried out separately for each site, the ultimate aim of the SP analysis is to achieve an integrated model across sites. While we would still expect the aircraft coefficients to vary by site, since the sound levels will vary, there is less reason to expect site-specific variation in either the time of day effects or in the response to money (though, of course, these could be further affected by variations in the sample of respondents between sites). The form of the multiplicative model enabled us to then develop a single 'pooled' model that combined the SP data across all 36 sites.

10.3.14    This (first) pooled multiplicative model derived relative values for each aircraft type and a money coefficient, for each site, as well as overall values for each of the 5 time periods relative to the (base) time period. The model form for this pooled multiplicative model is:

$$U = \sum_{t=1}^{T} \sum_{a=1}^{A} \beta_a^s . \beta_t . N_{ats} + \beta_m^s . M_s + \varepsilon_0 \qquad (10.4)$$

10.3.15    Note that compared with model (10.3) the time coefficient is no longer-site specific. As might be expected, given the large reduction in model coefficients, this did result in a significant deterioration in goodness of fit, suggesting that there **are** some site-specific variations in response by time of day. Unfortunately, attempts to define these more precisely were not successful.

10.3.16    The final step jn this preliminary analysis was a pooled multiplicative model that 'forced' the site-specific money coefficients to a single money coefficient. The further loss of goodness of fit when deriving a single money coefficient is most probably due to differences in (sound level  or other) circumstance across sites that are not specifically accounted for in the model, rather than actual differences in value of money. This model, which we have termed the pooled 'National Model', derived 102 relative aircraft coefficients (2 or 3 aircraft types x 36 sites) together with 5 relative time period coefficients and money. Despite the reduction in overall explanation, only two of the estimated coefficients are statistically insignificant.

10.3.17    In this model, which is discussed in more detail in the following section, the response to money and the time of day effects is the same across all sites, but the individual aircraft coefficients vary by site.

10.3.18    In this section, we have described the initial SP Models which were developed, enabling us to progress from site-specific SP Models to a pooled SP Model that kept the relative value of time periods and (separately) money constant across sites.   The modelling process is summarised in Figure 10.1.
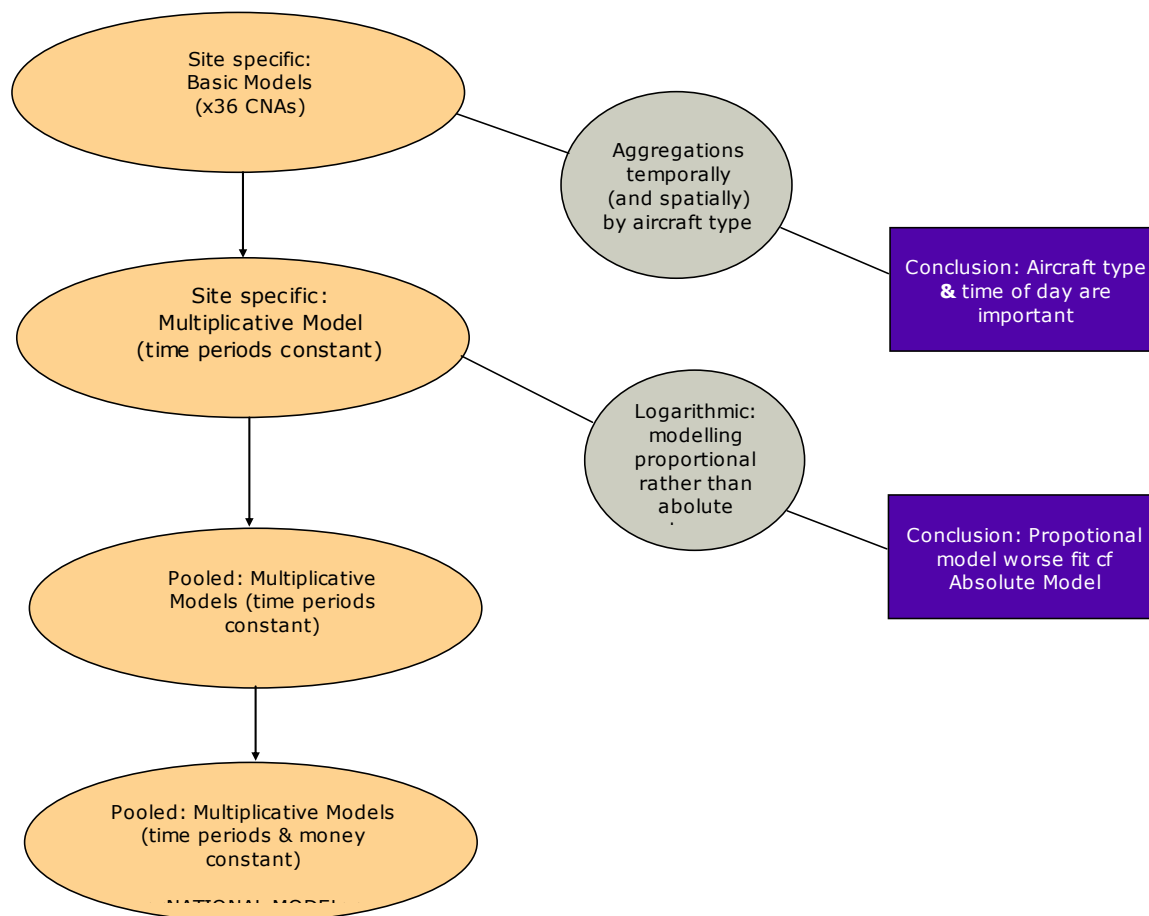


**Figure 10.1          Flow-Diagram of SP Modelling Process change**

10.3.19    The table below summarises the log-likelihoods for the various "pooled" models discussed here, in order of decreasing likelihood. This value is an indication of the goodness of fit of the model relative to a 'null' hypothesis that assumes that none of the modelled variables provides any explanation of the data.   By noting the difference between the modelled loglikelihood and the null loglikelihood we can measure improvements in goodness of fit against the null hypothesis.   The null and model loglikelihoods are negative so as the fit of a model improves, the model loglikelihood tends to zero.   If the difference is statistically significant then we can assume that the variables in the model are helping to explain respondents' stated preferences.   By comparing the model loglikelihoods between different SP models, and allowing for the number of explanatory variables, we can identify improvements in goodness of fit.

**mva**consultancy

**Table 10.1        Summary of log-likelihoods for alternative model formulations**

| Model formula | Model description | no. of parameters estimated | Log-likelihood |
|---|---|---|---|
| $U = \sum_{t=1}^{T} \sum_{a=1}^{A} \beta_{at}.N_{at} + \beta_{m}.M + \varepsilon_{0}$ (10.2) | site-level additive | 648 | -37975.35 |
| $U = \sum_{t=1}^{T} \sum_{a=1}^{A} \beta_{a}^{s}.\beta_{t}^{s}.N_{ats} + \beta_{m}^{s}.M_{s} + \varepsilon_{0}$ (10.3) | site-level multiplicative | 318 | -38423.96 |
| $U = \sum_{t=1}^{T} \sum_{a=1}^{A} \beta_{a}^{s}.\beta_{t}.N_{ats} + \beta_{m}^{s}.M_{s} + \varepsilon_{0}$ (10.4) | pooled with site-specific money | 143 | -38542.48 |
| $U = \sum_{t=1}^{T} \sum_{a=1}^{A} \beta_{a}^{s}.\beta_{t}.N_{ats} + \beta_{m}.M_{s} + \varepsilon_{0}$ | Pooled National | 108 | -38992.68 |
| $U_{s} = 0$ | Null | 0 | -45091.42 |

10.3.20    Even though these reductions are not justified by conventional statistical criteria, it can be seen, in relation to the "null" model, that they explain a significant amount of the variation in the data.

## 10.4    ANASE SP Research – Pooled (National) SP Model

10.4.1    The pooled National Multiplicative SP Model covering all 36 sites, with a single set of time period coefficients, and a single money coefficient is presented in Table 10.2.  This model provides the basis for deriving estimates of willingness to pay for a unit reduction of aircraft by aircraft type at each site.

10.4.2    The table reports the relative coefficient for each aircraft type at each site, and for each time period, as well as money; along with the corresponding t-statistic (in parenthesis) that indicates each estimate's statistical validity.  Blank cells indicate the absence of a certain aircraft type at the site concerned.  [Note 60 respondents contributed to the model coefficients for every site, except one – O6D - where only 21 interviews were collected].

**mva**consultancy

### Table 10.2 Pooled Multiplicative SP Model (National Model) (t-statistics shown in brackets)

| Site | Jumbo | | Underwing[29] | | Tailjet[30] | | Turboprop | |
|---|---|---|---|---|---|---|---|---|
| | Coeff't | £ Value | Coeff't | £ Value | Coeff't | £ Value | Coeff't | £ Value |
| H1H | -0.0425 (-0.9) | 1.18 | -0.0084 (-0.6) | 0.23 | -0.0425 (-5) | 1.18 | | |
| H1J | -0.4881 (-8.7) | 13.54 | -0.2266 (-7.8) | 6.29 | -0.1289 (-8.5) | 3.58 | | |
| H1K | -0.3687 (-6.6) | 10.23 | -0.1221 (-5.3) | 3.39 | -0.0921 (-6.4) | 2.55 | | |
| H1L | -0.4429 (-9.4) | 12.29 | -0.0694 (-3.5) | 1.93 | -0.0542 (-4.2) | 1.5 | | |
| H1M | -0.2608 (-5.2) | 7.24 | -0.0361 (-2.3) | 1 | -0.0706 (-5.3) | 1.96 | | |
| H1P | -0.607 (-9) | 16.84 | -0.1392 (-5.3) | 3.86 | -0.1193 (-7.4) | 3.31 | | |
| H3A | -0.3582 (-4.2) | 9.94 | -0.1013 (-3.9) | 2.81 | -0.1457 (-7.8) | 4.04 | | |
| H3B | -0.5439 (-7.8) | 15.09 | -0.3379 (-7.9) | 9.37 | -0.2059 (-8.8) | 5.71 | | |
| H3C | -0.4902 (-6) | 13.6 | -0.1028 (-3.3) | 2.85 | -0.0269 (-2) | 0.75 | | |
| H3D | -0.6128 (-9) | 17 | -0.1116 (-4.4) | 3.1 | -0.0864 (-6.1) | 2.4 | | |
| H3E | -0.1314 (-2.6) | 3.64 | -0.0436 (-2.5) | 1.21 | -0.0604 (-5.1) | 1.68 | | |
| H5A | -0.1498 (-5.9) | 4.16 | -0.132 (-5.6) | 3.66 | | | | |
| H5B | -0.2503 (-4.4) | 6.94 | -0.034 (-1.7) | 0.94 | -0.0943 (-6.3) | 2.62 | | |
| H5C | -0.5151 (-8.1) | 14.29 | -0.2075 (-6.9) | 5.76 | -0.1343 (-8.6) | 3.73 | | |
| H5D | -0.5075 (-7.9) | 14.08 | -0.1339 (-5.6) | 3.72 | -0.1737 (-11.1) | 4.82 | | |
| H5E | -0.3599 (-7.6) | 9.99 | -0.3918 (-8.7) | 10.87 | | | | |
| H5F | -0.4488 (-8.2) | 12.45 | -0.102 (-4.6) | 2.83 | -0.0701 (-5.1) | 1.95 | | |
| O2A | | | -0.1253 (-7.9) | 3.48 | -0.1342 (-7.8) | 3.72 | | |
| O2C | | | -0.4982 (-8.1) | 13.82 | | | -0.3464 (-7.5) | 9.61 |
| O2D | -0.7061 (-7.8) | 19.59 | -0.135 (-7.3) | 3.75 | -0.1543 (-10.5) | 4.28 | | |
| O2E | -0.6174 | 17.13 | -0.1004 | 2.79 | | | -0.112 | 3.11 |

---

[29] The underwing statistics are given for all sites, apart from sites O4G, O6D, O6E and O6F, where large underwing statistics are given.

