Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Transportation Research Part D 10 (2005) 305–322
TRANSPORTATION RESEARCH PART D
www.elsevier.com/locate/trd

Valuing noise level reductions in a residential location context

Patricia Galilea, Juan de Dios Ortúzar *

Department of Transport Engineering, Pontificia Universidad Católica de Chile, Casilla 306, Código 105, Santiago 22, Chile

Abstract

Although noise levels in urban areas frequently surpass the designated norms the consequences on health are only beginning to be examined. A stated preference experiment is used to estimate the willingness-to-pay for reducing noise levels in a group-based residential location context. The experiment considers variations in travel time to work, monthly house rent, sun orientation of the dwelling and subjective noise level inside it; objective noise levels are also measured after the experiment. Multinomial and mixed logit models are estimated based on a consistent microeconomic framework, including non-linear utility functions and allowing for various stratifications of the data. The more flexible models allow for the treatment of repeated observations problem common to stated preference data and provide a better fit to the data, although willingness-to-pay results remain almost invariant. Subjective values of time derived from these models are consistent with previous values obtained in the country, giving support to the experimental design and quality of the sample.

Keywords: Noise valuation; Willingness-to-pay; Stated preference; Discrete choice modelling; Value of time

1. Introduction

As urban transport systems grow externalities often become important. The social evaluation of transport projects usually includes the benefits associated with time savings, however the benefits

* Corresponding author. Tel.: +56 2 686 4822; fax: +56 2 553 0281.
E-mail address: jos@ing.puc.cl (J. de D. Ortúzar).

1361-9209/$ - see front matter
doi:10.1016/j.trd.2005.04.004

of reducing the number and severity of accidents are only considered seriously in first world nations, and the potentially large benefits associated with reductions in pollution and noise are seldom considered even in the developed world.

There are at least two ways to incorporate the social valuation of external effects. The first attempts to quantify the change in the aggregate product value in terms of the impact on social or individual productivity—the social accounting or shadow price approach. The second involves estimating the perception of users about the damage inflicted, either through observations of their actions (revealed preference) or through analysis of their stated willingness-to-pay (WTP), the so-called questionnaire methods.[1]

Here we deal with the problem of valuing reductions in noise levels. Noise is an endemic problem in large metropolis and has received relatively little attention as a potential health hazard in the developing world. For example, measurements carried out by the Environmental Health Service in Chile (SESMA, 1999) have shown that in most areas of its capital, Santiago, the relatively lax Chilean norm of a maximum of 55 dB(A) in residential areas, is violated during many hours of the day.

This is, however, a complex subject due to the diffused nature of the contribution from each agent and the complex causal relation between noise and health. Several methods have been applied to value noise effects in the past, including hedonic pricing (Vainio, 2001) and contingent valuation (Feitelson et al., 1996; Bjorner, 2004)—a comprehensive review is provided by Navrud (2002). Nonetheless, these methods have well-known deficiencies (Hausman, 1993; Azqueta, 1994).

On the other hand, stated preference (SP) methods have been widely used in marketing and in transport research to identify responses to choice situations that are not clearly revealed in the market. As such, they appear to be a promising tool to estimate monetary values for improvements to the environment. Here we report the first experience in using SP data to estimate the WTP for reducing noise levels in Latin America. Few studies of this type have been reported before (Saelensminde, 1999a; Daniels and Hensher, 2000; Wardman and Bristow, 2004).

Household location choice is studied for theoretical reasons. In valuation experiments, it is necessary to induce a direct perception of damages by the individual and this cannot be achieved in a mode or route choice context. To interpret the results of the SP approach, this work is based on the microeconomic residential location model developed by Perez et al. (2003). The group-based approach (i.e. making the whole family participants of the SP game) is generally seen as superior to more traditional individual—based studies (Molin et al., 1999) and has been used to measure WTP for improved accessibility and for lower environmental pollution (Ortúzar et al., 2000; Ortúzar and Rodriguez, 2002).

2. Survey design

As valuation of noise reductions is not an everyday consideration for individuals it is important to be careful in selecting an appropriate choice context for doing this, and the specific way, the *n* variable metric, in which it would be presented in the survey.

[1] Freeman (1993) and more recently Arsenio (2002), examine and classify the various methods that have been proposed in the literature.