[30] The tailjet statistics are given for all sites apart from sites O4G, O6D, O6E and O6F, where small underwing statistics are given

**mva**consultancy

| Site | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (-7.4) | | (-4.3) | | | | (-7.6) | |
| O2F | -0.525 (-7.4) | 14.57 | -0.0651 (-3.7) | 1.81 | | | -0.0801 (-7.1) | 2.22 |
| O4A | | | -0.1466 (-4.4) | 4.07 | | | -0.1127 (-3.3) | 3.13 |
| O4B | -0.5144 (-7) | 14.27 | -0.1263 (-6) | 3.5 | | | -0.0973 (-7.7) | 2.7 |
| O4C | -0.6336 (-8.4) | 17.58 | -0.1309 (-7.2) | 3.63 | | | -0.1425 (-10.6) | 3.95 |
| O4D | | | -0.2943 (-6.6) | 8.17 | -0.2041 (-5.1) | 5.66 | | |
| O4E | -0.2267 (-7.2) | 6.29 | -0.0832 (-3.3) | 2.31 | | | -0.0864 (-6.5) | 2.4 |
| O4F | -0.3896 (-6.6) | 10.81 | -0.1093 (-6.1) | 3.03 | | | -0.1262 (-9.8) | 3.5 |
| O4G | | | -0.8071 (-7.9) | 22.39 | -0.2905 (-8.5) | 8.06 | -0.2162 (-8.7) | 6 |
| O5A | | | -0.3787 (-6.1) | 10.51 | -0.3122 (-6.3) | 8.66 | -0.194 (-6.7) | 5.38 |
| O6A | | | -0.2688 (-6.5) | 7.46 | -0.0984 (-3.5) | 2.73 | -0.1431 (-6.3) | 3.97 |
| O6B | | | -0.2904 (-7.4) | 8.06 | -0.1045 (-3.5) | 2.9 | -0.1138 (-5.9) | 3.16 |
| O6C | | | -0.289 (-6.5) | 8.02 | -0.1124 (-4.2) | 3.12 | -0.0695 (-5.2) | 1.93 |
| O6D | | | -0.6735 (-4.2) | 18.69 | -0.2361 (-3.5) | 6.55 | -0.1614 (-4.2) | 4.48 |
| O6E | | | -0.5283 (-5.5) | 14.66 | -0.138 (-4.6) | 3.83 | -0.1855 (-7) | 5.15 |
| O6F | | | -0.607 (-7) | 16.84 | -0.2691 (-5.9) | 7.47 | -0.257 (-8.1) | 7.13 |

| | | |
|---|---|---|
| 2300-0300 | 1.0000 | FIXED |
| 0300-0700 | 0.7321 | 18.5 |
| 0700-1100 | 0.5454 | 19.9 |
| 1100-1500 | 0.5425 | 19.7 |
| 1500-1900 | 0.6027 | 20.5 |
| 1900-2300 | 0.6332 | 23.0 |
| MONEY | -0.0360 | -26.3 |

Null  Loglikelihood = -45091

Model Loglikelihood = -38993

10.4.3    The ratio of an aircraft coefficient and the money coefficient gives the monetary value that respondents assign to one extra aircraft (of type *a*) flying over their home in the 'base' time period.    For example, the results reported in Table 10.3 for Site O4F suggests that respondents would be willing to pay around £11 a month to have 1 less jumbo flying over their home between 11pm and 3am every day; £3.50 a month for one less turboprop; and £3 a month to have 1 less underwing.   The base time period is during the middle of the night, and values per aircraft can be derived for other time periods by applying the

**mva**consultancy

appropriate time period factor. So, the valuations for each aircraft type at O4F during the day-time (e.g time period 4, 11am - 3pm) are around £6, £2 and £1.50 for jumbo, turboprop and tailjet respectively. A fuller discussion of the implied willingness to pay values is deferred until later in the chapter.

10.4.4    For all but three of the modelled coefficients, the corresponding t-statistics have an absolute value greater than 1.96, so are statistically significant at the 95% level. Two of the three insignificant estimates are at the same site (H1H).

**Income Effects**

10.4.5    To test the relationship between household income and the money coefficient, a model was developed which segmented the money coefficient by "protest voters"[31], and for the remaining respondents, by six income bands (including a non-response category). The full model is shown in Table D7 of the Appendix A10.2, while the money coefficients are shown in Table 10.3.

**Table 10.3    Multiplicative model: All Sites, Money Segmented by Income and Protest Voters – Money Coefficients Only**

| Income/Protest Voter | Coefficient | t-statistic |
|---|---|---|
| LOWEST INCOME (<£10k) | -0.050 | -10.8 |
| LOW INCOME (£10-20k) | -0.041 | -12.6 |
| MID INCOME (£20-40k) | -0.030 | -10.8 |
| HIGH INCOME (£40-50k) | -0.030 | -5.1 |
| HIGHEST INCOME (£50k+) | -0.023 | -6.2 |
| MISSING | -0.029 | -7.3 |
| PROTEST | -0.044 | -11.7 |

10.4.6    The model loglikelihood is -38899, and represents an improvement of 93 loglikelihood points for an additional 6 coefficients, which is a statistically significant improvement, though it only covers a small degree of the apparent variation in the money coefficient between all 36 sites. This model shows the money coefficients by income band decreasing in magnitude in a generally plausible way: the coefficient for income non-response is close to the middle income values. Note that the "protest" respondents behave more like the low income groups.

10.4.7    It should also be noted that respondents were asked to consider the options within each SP choice-set from the perspective of themselves and the other members of their household (see paragraph 10.2.6 for details). However, further segmenting of the money coefficient by

---

[31] the definition of this category is given in section 10.8.6

mvaconsultancy

number of adults in the household suggested that there was little overall difference in money valuation by household size.

### 10.5    Implications for LAeq

10.5.1    The "national model" presented in Section 10.4 provides a well-founded and statistically sound basis for assessing the response to different kinds of aircraft at different sites, as well as a basis for the variation by time of day. However, this model does not in itself represent a truly national model, since it remains tied to the individual sites in the sample.  To obtain a generalised national model, we need to take explicit account of the sound levels associated with the aircrafts at each site.

10.5.2    It was recognised at the outset that this would be a major challenge.  Ideally, what we would like to do is to relate the individual aircraft type coefficients to the sound levels, in such a way that we could develop a model which would be based on the total sound level represented by each SP scenario.  If this could be achieved, it would allow us to derive SP-based valuations of changes in overall sound levels from aviation (as opposed to changes in the number of particular types of aircraft at particular sites).   The same approach would provide some assessment of the suitability of the LAeq index as a proxy for community annoyance.  This was the other objective of the SP work.

10.5.3    We began by considering how the SP model could be developed if LAeq were indeed an appropriate index.  As we discuss below, this leads to an implied interpretation of the SP aircraft coefficients in the "national model".

10.5.4    As discussed in Section 2.2, the LAeq value represents the total amount of 'acoustical energy' received at a point, and the same value can arise from a range of combinations of numbers of movements and average sound level per movement.  If, at a particular site $s$, the sound energy for a movement of aircraft type $a$ is $e_{as}$, then the implied total sound energy $E_{sk}$ of a SP option k comprising $N_{ask}$ movements for each of the aircraft types a is given by:

$$E_{sk} = \sum_a e_{as} N_{ask}$$

10.5.5    If LAeq is a good proxy for annoyance, therefore, ordering the SP options according to their $E_{sk}$ levels should correlate well with the "utility" ordering of the SP options, which will be determined by the weights ($\beta_a{}^s$) inferred for the different aircraft types, according to the standard formula:

$$U_{sk} = \sum_a \beta_a{}^s N_{ask} \ .$$

10.5.6    To establish the correlation between $U_{sk}$ and $E_{sk}$ that is needed to confirm a relationship between annoyance and LAeq, it is therefore implied that the ratios between the weights in the utility expression should equal the corresponding ratios of sound energy.  A corollary of this is that the $\beta_a{}^s/e_{as}$ ratios will be the same for all aircraft types, at all sites.

10.5.7    A straightforward way to test the equality of ratios of weights and sound energy is to fit a model in which the large number of variables representing the number of movements of different aircraft types $a$ at different sites $s$ are replaced by a single variable proportional to the total amount of energy per square metre $E_{st}$ produced by those movements in time

**mva**consultancy

period $t$.

10.5.8    For an SP scenario involving $A$ aircraft types, this sound energy is calculated from the number of movements ($N$) in time period $t$ and the Sound Exposure Level ($SEL$) of each aircraft type $a$ at each site $s$, as follows:

$$E_{st} = \gamma \sum_a N_{ast} . 10^{SEL_{as}/10}$$

where $\gamma$ is an arbitrary scaling factor related to the definition of the level (in decibels): for convenience, we set this to unity.

10.5.9    This model can thus be formulated as follows:

$$U_s = \beta_E . \sum_{t=1}^{T} \beta_t . E_{st} + \beta_m . M + \varepsilon_0 \qquad (10.5)$$

where $\beta_E$ is the single "sound energy" coefficient

10.5.10    If the goodness-of-fit of this model approached that of the "national model" in which separate weights (the $\beta_a{}^s$ weights) were estimated, it would indicate (within the bounds of statistical uncertainty) a one-to-one relationship between increasing sound energy and declining utility, and would be strong corroboration that LAeq was a good proxy for annoyance.

10.5.11    However, the model loglikelihood of this model is low, -43590.28 [Table D10 in A10.2 Annex B, also Table 2 in A10.3], compared to -38992.68 for the equivalent model [Table D6 in A10.2 Annex B] with separate aircraft coefficients, and -45091.42 for the null loglikelihood. Even allowing for the significant reduction in the number of coefficients, the loss of explanation implies a very weak relationship between disutility and sound energy, suggesting that LAeq may not be a good proxy for the response to different combinations of noise events.

10.5.12    One possibility may be that the utility formulation we are considering assumes that it is the **difference** between sound energies in the various SP options which is important in affecting choice, whereas it is generally considered that in sound terms it is the **ratio** (logarithmic difference) which is relevant.  For this reason we re-formulated the model as:

$$U_s = \beta_E . \sum_{t=1}^{T} \beta_t . [10.\log_{10} E_{st}] + \beta_m . M + \varepsilon_0 \qquad (10.6)$$

10.5.13    This in fact produced a significantly better log-likelihood (-42726.21), but it was still a poor result relative to the site-specific model [Table D6 in A10.2 Annex B].

10.5.14    In order to make progress, we therefore tried to develop **direct** relationships between the site-specific aircraft coefficients $\beta_a{}^s$ estimated in the "national model" (Table 10.3), and the recorded sound level for each aircraft (SEL).  The models developed as part of that investigation, including the all sites model using sound energy, a single set of time period coefficients and money coefficients, and the SEL data used in these models are provided in Appendix A10.3.

10.5.15    We began by investigating simple linear models of the form:

$$\beta_a{}^s = \alpha + \lambda . SEL_{as}$$

10.5.16    Although this produced a significantly negative coefficient on SEL, further investigation suggested that much of the explanation was coming from the within-site variation, rather than the between-site variation.

10.5.17    Since the $\beta$ coefficients in the SP analysis are multiplied by the number of aircraft, according to the LAeq formulation it is implied that a relationship of the form:

$$-\beta_a{}^s = \alpha'.10^{SEL_{as}/10}$$

would be appropriate.  This can be transformed to

$$\log_{10}(-\beta_a{}^s) = \alpha + \lambda.SEL_{as}$$

where $\alpha = \log_{10}(\alpha')$ and the implied value of $\lambda$ is 0.1.

10.5.18    As discussed in Appendix A10.3, when this equation was estimated allowing the constant term $\alpha$ to vary by site, we obtained a result in which the average value of the constant was −5.282 and the coefficient $\lambda$ had the value .0572, with a t-statistic of 8.43.  This is clearly significantly different from the value of 0.1 which would be required for the LAeq formulation.

10.5.19    Further work (discussed in Appendix A10.3) was carried out in order to reduce the site-specific effects, but this was only partly successful.  However, it suggested, if anything, that a **lower** value of $\lambda$ could be supported (at 0.0416).

10.5.20    Taking an average value of 0.05 from these investigations, we can tentatively suggest that the estimated aircraft coefficient bears the following average relationship to the SEL of the aircraft:

$$-\beta_{as} = \alpha'.10^{SEL/20}$$

10.5.21    Compared with the LAeq formula, this represents a significantly lower rate of increase in annoyance (or, here, decrease in "utility") for increasing sound levels.  The LAeq formula implies that, for the same number of aircraft, annoyance increases by 26% for each additional dB, so that the annoyance increases by 10 times for an additional 10 dB.  The formula just given, by contrast, implies that annoyance rises by about 14% for each additional dB, so that for an additional 10 dB the annoyance is 3.7 times as much.

10.5.22    To illustrate it in another way, LAeq implies that 100 events with SEL 80 is equivalent to 10 events with SEL 90, over the same defined period.  The results here imply that the equivalence is nearer to 32 events with SEL 90, in other words that the role of number should be upgraded relative to SEL.  Approximately, the implied relationship can be considered to be:

LAeqx = SEL + 20 $\log_{10}$ N − 10 $\log_{10}$ T

10.5.23    On this basis, therefore, for predicting changes in community disutility in response to changes in aircraft sound level, a ('k') weighting of 20 on the number variable would seem to be better than the weighting of 10 that currently exists in the LAeq formula.  This supports the finding reported in the previous chapter (§10.7), which found that for the ANASE

**mva**consultancy

annoyance data, a weighting of 21 on $\log_{10}(N)$ may be a better proxy than LAeq for predicting changes in community annoyance.

10.5.24    Having obtained what appears to be an improved relationship between the estimated coefficients $\beta_a{}^s$ and the sound levels of individual aircraft, we can now return to the SP data. In place of the earlier model (10.5) where we attempted to estimate a single coefficient on the term $E_{st}$, we now re-calculate the "energy term" as:

$$E'_{st} = \sum_a N_{ast}.10^{SEL_{as}/20}$$

We can then, as before, formulate the utility model as

$$U_s = \beta_E.\sum_{t=1}^{T} \beta_t.E'_{st} + \beta_m.M + \varepsilon_0 \qquad (10.7)$$

where $\beta_E$ is the single "sound energy" coefficient

10.5.25    Compared with the corresponding model using $E_{st}$ rather than $E'_{st}$, there is a significant improvement in the model loglikelihood, -41766.4, and this is also much better than the re-formulated model (10.6) using $10 \log_{10} E_{st}$. Interestingly, substituting $E'_{st}$ into this version of the model, ie using:

$$U_s = \beta_E.\sum_{t=1}^{T} \beta_t.[10.\log_{10} E'_{st}] + \beta_m.M + \varepsilon_0 \qquad (10.8)$$

led to a worse log-likelihood (-42536.9), though still better than the corresponding result with $E_{st}$.

10.5.26    On this basis, our preferred model is set out in Table 10.4 below.

**Table 10.4 Pooled SP Model (10.7)**

| Variable | estimate | t-statistic |
|---|---|---|
| $E'_{st}$/1000 | -0.0132 | -20.7 |
| Time period 1 (2300-0300) | 1.000 | FIXED |
| Time period 2 (0300-0700) | 0.591 | 13.5 |
| Time period 3 (0700-1100) | 0.438 | 14.5 |
| Time period 4 (1100-1500) | 0.441 | 14.2 |
| Time period 5 (1500-1900) | 0.480 | 15.8 |
| Time period 6 (1900-2300) | 0.514 | 17.0 |
| Money | -0.0280 | -21.6 |
| Null  Loglikelihood = -45091 | | |
| Model Loglikelihood = -41766.4 | | |

10.5.27    Note that this model essentially provides a valuation in terms of the ratio of the coefficients on $E'_{st}$ and "money", suggesting that a unit change in ($E'_{st}$/1000) (NB in period 1) is valued at £0.47.  Since the temporal units are implicitly the same, it can be assumed that this is the value per day over which the change occurs.

10.5.28    Unfortunately, it is extremely difficult to interpret this, given the units of $E'_{st}$!.  Purely as an example, we can consider the base level for the SP scenarios presented in Site H5F for time period 4:

**mva**consultancy

■      Jumbos        2      SEL value      83.2

■      Underwings    15     SEL value      81.0

■      Tailjets       1      SEL value      77.9

10.5.29   If we calculate $E'_{st}/1000$ for this example, we obtain a number approximately equal to 205. An increase of one aircraft of each type during the four-hour period will increase $E'_{st}/1000$ by 33.5.  Likewise, an increase of 1 dB for each aircraft type will increase $E'_{st}/1000$ by 25.0. Hence it can be seen that the implied valuation of these changes is of the order of £12-15 (when multiplying by the unit value of 47 pence).  Even allowing for the fact that this should be approximately halved to take account of the lower sensitivity in Time Period 4, this seems extremely high.  Note that the change in (4-hour) LAeq associated with these changes is 0.67 for the increase in the number of aircraft, and, of course, 1 for the (uniform) increase in sound level.  These would be considered small changes.

10.5.30   Hence, although the form of the $E'_{st}$ model makes it difficult to generalise, there is a suggestion that the implied valuation is higher than would be considered reasonable.  We discuss the question of valuation further in Section 10.7.

10.5.31   The table below summarises the log-likelihoods for the various "pooled" models discussed here, in order of increasing likelihood.  All these models are of the "multiplicative" form, and have a single money coefficient:

**Table 10.5        Summary of log-likelihoods for alternative model formulations**

| Model formula | Model description | no. of parameters estimated | Log-likelihood |
|---|---|---|---|
| $U_s = 0$ | Null | 0 | -45091.42 |
| $U_s = \sum_{t=1}^{T} \beta_E . \beta_t . E_{st} + \beta_m . M + \varepsilon_0$ (10.5) | $E_{st}$ based on LAeq formula | 7 | -43590.28 |
| $U_s = \sum_{t=1}^{T} \beta_E . \beta_t . [10.\log_{10} E_{st}] + \beta_m . M + \varepsilon_0$ (10.6) | ditto | 7 | -42726.21 |
| $U_s = \sum_{t=1}^{T} \beta_E . \beta_t . [10.\log_{10} E'_{st}] + \beta_m . M + \varepsilon_0$ (10.8) | $E'_{st}$ factors SEL by 20, rather than 10 as in LAeq formula | 7 | -42536.9 |
| $U_s = \sum_{t=1}^{T} \beta_E . \beta_t . E'_{st} + \beta_m . M + \varepsilon_0$ (10.7) | ditto | 7 | -41766.4 |
| $U = \sum_{t=1}^{T} \sum_{a=1}^{A} \beta_{as} . \beta_t . N_{ast} + \beta_m . M + \varepsilon_0$ ("national model") | aircraft- and site-specific coefficients | 108 | -38992.68 |

**mva**consultancy

10.5.32    We can summarise the investigations in this section along the following lines.  While the "National model" provides a good fit to the data, it is difficult to generalise because it has so many site-specific parameters.  We therefore investigated the possible relationship between these parameters and the sound levels.  This also provided a basis for testing to see whether LAeq was a good proxy for disutility (= annoyance).

10.5.33    Initial testing suggested only a weak relationship between disutility and LAeq.  However, more detailed investigations suggested that with a higher weight for $\log_{10}(N)$ a stronger relationship could be obtained.  The weight of 20 (as opposed to the implied value of 10 in the Laeq formula) was suggested, and it is of interest that this is consistent with the work reported in the previous chapter on the relationship between annoyance and Lav and log Nav.

### 10.6    Sensitivity to Aircraft Noise by Time of Day

#### Evidence from ANASE

10.6.1    As described above, each trade-off exercise that respondents undertook related to a specific 4-hour period of the day, and involved directly trading off aircraft noise (implied from type and number) and money (presented as a grant).

10.6.2    The pooled National Model (Table 10.3) is formulated to disaggregate the effect of aircraft type and time of day into separate coefficients.  The five coefficients which represent the relative sensitivity of experiencing an additional aircraft movement in Time Periods 2 to 6 relative to Time Period 1 are re-stated in Table 10.6[32].