2.1. The choice context

To estimate the WTP a setting is needed where a family can be exposed, credibly, to different noise levels. Previous experience with related problems have shown that a realistic context is offering respondents a choice of residential locations associated with different noise levels. When people choose a place to live they consider not only the dwelling characteristics but also the features of its location including noise levels and accessibility conditions (Hunt et al., 1994).

The noise surveillance done by SESMA (1999) shows that noise levels vary in different parts of Santiago suggesting families that have made their location choice recently have implicitly decided on an acceptable level of noise for their home activities. This has to be contrasted with other attributes associated to the selected dwelling, such as rent and accessibility.

Residential location represents a medium/long term decision and it involves a complex decision process. A rank-order is appropriate because it requires ordering options based on attractiveness criteria, instead of choosing a particular one.[2] This allows for in-depth family discussions (Ortúzar and Rodriguez, 2002; Perez et al., 2003). Maintaining context and format also has the advantage of allowing particular results to be compared with those of previous studies; e.g. subjective values of time.

2.2. Defining attributes and measurement units

Focus group and pre-tests were conducted to identify the set of attributes used to define the alternatives and how best to present them. Individuals of different ages, gender and socio-economic conditions were selected who had recently moved into a flat; previous studies had found that people who own, or had been renting a flat for more than a year, found it difficult to put themselves in the context of an hypothetical experiment that requires thinking about changing location.

2.2.1. Identification of the set of attributes

The first attribute normally considered by a family when choosing a flat is its location. In this sense, Santiago is an extremely segregated city; the rich and the poor live in different areas and very seldom mix. Second, come rent/mortgage consideration and then the apartment quality. After some prodding focus group participants recognized that the noise level was important, but also the direction the flat faced, whether it faced the sun or not; orientation preferences are not universal; some people prefer sun in the morning and others in the evening. Based on these results, the attributes chosen were: rent or mortgage paid, noise level, travel time to work and sun orientation. The quality of the apartment is assumed to be the same in all cases.

2.2.2. Selection of measurements units

This task is trivial for time and money attributes because they are familiar to most people. Unfortunately, this is not the case for variables such as pollution, noise level or spatial orientation. The focus group confirmed that although noise levels are measured on the decibel scale (dB), people do not know its meaning nor do they know it is logarithmic. Presenting the noise variable as a function of recalled levels at key intersections in Santiago was tried (e.g. some objectively

[2] Although the use of rank-order data is not favoured by all (Louviere, 2002), it can, however, be a very useful method (Jones and Hensher, 2003).

louder than others). For this, participants were asked to rank using a five-point scale. The results, however, did not show a clear pattern and none of the respondents came close to the objective data in their assessments.

Laboratory experiences were discarded, although they would allow the simulation of a wide range of situations, because the experiments are costly and there is no way to know if respondents feel 'at home' in this simulation or if they would be really annoyed by that noise level in practice (Arsenio et al., 2000). Eventually, a traditional 10-point rating scale was adopted. Finally, in the case of sun orientation cardinal points are used; this is not only the natural way of presenting the attribute but it was easily understood by participants and confirmed at the subsequent pre-tests.

2.3. Experimental design and statistical design

After selecting the attributes and their representation, the next step is to select their numbers of levels. Although more levels allow testing for non-linearities, the number of choice situations increases and so does respondent burden (Ampt, 2003; Caussade et al., 2005).

The factorial design chosen for this survey was a 2^4 experiment (Street et al., 2001). A full factorial needs 16 choice situations but has been found in past work to test respondent's patience. Two blocks with eight treatments each are thus used, confounding the four-way interaction ($A \times B \times C \times D$). Although two-way and three-way interactions can be estimated, quadratic effects cannot be estimated because this represents a 2^k design (Louviere et al., 2000).

Two blocks of eight treatments are generated by solving[3]:

Block 1: $A + B + C + D = 0$ (mod 2)	Block 2: $A + B + C + D = 1$ (mod 2)
0000 (alternative 1)	0001 (alternative 6)
1100 (alternative 8)	0010 (alternative 5)
1001 (alternative 7)	0100 (alternative 2)
1010 (alternative 3)	1000 (alternative 3)
0110 (alternative 5)	1110 (alternative 4)
0101 (alternative 4)	1101 (alternative 7)
0011 (alternative 2)	1011 (alternative 8)
1111 (alternative 6)	0111 (alternative 1)[a]

[a] This alternative was modified, as seen in Table 1, and replaced by (0000) to introduce more realism at the cost of loosing complete orthogonality.