**Table 10.6        Time Period Coefficients from National SP Model**

| Period Number | Time of Day | Coefficient | t-statistic |
|:---:|:---:|:---:|:---:|
| 1 | 2300-0300 | 1.000 | FIXED |
| 2 | 0300-0700 | 0.732 | 18.5 |
| 3 | 0700-1100 | 0.545 | 19.9 |
| 4 | 1100-1500 | 0.543 | 19.7 |
| 5 | 1500-1900 | 0.603 | 20.5 |
| 6 | 1900-2300 | 0.633 | 23.0 |

10.6.3    The results show that the disutility of an additional aircraft is lowest during the daytime, with the smallest coefficients relating to the eight-hour daytime period from 0700-1500. Comparing the relative coefficients that measure disutility of an additional aircraft during the

---

[32] The time-period values are slightly different from the model (10.7) reported in the previous section.  On grounds of overall model fit, we think it is better to base the discussion in this section on the "National" model reported in Section 10.4

**mva**consultancy

period 2300-0300 (1.000) and the daytime period (averaged at 0.544), we can infer that the noise of an aircraft at night provides the same disutility to respondents as (1.000/0.544 =) 1.84 of the same aircraft during the daytime.  A further inference is that respondents consider the same noise to be 84% more annoying at night than the same noise during the day.

10.6.4    Relative to the **daytime**, and with some rounding, the sensitivity to the same aircraft noise at other periods are:

- 2300-0300: **80%** more annoying;

- 0300-0700: **35%** more annoying;

- 1900-2300: **15%** more annoying; and

- 1500-1900: **10%** more annoying.

10.6.5    The above reflects society's sensitivity overall.  However, implicit in these weightings are the proportion of people at home exposed to the noise – these weightings would be expected to differ for an individual at home throughout the day.  To explore this further, we have compared these time of day sensitivities with the proportion of the sample at home during each time period.

10.6.6    Figure 10.2 suggests a pattern between presence in the home and sensitivity by time of day, and further investigation showed the correlation to be statistically significant at the 90% level.
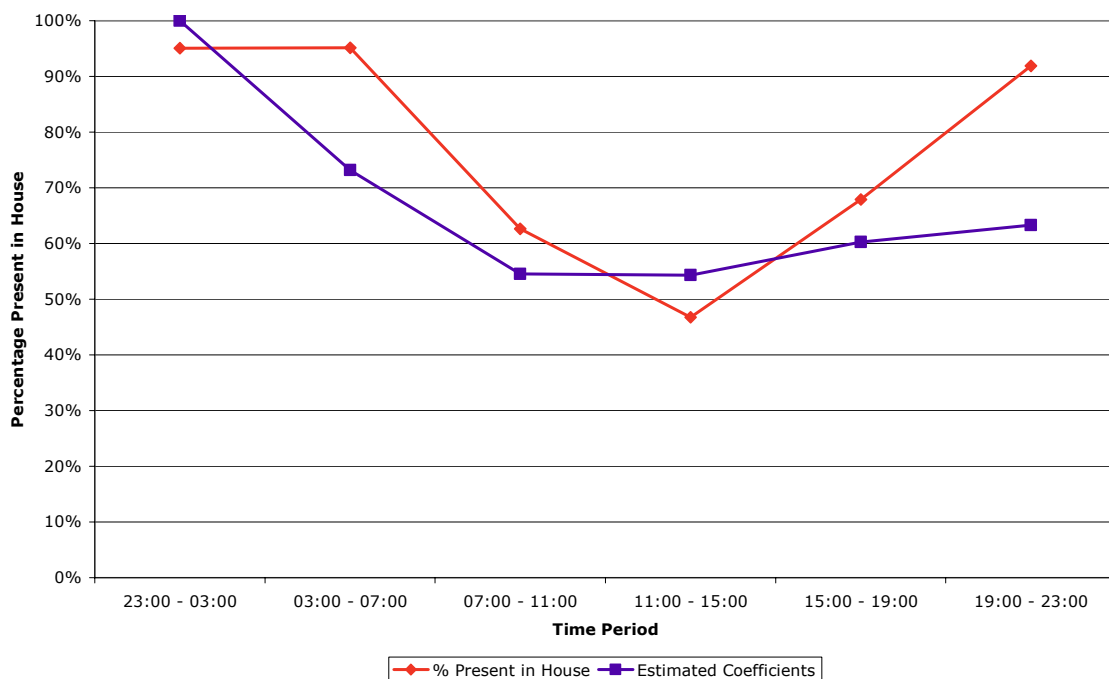


**Figure 10.2**        **Comparison of Time of Day Sensitivities and Presence in the Home**

10.6.7    The apparent pattern between presence in the home and sensitivity by time of day suggested that a model in which two parameters were estimated for each time period, one for respondents who were normally present in the home in that period and one for those who were not, may give a better fit to the data than the National Model.  (In practical terms, it

**mva**consultancy

was actually necessary to fit a model with a third set of time period parameters, for those respondents who did not specify whether or not they were present).  The loglikelihood of this model, at -38977, does indicate a statistically significant improvement in the explanation of the data over the National Model taking into account the number of additional explanatory variables.  (The model is reported in Table D21 of Appendix A10.2).  However, though most parameter coefficients are statistically significant (i.e. t-statistics exceeding 1.96), the 'at home' coefficients are not always greater than the 'not at home' coefficients.  Thus, in spite of the apparent correlation, it is **not** the case that the time period effects are being dominated by presence in the home.

10.6.8    The finding that people are differentially annoyed at different times of day regardless of whether they are at home or not is not necessarily counter-intuitive.  Firstly, those 'not at home' have indicated that they were unlikely to be at home (or in the immediate neighbourhood close to their home) for a given period, but they may well be at home for at least some of the 4-hour period sometimes.  Secondly, they may still be in a neighbourhood that is affected by aircraft noise at this time.  Thirdly, they may have relatives at home during this period so, again, they may reasonably attach disutility to aircraft noise at this time.  These results suggest there may be some other external factor that is also correlated with presence in the home and annoyance with aircraft noise.

10.6.9    A model that omitted respondents who stated that they were not annoyed by aircraft noise in response to the standard ISO noise questions gave very similar relative values to those derived in the National Model.  That is, the relativities between aircraft types within and across sites, time period and money coefficients all remain fairly unchanged.

10.6.10    There was some evidence to suggest that a minority of respondents had time of day sensitivities that were different to those identified in paragraph 10.6.4, but no explanatory factor could be identified that isolated this effect.

### Evidence from Other Research

10.6.11    A number of other studies have considered the relative annoyance caused by aircraft noise at different times of day.

10.6.12    **Hume et al (2003)**[33] found that the ratio of complaints to aircraft movements in Manchester was highest overnight (between 2100 and 0500), and at lunchtime (1100-1300). On average, the ratio was over 4.5 times higher for night-time periods than for the rest of the day.

10.6.13    **Carlsson et al (2004)**[34] found that values varied by time of day and between weekday and weekend, but were generally high in the evening: the period between 2200 and 0700 was not covered.   At the weekend, the ratio between evening and morning (0900-1200) coefficients was between 1.3 and 2.6.   Weekday results were more variable, with a coefficient over 8 times higher in the evening than the morning in one case, but annoyance apparently higher in the morning than in the evening in others.

---

[33] "Complaints caused by aircraft operations: an assessment of annoyance by noise level and time of day", Hume, Gregg, Thomas, Terranova, Journal of Air Transport Management 9 (2003)

[34] "The marginal values of noise disturbance from air traffic: does the time of day matter?", Carlsson, Lampi, Martinsson, Transportation Research Part D (2004)

**mva**consultancy

10.6.14   **Bristow and Wardman** (2004[3] and 2006[35]) use a fairly broad breakdown by time periods (early mornings, daytimes, evenings, night) and report values in Euros per aircraft from research at three airports.  Different SP exercises yielded different sets of values, depending on:

- the airport at which the research took place;

- whether the purpose of the study was 'masked' by asking people to trade other aspects of quality of life as well as aircraft noise;

- whether respondents were trading explicitly between time periods, or whether results for single time period exercises were being combined; and

- whether the changes being considered were improvements or deteriorations, or both.

10.6.15   Results varied between airports, but night values were high in all cases, and evenings and weekends generally valued more highly than daytime periods.  Excluding the results from Bucharest (where conditions were very different), the ratio of night-time (2200-0600) to daytime (0900-1800) coefficients ranged from around 2.5 to over 16 in Lyon, and from 3.7 to 7.3 in Manchester.  The relative values vary considerably, but are significantly higher than those found in the ANASE SP results.

10.6.16   Results for the evening period were both lower and more consistent than for night-times, generally around 2-3 times higher than the daytime period.  Again, the ANASE results show a lower weighting of time periods outside the main working day.

## 10.7    SP-based Willingness to Pay

10.7.1   We can use the results in Table 10.3 (that assumes a single monetary coefficient for all respondents), to give some overall impression of the implied willingness to pay per month per household for one less aircraft per day for different sound levels and for a given 4-hour period.  The implied willingness to pay values for the 4-hour **day-time period** of 1100-1500 (i.e. applying the time period 4 factor of 0.5425) by aircraft type, and across LAeq bands, is (per household per month):

*LAeq >60 dB*

- Jumbo:      £5 - £9 per aircraft      (min SEL =  84dB, max SEL =  95dB)
- Underwing:  £2 - £6 per aircraft      (min SEL =  82dB, max SEL =  89dB)
- Turboprop:  £2 - £3 per aircraft      (min SEL =  77dB, max SEL =  84dB)
- Tailjet:    £2 - £5 per aircraft      (min SEL =  67dB, max SEL =  84dB)

---

[35] "Valuation of aircraft noise by time of day: a comparison of two approaches", Bristow and Wardman, Transport Reviews, Vol 26, No. 4, 2006

**mva**consultancy

*LAeq 55.1-60 dB*

- ■    Jumbo:      £7 - £10 per aircraft     (min SEL =  79dB, max SEL =  90dB)
- ■    Underwing: £1 - £4 per aircraft     (min SEL =  77dB, max SEL =  87dB)
- ■    Turboprop: £1 - £2 per aircraft     (min SEL =  75dB, max SEL =  80dB)
- ■    Tailjet:      £2 - £3 per aircraft     (min SEL =  74dB, max SEL =  81dB)

*LAeq 50.1-55 dB*

- ■    Jumbo:      £4 - £9 per aircraft     (min SEL =  75dB, max SEL =  88dB)
- ■    Underwing: £1 - £2 per aircraft     (min SEL =  70dB, max SEL =  85dB)
- ■    Turboprop: £2 - £3 per aircraft     (min SEL =  75dB, max SEL =  76dB)
- ■    Tailjet:      £1 - £2 per aircraft     (min SEL =  63dB, max SEL =  83dB)

*LAeq up to 50 dB*

- ■    Jumbo:      £2 - £7 per aircraft     (min SEL =  75dB, max SEL =  84dB)
- ■    Underwing: £2 - £7 per aircraft     (min SEL =  69dB, max SEL =  84dB)
- ■    Turboprop: £2 - £5 per aircraft     (min SEL =  70dB, max SEL =  77dB)
- ■    Tailjet:      £2 - £3 per aircraft     (min SEL =  67dB, max SEL =  75dB)

10.7.2    The average monetary value of each aircraft generally declines in line with aircraft SEL range.   Furthermore, inspection of the within-site aircraft valuations within the National Model shows the vast majority to be consistent with relative SEL values - i.e. aircraft with a higher SEL has a higher relative value than aircraft at the same site with a lower SEL.  (The relationship between respondents' valuations and SEL was briefly explored earlier in this Chapter where we investigated the suitability of LAeq as a proxy for community annoyance).