Table 1 details these results. If a respondent answers block 1 seriously, the person should rank alternative 8 first and alternative 2 last. On the other hand, a respondent answering block 2 seriously should put alternative 3 before alternatives 1, 5, 6 and 8. These are just examples; blocks 1 and 2 imply 12 and nine dominated pairs of alternatives.[4] This enables checking of the data for inconsistencies (Saelensminde, 2001, 2002).

[3] Variables with two levels are represented by zero at the lower level and by one at the higher level.

[4] Dominance appears when every attribute of an alternative is in an equal or better level than those of the other; therefore, the former should be placed always higher than the latter in the ranking.

Table 1
Attributes levels in each block

Alternative	Block 1				Block 2			
	Time	Noise	Sun	Rent	Time	Noise	Sun	Rent
1	High	High	Best	Low	High	High	Best	Low
2	High	High	Worst	High	High	Low	Best	Low
3	Low	High	Worst	Low	Low	High	Best	Low
4	High	Low	Best	High	Low	Low	Worst	Low
5	High	Low	Worst	Low	High	High	Worst	Low
6	Low	Low	Worst	High	High	High	Best	High
7	Low	High	Best	High	Low	Low	Best	High
8	Low	Low	Best	Low	Low	High	Worst	High

2.4. Pilot study

A pilot study was conducted for a sample of 12 households with one, two, four, five and eight family members. This allowed additional examination of a seven-point scale based on the Chilean school marking system. This scale was found to confuse respondents; many family members forgot that grade seven (i.e. the best mark in the Chilean school system) represented a 'good' noise level (almost silence) and took it as the highest noise level. For this reason a traditional 10-point scale was finally chosen, where grade one represented a noise level 'as in the countryside' and grade ten an 'unbearable noise'. So each household indicated the noise level grade they thought their dwelling was in; the 'current level' referred to below.

The pilot also helped define variations in attribute levels for the final survey (Table 2). As the sun orientation variable cannot vary in percentage terms each household was asked for its best and worst orientations, and these were used as levels; although this definition worked well in the pre-tests and pilot, it was found to be somewhat extreme in the final analysis.

Finally, the pilot study helped in define minimum levels and threshold variations for some attributes; e.g., a variation of less than 10 min in travel time between two situations emerges as insignificant (in comparison with variations in the other variables). We also established that the minimum noise level should be grade two and that the minimum travel time should be 8 min (Galilea, 2002).

3. Data collection and analysis

3.1. Sample strategy

The data collection strategy involves two stages and a small group of well-trained interviewers making personal visits to each family (Ortúzar et al., 2000; Ortúzar and Rodriguez, 2002).[5] The

[5] Each interviewer was provided with a survey manual containing the precise set of questions to be formulated and exact definitions of the data required; e.g., the manual specified that an adult family member should be the first contacted to ask for the general household information.

Table 2
Variation levels for each attribute

Attributes	Variation over the current level	
	Level 0 (%)	Level 1 (%)
Time travel to work (TTW)	15	−15
Monthly rent (MR)	−10	10
Noise level (NL)	15	−15

first stage involved obtaining details of the main characteristics of the dwelling and basic information about family members. Socio-economic characteristics required were: first name, relation to household head, gender, age, educational level, possession of a driving license, and occupation. The household data were: borough where the dwelling was located, nearest street intersection, monthly rent/mortgage paid, origin of this money, number of household vehicles and family income. The interviewer also gathered trip data from every worker in the family: modes used, borough where the work place was located, nearest street intersection to the work-place, current travel time and weekly number of trips to work. Finally, the family was asked about which sun orientations were considered best and worst, and requested them to grade the current level of noise inside the dwelling according to the 10-point scale.

The second stage involved a second visit two days later where a customized SP exercise was presented. The experiment was generated on the basis of the data collected at the first stage and included a set of complementary questions. The interviewer made an introductory description of the SP context and varying attributes and then delivered 8 cards representing different residential locations (Fig. 1). The family was asked to rank these cards and after completing the process the interviewer asked questions designed to see if the family members had played the game consistently, how important they considered the variable noise, and if the attribute levels had been considered realistic (Galilea, 2002).

Because the main purpose of the exercise was to value reductions in noise it was felt important to measure the noise level of each dwelling. Due to the high cost of this it was decided to interview families living only in predetermined buildings. An important step was to get permission to do the survey in each of these buildings; a registry of noise levels was offered in return.

Finally 150 flats were surveyed in nine buildings in different areas of Santiago (four in the high income district, two in a medium-high income sector and three on the boundary between middle and low income sectors). The buildings were selected on the basis of their socio-economic characteristics and noise levels. The number of households interviewed at each building varied between 3 and 4 in high-income buildings to between 27 and 33 in low-medium income buildings.[6]

The socio-economic information of the sample is representative of middle and high-income people in Santiago. Table 3 shows the distribution by rent paid and family income; 63.5% of households answering the income question stated an income level over 850,000 Ch$/month. This

[6] The sampling strategy was to get as many households who have moved in the last year as possible in each building. No record of detailed response rates was kept, although of 12 preliminary selected buildings, we were given permission to conduct our survey in 9.

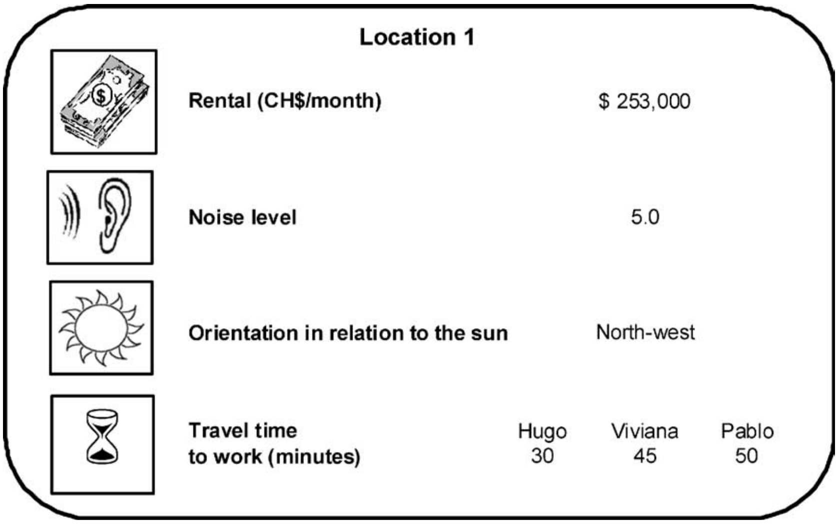

Fig. 1. Example of personalized ranking card.

Table 3
Rent/mortgage distribution by family income

Rent/mortgage (10^3 Ch$/month)	Family income (10^3 Ch$/month)						Total
	100–350	351–550	551–850	851–1250	Over 1250	No answer	
0–99	3	2	1	1	1	1	9
100–199	5	15	24	21	11	4	80
200–299	2	2	4	9	14	3	34
300–399	–	1	1	4	11	4	21
Over 400	–	–	–	–	5	1	6
Total	10	20	30	35	42	13	150

is among the highest 10% of income in the country—the minimum wage being a little over 100,000 Ch$/month (at the time $1 = Ch$650).

The work on measuring noise levels started as soon as the second visit was over for all households. Attempts were made to measure inside flats to spot differences resulting from street-level height or orientation, however, only 64% of flats were measured—the rest either refused entrance or nobody was in when the noise specialist arrived.

3.2. Consistency and lexicographical behaviour

There is a need to detect observations that are not internally consistent or that do not correspond to the assumed population behaviour. Those households that did play the SP game as carefully as required are first identified. A second task was to identify those cases where responses suggested that the household decision strategy could be inconsistent with the compensatory decision making protocol assumed.

Comparing eight alternatives with four attributes each is not easy and it is to be expected that some households would make mistakes and be inconsistent. A maximum of two inconsistent responses per household is seen a as a reasonable indicator of a carefully completed exercise. This led to the elimination of 65 individual responses and all observations from 18 households with three or more inconsistencies.

Forty households exhibiting potentially lexicographical behaviour (i.e. ranking on the basis of a single attribute) were detected; 16 on the attribute rent, 23 on sun orientation, and 1 on travel time to work.[7] This is a slightly lower proportion than reported in previous studies (Saelensminde, 1999b; Ortúzar and Rodriguez, 2002; Rizzi and Ortúzar, 2003; Iraguen and Ortúzar, 2004). The relatively high presence of 'lexicographical households' with respect to sun orientation may be due either to a real family concern for this attribute or, and more probably, to the use of the best and worst orientations as levels for this attribute which offered only extremes.