10.7.3    However, there is considerable over-lap between SEL values and LAeq bands, with aircraft SEL values often higher at lower LAeq sites.  It would seem that the SP valuations reflect the stimulus material (which was a play-back of the actual aircraft SEL for the site) more than the general noisiness of the area, defined by LAeq.  The relationship between SP values and LAeq is, therefore, less strong than that between SP value and SEL.

10.7.4    In comparison with Bristow and Wardman[36], these willingness to pay values seem high. They quote their SP results in terms of Euros per week per aircraft per hour (the periods to which the aircraft relate vary in length): no analysis is reported relating to the separate aircraft types.  For Manchester, the daytime value is quoted as €0.84 per aircraft.  Dividing this value by 4 (to correct for the length of the ANASE period) and adjusting to sterling per month, we have a comparative monthly willingness to pay value of just under £1 per (generic) aircraft – at the bottom end of the ANASE valuations.

---

[36] "Valuation of Aircraft Noise using stated preference techniques", Bristow and Wardman, InterNoise 2004

**mva**consultancy

10.7.5    The ANASE SP values are very high when one considers the number of aircraft – even Jumbos – that would need to stop flying overhead in order to reduce the overall LAeq by 1dB at the site.   We have not been able to explain the cause of this considerable disparity between the ANASE SP-based valuation and valuations based on hedonic pricing and contingent valuation.  We conjecture that the reason is one or more of the following:

- SP is, in fact, an inappropriate tool for deriving monetary values in the context of aircraft noise;

- faults in the adopted SP design, data capture or analysis;

- an inability, on the respondents' part, to correctly associate a reduction of aircraft at a given time of day on their overall level of annoyance with aircraft noise; and

- an inability, on the researchers' part, to correctly associate a reduction of aircraft at a given time of day on respondents' perceived noise exposure (i.e. translating a marginal value per aircraft into a marginal value per dB Leq).

10.7.6    In essence, the questions seems to be: what in the data collection/analysis/interpretation processes might generate results that are a magnitude out?; Or, is there a generic and fundamental problem with SP?.  In what follows, we reflect on these various possibilities.

10.7.7    In the authors' view, there is no reason to think that SP will inherently produce vastly over-estimates of monetary valuations of goods and services.  The SP technique has been used for more than twenty years and has been validated (through the use of observed data) on many occasions.  SP, and other survey-based techniques such as CVM, can provide over-estimates of willingness to pay if derived in the context of separate, individual components, which are then summed to give an overall valuation.  This is known as the 'package' effect or 'part-whole bias' and recognises that an overall value of a good or service may not equal the sum of the values of its individual parts.   However, though there is an element of disaggregation within the ANASE SP survey (namely the valuations by 4-hour time period), this phenomenon cannot be the main problem since the valuations for an individual period is still very high.

10.7.8    We are also of the view that the SP design, data collection and analysis accurately captures the views and preferences of respondents.  This is based on anecdotal information gained throughout the study (i.e. cognitive assessment of respondents' decision processes when considering the SP trade-offs), and we are confident that respondents considered the SP options that they were presented with to be realistic, and that they stated their 'true' preference from each choice-set.   This view is supported by the degree of internal consistency of the SP data (within individual responses, within individual sites and across sites).

10.7.9    We believe the area of greatest uncertainty is the link between respondents' willingness to pay for a reduction in aircraft (e.g. around £5 a month for one less jumbo every day during a certain 4-hour period) and their assumed improvement in their quality of life (through reduced annoyance by aircraft noise).  The ANASE study has revealed that a change in the number of aircraft is perceived to have the greatest effect on reducing aircraft noise annoyance.  However, more research is needed to explore how accurately people associate a reduction in aircraft numbers with a change in overall sound levels.   It may be that respondents perceive that a reduction of a few jumbos during a particular period of the day would have considerable impact on their overall sound levels yet not even notice the

**mva**consultancy

reduction, in reality.

10.7.10    A further uncertainty is the ability to accurately represent people's willingness to pay for reduced aircraft noise in the units most commonly used by economists and policy-makers – i.e. per dB LAeq.  In fact, it is extremely difficult to estimate the change in aircraft that would correspond to a reduction of 1dB LAeq at a particular site (even for acoustics experts).  The difficulty in translating the SP (marginal aircraft) values into an equivalent per unit per decibel change, such as NSDI was known.  Unfortunately, as noted in paragraph 10.3.12, the proportional SP model did not explain the data well, and we therefore did not pursue this model further.  However, this form of model might have been more suitable for interpreting SP money values for comparing against equivalent NSDI values derived from CVM or hedonic pricing.

10.7.11    A further possibility allowed for in ANASE was to adopt an overall willingness to pay value for reduced aircraft noise, derived from the CVM data, and in some way scale the SP-derived values.

## 10.8    ANASE CVM Research

10.8.1    As noted earlier, CVM has been widely used within the environmental research community, for valuing environmental and amenity factors.  It aims to obtain a **direct** assessment of the survey respondent's valuation of an improvement of interest (such as reduction in noise), in terms of willingness to pay (WtP) or willingness to accept (WtA).

10.8.2    In our study we have endeavoured to reduce some of the more obvious opportunities for bias encountered with so-called "open-ended" form of CVM questions, by offering respondents a sequence of money amounts and noting when the response changes from willing to not-willing to pay.  The question asked of respondents was as follows:

> *"Now imagine that there is a house exactly like the one you live in.  It has access to all the same facilities such as shops, personal services, doctor, schools, etc.  These facilities are the same distance away from this imaginary house as they are from your home.*
>
> *There is one difference between this imaginary house and where you live now.  This imaginary house has NO NOISE FROM AIRCRAFT and it would cost your household £X a week more live to there"……So, there would be no noise from aircraft, the same amount of noise from road traffic and your household would have £X less to spend each week – or approximately £Y less each year to spend on other things.*
>
> *Would you prefer to get rid of all aircraft noise, but have £Y less per year for your household to spend on other things?  Or would you prefer to put up with the aircraft noise and be able to spend that £Y per year on other things?."*

10.8.3    In case the WtP valuation obtained is an artifact of the starting amount in the bidding process, we presented a range of 'starting' values for "£X", randomly offered to different respondents:

- £1 a week (£50 a year);

- £5 a week (£250 a year);

- £10 a week (£500 a year); and

- £20 a week (£1000 a year).

10.8.4    Given their response to this first CVM question in terms of £X, respondents were then asked to look at a range of monetary values and state which would be the maximum that they would be willing to pay in order to eradicate all aircraft noise.

> "*Please look at this card* [depicting a range of monetary values] *and tell me what is the MOST you would be prepared to give up in terms of your household spending on other things to get rid of all aircraft noise heard at your home.*"

10.8.5    Respondents who said that to have no aircraft noise was worth at least the value originally presented were asked to consider a choice-set of higher money values rising to more than £60 a week.  Respondents who said that to have no aircraft noise was not worth at least the value originally presented were presented with a choice-set of lower money values down to zero.  Therefore, following a second iteration, a maximum willingness to pay value (including £0) to go from the current noise situation to a 'no noise' situation was obtained for each respondent.  This use of a range of alternatives from which the respondent identifies their household's maximum willingness to pay is sometimes referred to as the 'Payment Card' format, and is the basis for the monetary valuations reported in the remainder of this section.

10.8.6    The CVM questions were included within the ANASE survey to provide a direct monetary value to (the lack of) overall aircraft noise.  It was further considered that this would then provide a comparison against the more detailed SP valuations.  The CVM questions also provide a direct mechanism for identifying potential **protest voters.**  These are defined as respondents who assign zero valuations because they do not accept the principle that anyone should have to pay for peace and quiet, irrespective of to what extent they might value peace and quiet for themselves.  Protest voters might respond differently to the SP trade-offs, compared with other respondents - for example, they may choose to ignore the money variable when providing responses to the SP.

**CVM WTP Results**

10.8.7    Table 10.7 shows the distribution of the maximum amount that respondents stated that they were willing to pay in order to eradicate aircraft noise, classified by their response to the ISO standard annoyance question.

**mva**consultancy

**Table 10.7   Overall Household Willingness to Pay to Eradicate Aircraft Noise, and Variation by Reported Level of Annoyance**

| Weekly amount | Reported Annoyance Level | | | | | |
|---|---|---|---|---|---|---|
| | Overall | Slightly | Moderately | Very | Extremely | Not at all |
| £0 | 57% | 63% | 47% | 40% | 34% | 84% |
| £1 - £10 | 25% | 23% | 34% | 32% | 31% | 10% |
| £11 - £20 | 7% | 5% | 7% | 11% | 12% | 1% |
| £21 - £30 | 3% | 3% | 4% | 4% | 6% | 0% |
| £30+ | 3% | 2% | 3% | 6% | 9% | 0% |
| Missing | 5% | 4% | 5% | 7% | 8% | 4% |
| Base (N) | 2,730 | 523 | 603 | 429 | 435 | 740 |

10.8.8   Across the whole sample, 57% of respondents are **not** prepared to pay anything to get rid of aircraft noise.  For many of these respondents, the zero valuation appears to be due to not being annoyed by aircraft noise (84% of respondents who report that they are not at all annoyed with aircraft noise are not willing to pay anything) but for other respondents this is not the case (34% of people reporting to be extremely annoyed with aircraft noise would not be willing to pay anything to eradicate it).

10.8.9   The reasons given why respondents were not willing to pay anything to eradicate aircraft noise were as follows:

■   32% not bothered by aircraft noise;

■   14% should not pay for it on principle;

■   4% cannot afford to pay;

■   4% it is not a priority;

■   2% other; and

■   1% not answered.

10.8.10   This finding suggests a minority (14% of all respondents) were not prepared to pay anything "out of principle"; these respondents have been classified as "protest voters".   These respondents gave a range of responses to the ISO annoyance question (49% very or extremely annoyed; 27% moderately annoyed; and 24% slightly or not at all annoyed) and reside in areas across the ANASE LAeq range (24% <50dB; 21% 51-55dB; 33% 56-60dB; and 22% >60dB).  Protest voters have been excluded from further CVM analysis.  All other zero willingness to pay responses have been included.

**mva**consultancy

10.8.11    Figure 10.3 shows the relationship between mean willingness to pay of (non-protest-voting) respondents by reported annoyance level (as defined in Chapter 7) at site (Model 10.9) and individual (Model 10.10) level.  (The willingness to pay and annoyance values are obtained for the individual so the individual model is preferred over the site-level model, which is retained for comparison with other reported graphs).

**Model 10.9  Willingness to Pay (Site level) =  -2.3 + 0.17 x Mean Annoyance**

| R² | Adjusted R² | N | Intercept | Annoyance |
|---|---|---|---|---|
| 0.454 | 0.447 | 76 | -2.33 | 0.17 |
| | | | -2.56 | 7.84 |

**Model 10.10    Willingness to Pay (Individual level) =  -0.65 + 0.15 x Mean Annoyance**

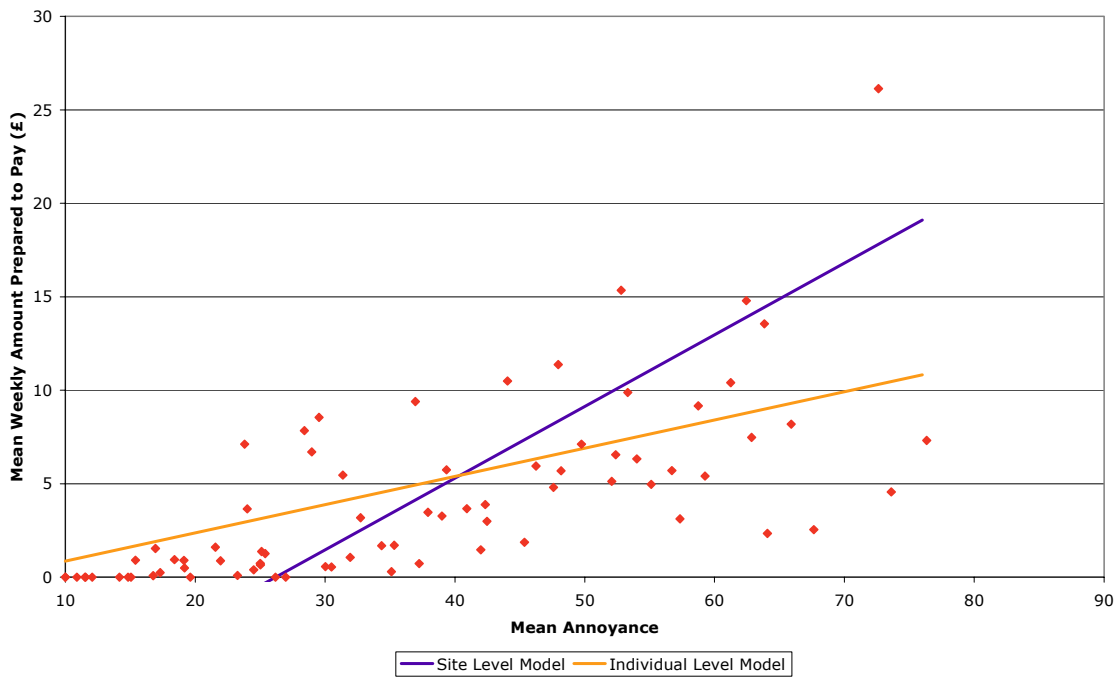| R² | Adjusted R² | N | Intercept | Annoyance |
|---|---|---|---|---|
| 0.125 | 0.125 | 2355 | -0.65 | 0.15 |
| | | | -1.55 | 18.38 |



**Figure 10.3 Mean Amount Respondents Willing to Pay (excluding protest voters) against Mean Reported Annoyance**

10.8.12    As might have been expected, there is clear correlation between the amount people who are not 'protest voters', state they would be willing to pay to eradicate aircraft noise and their degree of annoyance reported with aircraft noise.

10.8.13    Given that the CVM question relates to the "removal of all noise due to aircraft", the implied reduction will depend on the **current** level of noise.  Thus, we should expect the difference in valuations to be related to the difference in the current level, implicitly providing a "per dB"

**mva**consultancy

estimate.

10.8.14    Figure 10.4 plots the mean weekly amount that (non-protest-voting) respondents at each site were prepared to pay to remove aircraft noise against the site LAeq value.  There is some indication that as LAeq increases, the amount of money that respondents were prepared to pay also increases.  [Note, we have again plotted the site-level (Model 10.11) and individual-level (Model 10.12) regression lines, but only the willingness to pay values vary at an individual level].

### Model 10.11  Willingness to Pay (Site level) =  -12.5 + 0.34 x LAeq

| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq |
|---|---|---|---|---|
| 0.220 | 0.206 | 56 | -12.45 | 0.34 |
|  |  |  | -2.75 | 3.91 |

### Model 10.12  Willingness to Pay (Individual level) =  -8.3 + 0.27 x LAeq

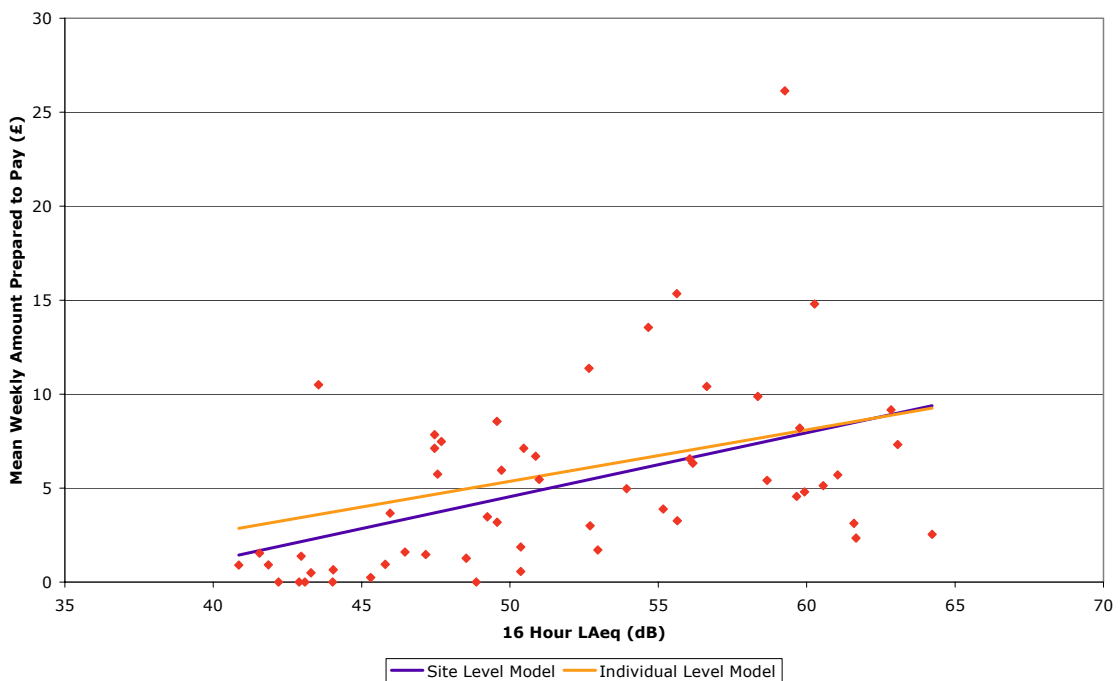| $R^2$ | Adjusted $R^2$ | N | Intercept | LAeq |
|---|---|---|---|---|
| 0.018 | 0.018 | 2096 | -8.33 | 0.27 |
|  |  |  | -3.51 | 6.23 |



**Figure 10.4 Mean Weekly Amount Non-Protest Voters are Willing to Pay to Remove Aircraft Noise**

10.8.15    There is a wide range of mean willingness to pay at each LAeq value.  For areas with a LAeq less than 45 dB, the mean amount that respondents are willing to pay to remove aircraft noise is not significantly different from £0 at the 95% confidence level.  This corresponds with DfT research reported in WebTAG[37] that showed that for road and rail noise, the

---

[37] WebTAG, TAG Unit 3.3.2, Noise, Section 1.3

**mva**consultancy

monetary value that people place on sound levels below 45 dB was not significantly different from zero at the 95% confidence level. At sites with a LAeq of around 55 dB, the mean willingness to pay values range from less than £5 to more than £15 a week.

10.8.16    The 'best-fit' linear regression line at an individual level has a gradient (i.e. coefficient on LAeq) of 0.270 suggesting a monetary value of £0.27 per week, or **£14 per household per annum per dB** change in LAeq. Although this coefficient is significant at the 1% level, the adjusted $R^2$ is very small so a large proportion of the variation is not accounted for by LAeq.