The potentially lexicographical responses were kept in the final estimation process for various reasons. First, there is never a certainty that respondents are truly lexicographic (see the discussion by Hojman et al., 2004). Second, their inclusion allows to compare SP results with those of revealed preference (RP) surveys, where lexicographical answers cannot be detected. Finally, in this case the models including lexicographical answers fit the data as well as those excluding them.

4. Discrete choice modelling

As is customary, we first searched for the best Multinomial Logit (MNL) model specification and then we relaxed the assumptions of fixed coefficients and independence of observations by the same household required by the model, and estimated more flexible Mixed Logit (ML) models (Train, 2003).[8] We proceeded from simple specifications to more interesting structures allowing for interaction effects and parameterised main effects.

The definition of the variables used at the modelling stage are: NL_i is the noise level at location i (1–10, where ten is an unbearable level of noise); RM_i is value of the flat rent or mortgage (thousand of Ch$ per month); SUN_i is a dummy variable that takes the value 1 if option i has the best orientation in relation to the sun, as declared by household h, and 0 if it has the worst; TTW_i is travel time to work by all family workers (minutes per week)-this household accessibility variable is defined following Perez et al. (2003):

$$\mathrm{TTW}_i = \sum_{h \in H_i} f_{ih} \mathrm{TTW}_{hi}, \tag{1}$$

where TTW_{hi} is the travel time to work by individual h from location i (minutes per trip) and f_{ih} is the frequency of trips to work by individual h, from location i (trips per week).

[7] We label these potentially, because it could be that although they actually choose in a compensatory fashion, the levels of the experiment do not imply a real trade-off for them.

[8] Ortúzar and Willumsen (2001) provide summary of the economic and statistical background for analysing discrete choice data.

4.1. Linear and stratified MNL models

Initially models were based on a linear-in-parameters indirect utility function:

$$V_i = \theta_{\mathrm{RM}}\mathrm{RM}_i + \theta_{\mathrm{NL}}\mathrm{NL}_i + \theta_{\mathrm{SUN}}\mathrm{SUN}_i + \theta_{\mathrm{TTW}}\mathrm{TTW}_i. \tag{2}$$

Two MNL models were estimated; the first included all consistent responses by every household, and the second excluded households exhibiting potential lexicographical behaviour. The maximum likelihood results show that both models have a satisfactory adjustment in comparison to the market shares model, correct signs and significant parameters (Galilea, 2002). This differs from previous experience when models excluding potentially lexicographical respondents had been found superior (Rizzi and Ortúzar, 2003; Iraguen and Ortúzar, 2004). For this reason models for the complete sample are considered.

Segmentation is tested to quantify the influence of certain household characteristics in the valuation of attributes, but only flat ownership was significant. To identify the best specification at each stratum, a step-by-step methodology proposed by Ortúzar and Rodriguez (2002) is used:

1. Estimate a general model with different parameters in both segments; identify statistical similarity between parameters considering magnitudes and t-tests.
2. Estimate a restricted model under the null hypothesis that the most similar pair of parameters can be replaced by a single parameter in the estimation process.
3. Perform a likelihood-ratio test to ensure that both models (general and restricted) are statistically equivalent and that, consequently, there is a gain in parsimony.
4. Recheck parameter similarity, identify the closest pair of specific parameters remaining and go back to step 3.
5. Stop when no potential pair similarities or model reductions are found.

Owners and tenants are considered. Although their numbers were not equivalent (43 and 107, respectively), the results in Table 4 are solid. A likelihood ratio-test shows that the stratified structure MNL-2 is significantly superior to MNL-1, the original model without stratification.

WTP values and their 95% confidence intervals are calculated following Armstrong et al. (2000) (Table 5). The subjective values of time (SVT) are in complete agreement with values estimated in previous studies (Ortúzar et al., 2000; Perez et al., 2003) and this gives credibility to the experiment, i.e. not only is it well designed and understood, but also respondents answered with seriously. Secondly, the subjective value of reducing noise (SVN) has the same order of magnitude as the money needed to double-glaze in dwellings; these are values based on people's perceptions of noise and not objective values.

On the other hand, model MNL-2 implies that flat owners value sun orientation and travel time higher than tenants, but their perception of noise and rent values are not significantly different; consistent with the findings of Bjorner (2004). The first result can be explained by considering that when a person is buying an apartment there is a longer-term commitment than when choosing one to rent, so sun orientation in a warm country should be more important. But the much larger parameter associated to owners' travel times to work is probably due to their higher income level. Finally, although the confidence intervals for the SVT overlap, the interval for the tenants does not include the owners' point estimate. Their confidence interval being wider could be explained

Table 4
Model stratified by dwelling's ownership

Attributes		Parameters (t-test)	
		MNL-1	MNL-2
RM (10^3 Ch$/month)		−0.043 (−16.3)	−0.0436 (−16.4)
NL		−0.660 (−14.3)	−0.665 (−14.3)
SUN	Tenants	1.697 (16.2)	1.566 (13.5)
	Owners		2.142 (10.4)
TTW (min/week)	Tenants	−0.006 (−7.3)	−0.006 (−6.1)
	Owners		−0.009 (−4.5)
$l(\theta)$		−967.03	−963.28
ρ_c^2		0.213	0.216
Sample size		859	859

Table 5
Subjective values by dwelling's ownership

Attributes		Subjective values (95% confidence interval)	
		MNL-1	MNL-2
Noise level (US$/NL per month)		23.54 (20.26–27.24)	23.45 (20.17–27.09)
TTW (US$/h)	Tenants	3.10 (2.30–3.92)	2.81 (1.93–3.70)
	Owners		4.43 (2.53–6.37)

by the fact that, although most owners had a higher income level than the tenants, this was not true for every owner.

4.2. MNL models including interaction effects[9]

The factorial design allows estimation of models involving interactions between variables. The specification searches began with a general model including all main effects, two-way and three-way interactions and then less significant effects were taken out, one by one, until a model was reached with significant variables and intuitively correct signs. The results were checked by choosing different ways to take out the less significant effects; this model had a significantly better fit than MNL-1 (Galilea, 2002).

[9] All models with interactions and with parameterised main effects were estimated deviating each variable by its mean. This helps estimating the confidence intervals and does not change the results. The variable means were calculated based on the attributes effectively available for each observation at the estimation process.

On the other hand, Eq. (2) does not allow incorporation of individual tastes in a MNL. One way of doing this is to parameterise the coefficients of the main effect variables by means of the individuals' socio-economic and journey characteristics:

$$V_{ij} = \left(\alpha_0 + \sum_l \alpha_l s_{lj}\right)\mathrm{RM}_{ij} + \left(\beta_0 + \sum_l \beta_l s_{lj}\right)\mathrm{NL}_{ij} + \left(\gamma_0 + \sum_l \lambda_l s_{lj}\right)\mathrm{SUN}_{ij} + \left(\delta_0 + \sum_l \delta_l s_{lj}\right)\mathrm{TTW}_{ij}. \quad (3)$$

In contrast with the traditional specification of socio-economic variables, (3) applies to both alternatives; also, since the same additional variable can be related to more than one attribute, it can be specified with different coefficients in each case. Therefore, as every individual has different socio-economic and journey characteristics, each may end up with different valuations for the same attributes. Rizzi and Ortúzar (2003) provide a microeconomic rationale for Eq. (3).

The binary variable s_{lj} represents socio-economic (SE) feature l of individual j. This is an interesting way of incorporating additional variables and allow us to use additional individual data to estimate WTP. As there may be different coefficients for each attribute depending on the special characteristics of each household, this specification allows estimating models that are almost unique to each household helping to reduce the problem of taste variations. After a specification search the most significant variables were (Galilea, 2002):

- *No. people/Income*$_{\mathrm{RM}}$: Number of household members divided by family income level; as it was added to the rent coefficient its value should be negative, if the number of members increases or if income decreases (*ceteris paribus*), an increase in rent should affect them more.
- *Floor*$^2_{\mathrm{NL}}$: Takes the quadratic value of the floor where the apartment is located. Its value should be negative, because the noise level is higher in the ground floor (closer to the source of noise) and in the upper floors (no shield from other houses or smaller buildings), so people living in these floors should be more sensitive to higher noise levels.
- *Owner*$_{\mathrm{SUN}}$: Takes the value of one if the household owns the flat. Its value should be positive because owners attach more importance to the sun orientation than the rest.