10.8.17    Some of this variation appears to be directly related to the starting value each respondent was presented with. Figure 10.5 demonstrates the considerable difference in the amount residents are willing to pay to eradicate aircraft noise. Respondents initially presented with a £20 per week trade-off, were willing to pay between £1 and £3 more than those initially presented with a £5 per week trade-off. This difference represents a doubling in valuation in low noise areas, and means that the CVM-derived willingness to pay values should be treated with some caution.
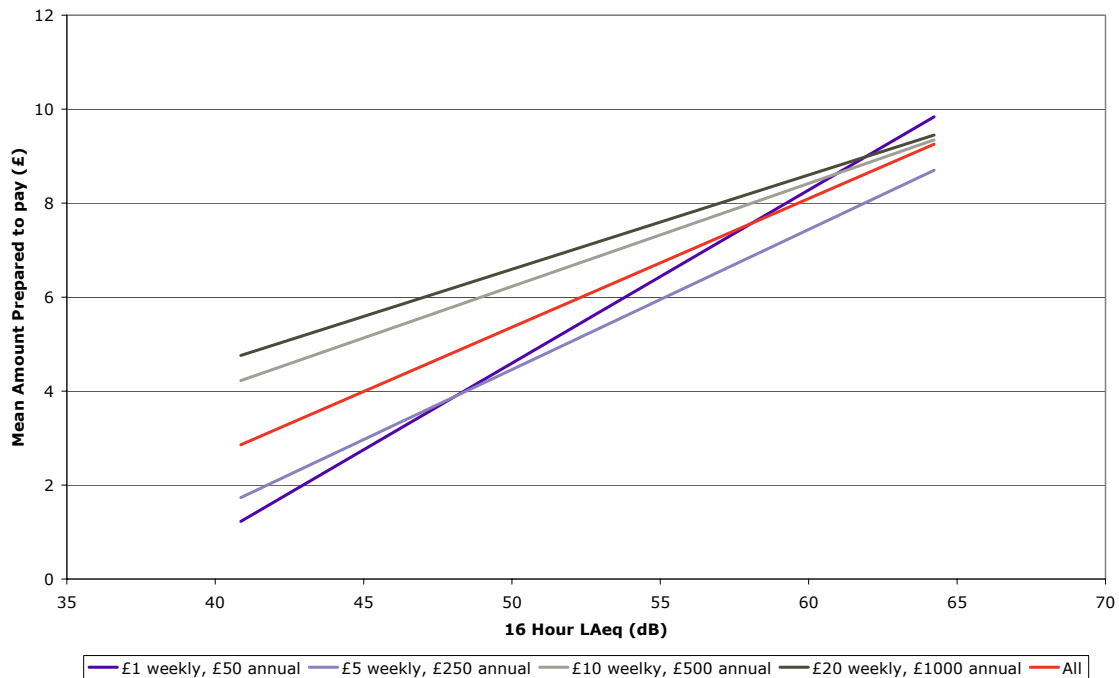


**Figure 10.5 Weekly Amount Non-Protest Voters are Willing to Pay to Remove Aircraft Noise by Initial Starting Point**

10.8.18    Given this variation in willingness to pay by starting value, and the fact that more than half of respondents (57%) indicated a zero value, we have only analysed this type of direct trade-off data using simple (Ordinary Least Squares) regression techniques. However, we have investigated improvements of fit of the individual-level regression model at a segmented level. Regression equations tested for goodness-of-fit were of the form:

Mean amount willing to pay (excluding protest voters) = a + b LAeq + c Variable

Mean amount willing to pay (excluding protest voters) = a + (b + c Variable) LAeq

**mva**consultancy

10.8.19   The best individual-level model (in terms of statistical goodness-of-fit) was with LAeq and mean income, as shown in Model 10.13.

**Model 10.13  Amount Willing to Pay =  0.92 + 3.69x10$^{-6}$x Household Income x LAeq**

| R$^2$ | Adjusted R$^2$ | N | Intercept | LAeq x Income |
|---|---|---|---|---|
| 0.084 | 0.084 | 1737 | 0.92 | 3.69 x 10$^{-6}$ |
|  |  |  | 1.69 | 12.64 |

10.8.20   This suggests a willingness to pay of **£270 per household per annum** to eradicate aircraft noise for those in households with an income of £20,000 and at 57 LAeq; and a marginal change is valued at **£3.84 per household per annum per dB** LAeq for these households. Similarly, for households with an income of £60,000 and at 57 LAeq, there is a suggested willingness to pay of **£700 per household per annum**, and a marginal change is valued at **£11.50 per household per annum per dB** LAeq.  (Note, the linear nature of the model means that the above marginal values apply for all levels of LAeq.)

10.8.21   Non-linear modelling of the CVM data gave no support for willingness to pay exponentially rising as noise exposure levels increase.  Indeed, a simple quadratic model suggested that valuations plateau at very high noise areas.

10.8.22   Overall, the results of the ANASE CVM research suggest a statistically valid relationship between willingness to pay to eradicate aircraft noise, and the level of current aircraft sound levels in the form of LAeq (for the LAeq range 40-65 dB covered in the research).  The implied willingness to pay of £3.80 – 11.50 per dB LAeq reduction per annum for respondents, depending upon household income level, is within the range of values derived in other CVM and Hedonic Pricing studies.

10.8.23   At an aircraft noise level of around 60 LAeq, the CVM results imply a willingness to pay of around £250-700 per annum, depending upon household income.  This is equivalent to somewhere around £2,000 to £8,000 in terms of equivalent capital value.  This works out at 1-4% on a typical £200,000 house, which could then be considered equivalent to an NSDI (noise sensitivity depreciation index) of between 0.2 and 0.8% per dB.

10.8.24   Bjorner's CVM study[38] quotes average valuations per person per year of nuisance from traffic noise at each level of reported annoyance.  At high sound levels (75 dB), Bjorner's study produces estimates of willingness to pay of €10 per dB per annum per person.  A hedonic pricing study of transport-related noise by Bateman et al[39] indicates willingness to pay levels of £30 per annum per dB at medium sound level sites (55 dB baseline), rising to more than £100 per annum per dB at higher sound level sites.

---

[38] "Combining socio-acoustic and contingent valuation surveys to value noise reduction", Thomas Bue Bjorner, Transportation Research Part D (2004) 341-356

[39] "The valuation of transport-related noise in Birmingham: non-technical report to the DfT", Bateman, Day and Lake, Transportation Research Part D 9 (2004)

**mva**consultancy

## 10.9    Summary of Main Points

■    There is an implied willingness to pay to remove all aircraft noise of £4 - £11 per annum per dB reduction in LAeq per household, depending upon household income level.  This value is in the same ball-park as recent valuations based upon Hedonic Pricing;

■    The results of the SP survey has shown strong internal consistency and statistical validity, with a clear indication that aircraft SEL, aircraft type, time of day and personal characteristics (in particular household income) greatly influence community annoyance and willingness to pay to reduce it;

■    A single jumbo has the same disutility as approximately 3 underwings or turboprops, or 4 tailjets;

■    The SP results have shown people to be much more sensitive to aircraft noise at night (particularly around midnight and the early hours thereafter).  In contrast, people are least sensitive to aircraft noise in the morning and early afternoon;

■    A single jumbo during the sensitive 2300-0300 period, has the same disutility as approximately 5 underwings; 6 turboprops; or 7 tailjets during the daytime;

■    These time of day sensitivities seem intuitively plausible and were internally consistent showing significant correlation with presence in the home.  The findings are also comparable with other research, which also suggest the ordering of night-time; evening, daytime in descending order of sensitivity.  However, the ANASE results indicate a lower ratio of annoyance than the other studies;

■    As a proxy for predicting changes in community annoyance in relation to a change from  current  sound levels, our SP research supports the view that the role of **number** of events needs to be higher than that implied in the LAeq index.

**mva**consultancy

# 11 Conclusions

## 11.1 Introduction

11.1.1 This chapter presents our conclusions from the ANASE study. The structure of the presentation is based on addressing the three research objectives described in Chapter 1.

## 11.2 Objective 1: Re-assess Attitudes to Aircraft Noise in England

11.2.1 Analysis of the ANASE survey data has shown that as the sound level indicator LAeq increases, the annoyance levels of respondents also increase, and that a large proportion of measured variation in annoyance can be accounted for by LAeq. However for a given LAeq, there is a range of reported annoyance indicating that annoyance is not determined solely by the amount of aircraft sound as measured by LAeq.

11.2.2 Our analysis showed that respondent's household income and SEG were the main additional influences on the level of annoyance. Once these factors are accounted for there are no further significant location effects (ie those affected by aircraft at Heathrow, for a given LAeq and income, are no more annoyed than those living close to other airports covered in the study).

11.2.3 For both this study (ANASE survey work carried out in 2005), and the ANIS survey (undertaken in 1982), mean annoyance scores have been calculated, and these have been compared to the LAeq metric.

11.2.4 In both studies, LAeq is effective at explaining much of the variation in respondents' reported annoyance.

11.2.5 However, this comparison has also shown that for the same amount of aircraft noise, measured in LAeq, people are more annoyed in 2005 than they were in 1982. For an LAeq of 57 (identified in the DORA report as the onset of significant annoyance), the modelled value of annoyance for ANIS is 39 (slightly higher than "a little annoyed" on the ANIS scale), whereas for ANASE it is 53 (somewhat higher than "moderately annoyed" on the ANASE scale). Thus annoyance is about 14 points greater in ANASE than it was 23 years ago (where a difference of one category on the ANASE annoyance scale is allocated 20 points).

11.2.6 Sensitivity tests confirm the conclusion that, measured in LAeq, people are more annoyed in 2005 than they were in 1982, though the size of the difference is affected by assumptions made.

11.2.7 A particular issue affecting the size of the difference is whether the introduction of noise playback equipment into the respondent's home generated an exaggerated response to the annoyance rating question. There is no statistical support for that effect from the ANASE data; substitution of the CAA LAeq estimates at Heathrow produces a significant effect.

11.2.8 Taking the worst case assumptions about the "equipment effect", the level of difference in annoyance reduces from our central case of 14 points to 7 points. Conversely, assumptions about "aligning" the ANASE and ANIS annoyance ratings increases the level of difference in annoyance to 21 points.

**mva**consultancy

11.2.9	The sensitivity analyses demonstrate that the relationship between LAeq and reported annoyance is **not** stable over time, and that this is a robust result.

11.2.10	If LAeq **is** an appropriate proxy measure of annoyance at a point in time, one possible explanation of the increase in reported annoyance for a given LAeq, between 1982 and 2005, may be a combination of changes in income/standard of living (which were significant cross-sectional factors in ANASE) and changes in attitudes within society.

11.2.11	The evidence from ANASE suggests that income growth alone is not sufficient to account for the difference.  There is common agreement that people today have higher expectations of a peaceful living environment, are less tolerant of environmental intrusion, and might consequently be less accepting of aircraft noise.  This view is supported by social trend data.  While both income and taste effects are likely to be important, it is not possible to identify their relative strength from our research: they are, of course, closely correlated.

11.2.12	An alternative hypothesis is that LAeq **is not** the appropriate measure, and that annoyance in both studies would correlate better with another sound level indicator.  This hypothesis is examined in the discussion relating to Objective 2, below.

11.2.13	The SP results have shown people to be much more sensitive to aircraft noise at night (particularly around midnight and the early hours thereafter).  In contrast, people are least sensitive to aircraft noise in the morning and early afternoon.  Ideally, therefore, a metric that reflects attitudes to aircraft noise should reflect these time of day sensitivities better than the existing LAeq - which does not weight by time of day.

## 11.3	Objective 2: Re-assess their Correlation with The LAeq  Index

11.3.1	Models were estimated which predicted mean annoyance values using LAeq.  These showed that the best fitting model, with around three-quarters of the variation explained, is a linear relationship between annoyance and LAeq.  However a logistic model, which produces an almost identical fit to the basic linear model, has the added advantage that it is bounded to the mean annoyance scores.

11.3.2	The modelling work also showed that respondents were less sensitive to changes in sound level below 42 LAeq and above 59 LAeq, adding support to a logistic form.  There was no threshold, or discontinuity, in the relationship between mean annoyance and LAeq.