The best model (MNL-4) incorporating both interactions and parameterised main effects is shown in Table 6. To compare it with the linear MNL we had to estimate a new model (MNL-3) using only data from families reporting their income and who passed the consistency test (119 households). As can be seen, all parameters in both models have correct signs but a likelihood-ratio test rejects comfortably the null hypothesis that both models are equivalent (i.e. $\mathrm{LR} = 33.7 > \chi^2_{6;95\%} = 12.6$).

The results indicate that, for example, a household with a given income level will increase its valuation of the rent in 0.031 (i.e. almost 85%) for each new member added to it. On the other hand, flat owners value sun orientation 36% higher (i.e. 1.674 plus 0.609 over 1.674) than do tenants.

It is surprising to find more three-way than two-way effects present in the preferred specification, as it has been generally assumed that the latter explain a greater proportion of the data

Table 6
Model with interactions and parameterised main effects

Attributes	Parameters (t-test)	
	MNL-3	MNL-4
RM (10^3 Ch$)	−0.0447 (−15.9)	−0.0360 (−6.8)
No people/Income$_{RM}$	–	−0.0307 (−2.6)
NL	−0.6789 (−13.8)	−0.6470 (−10.3)
Floor$^2_{NL}$	–	−0.0026 (−2.2)
SUN	1.743 (15.7)	1.6742 (13.5)
Owner$_{SUN}$	–	0.6086 (2.6)
TTW (min)	−0.00621 (−6.9)	−0.00668 (−7.1)
RM × NL	–	−0.00152 (−2.7)
RM × SUN × TTW	–	−0.000025 (−2.8)
NL × SUN × TTW	–	−0.000481 (−2.0)
$l(\theta)$	−877.63	−860.78
ρ^2_c	0.222	0.236
Sample size	786	786

variation (Louviere et al., 2000); but models with only two-way interactions were consistently inferior (Galilea, 2002).

Table 7 presents estimates for the subjective values derived from MNL-4. Although the specification is much improved, the results remain almost invariant. These values are computed taking account of interactions and parameterised effects, so we need to derive utility with respect to the attributes as

$$\frac{\partial U}{\partial \mathrm{RM}} = \theta_{\mathrm{RM}} + \theta_{\mathrm{No.\ people/Income}} \times \frac{\mathrm{No.\ people}}{\mathrm{Income}} + \theta_{\mathrm{RM}\times\mathrm{NL}} \times \mathrm{NL} + \theta_{\mathrm{RM}\times\mathrm{TTW}\times\mathrm{SUN}} \times \mathrm{TTW} \times \mathrm{SUN}, \tag{4}$$

$$\frac{\partial U}{\partial \mathrm{NL}} = \theta_{\mathrm{NL}} + \theta_{\mathrm{Floor}^2} \times \mathrm{Floor}^2 + \theta_{\mathrm{RM}\times\mathrm{NL}} \times \mathrm{RM} + \theta_{\mathrm{RM}\times\mathrm{SUN}\times\mathrm{TTW}} \times \mathrm{RM} \times \mathrm{SUN} + \theta_{\mathrm{NL}\times\mathrm{SUN}\times\mathrm{TTW}} \times \mathrm{SUN} \times \mathrm{TTW}, \tag{5}$$

$$\frac{\partial U}{\partial \mathrm{TTW}} = \theta_{\mathrm{TTW}} + \theta_{\mathrm{NL}\times\mathrm{SUN}\times\mathrm{TTW}} \times \mathrm{NL} \times \mathrm{SUN} + \theta_{\mathrm{RM}\times\mathrm{SUN}\times\mathrm{TTW}} \times \mathrm{RM} \times \mathrm{SUN}. \tag{6}$$

Then, the subjective values of time and noise level are calculated as follows[10]:

$$\mathrm{SVT} = \frac{\partial U/\partial \mathrm{TTW}}{\partial U/\partial \mathrm{RM}} \times \frac{12\,000}{52} \times \frac{60}{650} \quad [\mathrm{US\$/h}], \tag{7}$$

$$\mathrm{SVN} = \frac{\partial U/\partial \mathrm{NL}}{\partial U/\partial \mathrm{RM}} \times \frac{1000}{650} \quad [\mathrm{US\$/\deg}]. \tag{8}$$

[10] In Eqs. (4)–(6) we used sample averages for the variables Income, Sun, Floor, Rent (RM) and travel time to work (TTW).