11.3.3	The ANIS and ANASE surveys allowed us to compare the correlation of reported annoyance with LAeq at two points in time.  Over the period between the two surveys, there has been a substantial change in the make-up of aircraft, with many more aircraft in 2005 but with a lower (average) sound level than in 1982.

11.3.4	We found that the relationship between annoyance and sound level was strong for ANIS, but there was little relationship between annoyance and aircraft numbers.  The converse was the case for ANASE.  Therefore, the changes in reported annoyance for a given LAeq between 1982 and 2005 may reflect the changes in the composition of number and sound level that people are exposed to, suggesting a **different** formulation to that implied by LAeq.

11.3.5	An NNI-type measure gives a larger weight to the number of aircraft relative to the sound level than LAeq, and comparisons of the ANIS and ANASE mean annoyance against the NNI-

**mva**consultancy

type metric showed that the two datasets were much more closely aligned with the NNI-type measure than LAeq.

11.3.6    However, the relationship between reported annoyance, sound level and the number of aircraft has not been stable over time.  The weight on aircraft numbers (relative to sound level) has risen from 6 in ANIS to over 20 in ANASE, so the contribution of aircraft numbers to annoyance has increased quite markedly

11.3.7    Because of its instability over time, use of the LAeq measure to predict future levels of annoyance may be misleading.  In particular, where numbers of aircraft are increasing significantly, the ANASE data suggest that under-prediction of annoyance is likely.

11.3.8    Although the NNI-type index is also not stable over time, with the later ANASE result giving greater weight to aircraft numbers, the ANASE result is relatively insensitive to a weight greater than 20, so an NNI type measure may provide a better tool for predicting annoyance from aircraft noise.

11.3.9    **Overall**, we consider that while LAeq continues to be a good proxy for measuring community annoyance at a point in time, the relationship between LAeq and annoyance is not stable over time.  Income growth has led to some increase in reported annoyance between the ANIS and ANASE surveys, but is unlikely to provide a full explanation for the difference in attitude which is apparent.  There is evidence that intolerance of aircraft noise has grown.

11.3.10    The results from the attitudinal work and the SP analysis both suggest that LAeq gives insufficient weight to aircraft numbers, and a relative weight of 20 appears more supportable from the evidence than a weight of 10, as implied by the LAeq formulation.  An NNI – type measure appears to offer a stronger basis than LAeq for estimating future levels of annoyance in response to changing numbers and types of aircraft.

## 11.4    Objective 3: Examine Willingness to Pay to Remove Aircraft Noise

11.4.1    The results of the SP survey have shown strong internal consistency and statistical validity, with a clear indication that aircraft SEL, aircraft type, time of day and personal characteristics (in particular household income) influence annoyance and willingness to pay to reduce it;

11.4.2    On average, a single jumbo has the same disutility as approximately 3 underwings or turboprops, or 4 tailjets;

11.4.3    The SP results have shown people to be more sensitive to aircraft noise at night (particularly around midnight and the early hours thereafter).  In contrast, people are least sensitive to aircraft noise in the morning and early afternoon;

11.4.4    These time-of-day sensitivities seem intuitively plausible and are also comparable with other research, which also suggests the ordering of night-time, evening, daytime in descending order of sensitivity.  However, the ANASE results indicate a lower ratio of annoyance than the other studies;

11.4.5    As a proxy for predicting changes in community annoyance in relation to a change from the current noise environment, our SP research supports the view that the role of **number** of

**mva**consultancy

events needs to be higher than that implied in the LAeq index.

11.4.6    Unfortunately, despite the internal consistency, the implied valuations from the SP are much higher then may be considered plausible, when translated into a "per dB" value.

11.4.7    Valuations are also available from the CVM analysis. Here, there is an implied willingness to pay to remove all aircraft noise of £11-18 per annum per dB reduction in LAeq for respondents, depending upon household income level.  However, although this value is in the same ball-park as recent valuations based upon Hedonic Pricing, we have some reservations about the data, both because of the large proportion of respondents professing zero willingness to pay, and the apparent influence of the initial starting point in the "bidding" process.

11.4.8    Overall, therefore, we do not think that the valuations from either method are safe, and it will probably be necessary to rely on sources based on Hedonic Pricing.  Nonetheless, the relative valuations – particularly those relating to time of day variation – can be used.

### 11.5    Recommendations for Further Research

11.5.1    The ANASE study has produced a range of interesting results from the survey data collected. However the research has raised a number of issues, some of which can be addressed with more detailed analysis of the current data, and some which will require supplementary data collection.  We set out below a number of research areas where further work is likely to be fruitful:

- further analysis at the level of the individual household.  The majority of our analysis of the ANASE data has been at the level of the site, since that is the level at which sound data was estimated.  However the dataset contains 2733 interviews and it would be very informative to conduct further research at the level of the individual household.  This would take the form of formal statistical modelling, in contrast to the CHAID analysis done so far.  Such research on the larger sample would enable the influence of individual variables such as income to be better understood, and coefficients to be estimated with greater precision.  Such analysis would also enable the potential "equipment effect" to be explored more thoroughly.  It seems feasible to calculate sound levels for individual households by using the INM models to estimate the measures at the household address point.  In theory, this could also be achieved at each ANIS site, depending on the input data available.

- development of an annoyance index based on NNI incorporating time-of-day effects to reflect the SP findings of relative annoyance by time of day.  The Lav and Nav data would need to be calculated for different time periods for the full survey sites where both SP and attitudinal data were available.

- further time of day SP analysis to better understand non-zero values for periods when the respondent is 'not at home'.  The ANASE research was required to capture time of day sensitivities for all residents living in close proximity to aircraft noise, not just those who stay in all/most of the time.  However, in the spirit of research it would be very interesting to explore whether those living at home have zero values in time periods when they are away from the home; and time of day relatives amongst full-time workers who work away from home compared with retired people, etc.  An extension of this exploration would be to exclude 'zero-valuers' from the analysis to

**mva**consultancy

see how the time of day weightings change, and may help to explain any differences compared with other research concerning time of day research;

- more sophisticated analysis of the CVM data;

- further (questionnaire-based) study of issues arising from ANASE – including re-interviewing some respondents to ask for their underlying rationale for their attitudes – to test different interpretations of findings.

- re-examination of the implications of the LAmax cut-off in relation to Lav and Nav - and test whether event duration (such as in SEL), variability, etc. have a confounding impact.

- research to explore how accurately people associate a reduction in aircraft numbers with a change in overall noise exposure and perceived change in quality of life/annoyance.  An extension of this would be to explore associations by time period; and examine any possible 'package' effect (in particular, how valuations for a given 4-hour period should be aggregated to give values across a 34-hour period).  This would assist in interpreting and translating SP-derived values from ANASE.

- research to explore how people react to different **experienced** (and subsequently measured) levels of noise.  Respondents do not actually experience different levels of noise and there is, therefore, considerable uncertainty as to how people weigh-up the likely impact on their annoyance of a change in noise exposure.  This would be extremely challenging but will be an essential means of validating and interpreting SP results in future.

- further investigation of the functional relationship between utility and the combination of sound level and number of events, including attempts to derive a robust model that contains SEL and is based on movements.  There is potential to include site-weightings and/or aircraft type weightings to reflect possible other acoustical differences between sites and aircrafts.  This would reduced the unwanted site specific effects obtained in ANASE and the noted 'diluting' of the effect of SEL on utility.

- research to explore the existence of any thresholds in valuations, and non-linearities such as gains and losses in movements having different valuations; and unit values being proportional to the level of movements.

- additional restricted survey interviews with people working from home to further explore the possible confounding effect on annoyance of working from home.

- additional restricted survey interviews with residents around Stansted to further explore the possible 'special' effects at this airport.

- additional restricted survey interviews with residents around Gatwick (and, possibly, East Midlands) to test for the possible confounding effect on annoyance of greater night-time aircraft noise.

- as above but undertaking full SP interviews so as to inform further the time of day relativities, derived from the ANASE SP work.

- undertaking national annoyance/sound measurement studies on a more frequent basis.  This could be on a smaller scale (say, with a sample size of 500) restricted survey interviews every 5 yrs.

mvaconsultancy

11.5.2    Further research of this nature would assist the industry's understanding of how community annoyance changes over time, and provide additional insight into potential confounding factors, and special cases, which policy-makers may need to take into account in future planning.

**mva**consultancy

MVA Consultancy provides advice on transport and other policy areas, to central, regional and local government, agencies, developers, operators and financiers.

A diverse group of results-oriented people, we are part of a 350-strong team worldwide.  Through client business planning, customer research and strategy development we create solutions that work for real people in the real world.

For more information visit www.mvaconsultancy.com

**Head Office**
MVA House, Victoria Way
Woking, Surrey GU21 6DD  United Kingdom
T: +44 (0)1483 728051  F: +44 (0)1483 755207

**Birmingham**
Second Floor, 37a Waterloo Street,
Birmingham, B2 5TJ, United Kingdom
T: +44 (0)121 233 7680  F: +44 (0)121 233 7681

**Dubai**
PO Box 123166, 803-805 Arbift Tower, Baniyas Road,
Deira, Dubai, UAE
T: +971 (0)4 223 0144  F: +971 (0)4 223 1088

**Dublin**
1st Floor, 12/13 Exchange Place, IFSC, Dublin 1, Ireland
T: +353 (0)1 542 6000  F: +353 (0)1 542 6001

**Edinburgh**
Stewart House, Thistle Street, North West Lane
Edinburgh EH2 1BY  United Kingdom
T: +44 (0)131 220 6966  F: +44 (0)131 220 6087

**Glasgow**
Seventh Floor, 78 St Vincent Street
Glasgow G2 5UB United Kingdom
T: +44 (0)141 225 4400  F: +44 (0)141 225 4401

**London**
Second Floor, 17 Hanover Square
London W1S 1HU  United Kingdom
T: +44 (0)20 7529 6500  F: +44 (0)20 7529 6556

**Lyon**
11 rue de la Republique, 69001 Lyon, France
T: +33 (0)4 72 10 29 29  F: +33 (0)4 72 10 29 28

**Manchester**
25th Floor, City Tower, Piccadilly Plaza
Manchester M1 4BT  United Kingdom
T: +44 (0)161 236 0282  F: +44 (0)161 236 0095

**Marseille**
13, rue Roux de Brignoles, 13006 Marseille, France
T: +33 (0)4 91 37 35 15  F: +33 (0)4 91 54 18 92

**Paris**
12-14, rue Jules Cesar, 75012 Paris, France
T: +33 (0)1 53 17 36 00  F: +33 (0)1 53 17 36 01

**Email: info@mvaconsultancy.com**

**Offices                    also                    in**
Bangkok, Beijing, Hong Kong, Shenzhen and Singapore

**mvaconsultancy**