Table 7
Subjective values for models with interactions

Attributes	Subjective value	
	MNL-3	MNL-4
Noise level (US$/NL per month)	23.37 (20.11–27.04)	23.68 (20.43–39.25)
Travel time to work (US$/h)	2.96 (2.18–3.74)	3.00 (2.68–5.85)

The SVT and SVN values were positive for all individual households, and this is not always the case in models of this type (Brownstone, 2001; Daniels and Hensher, 2000).

4.3. Estimation of Mixed Logit models

ML models were finally estimated to examine the importance of allowing heterogeneity in individuals tastes explicitly, as well as a correct treatment of the repeated observations problem associated to SP data. In the ML model, apart from the random errors (i.e. white noise) that distribute independent and identically (iid) Gumbel as in the MNL, the systematic part of the utility function, V_{iq}, may also have randomly distributed parameters. This flexible specification allows to consider heteroscedasticity, correlation and variations in tastes but it is hard to estimate (Train, 2003).

To estimate the ML we used a non-commercial code implemented in GAUSS (it can be downloaded from the web page of Kenneth Train: http://elsa.berkeley.edu/~train). We considered independent Normal distributions for the attributes, as previous experience with multivariate functions had shown results to vary little but at non-negligible cost (Sillano and Ortúzar, 2005). For the simulated maximum likelihood search we used sequences of 125 Halton numbers as is usual practice (Train, 2003). A battery of ML models, associated to several of the MNL models estimated during the research is discussed by Galilea (2002); in every case the ML gave a substantially better fit to the data than the corresponding MNL. In what follows we just present two such models, one associated to MNL-1 and another to MNL-4.[11] Comparing ML-1 in Table 8 and MNL-1 in Table 4, it can be seen that the mean values of the ML parameters correctly increase in size due to the scale factor effect; as we are allowing for random parameters the white noise variance (inversely related to the scale factor) decreases significantly (Sillano and Ortúzar, 2005). Note that this also happens with ML-4 and MNL-4.

The results in Table 8 suggest that ML-4 is superior to MNL-4 (although it is not strictly comparable). Note that although most interactions finally received fixed parameters the main effects remained consistently variable among individuals. The table also shows that ML-4 is closer but still superior than ML-1.

Table 9 presents the willingness-to-pay point estimates for these final models. As can be seen, those for ML-4 are larger than those of ML-1 and MNL-4, but still the confidence intervals for each model contain all the point estimates of the competing functions; finally, again no households were found to have individual subjective values with an incorrect sign. As this is seldom

[11] Note that we could not estimate a close enough function as an ML with both interactions and parameterised main effects did not converge.

Table 8
Mixed Logit model including interactions

Attributes	Parameters (t-test)				
	MNL-4	ML-1		ML-4	
		Mean	Standard deviation	Mean	Standard deviation
RM	−0.0360 (−6.8)	−0.1398 (−5.5)	0.0863 (4.7)	−0.120 (−8.6)	0.0888 (6.9)
No people/Income$_{RM}$	−0.0307 (−2.6)	–	–	–	–
NL	−0.6470 (−10.3)	−2.2852 (−10.3)	1.1504 (8.0)	−2.570 (−8.6)	1.283 (7.8)
Floor$^2_{NL}$	−0.0026 (−2.2)	–	–	–	–
SUN	1.6742 (13.5)	3.718 (9.4)	4.699 (7.9)	4.566 (7.5)	5.052 (7.0)
Owner$_{SUN}$	0.6086 (2.6)	–	–	–	–
TTW	−0.00668 (−7.1)	−0.0241 (−7.4)	0.0137 (5.4)	−0.0284 (−8.4)	0.0097 (5.7)
RM × NL	−0.00152 (−2.7)	–	–	−0.00456 (−2.6)	0.00662 (3.8)
RM × TTW	–	–	–	−0.00007 (−3.5)	0.00013 (6.5)
NL × SUN × TTW	−0.000481 (−2.0)	–	–	–	–
RM × SUN × TTW	−0.000025 (−2.8)	–	–	–	–
$l(\theta)$	−860.78	−787.74		−777.62	
ρ^2_c	0.236	0.359		0.367	

Table 9
Subjective values for Mixed Logit model including interactions

Attributes	Subjective values		
	MNL-4	ML-1	ML-4
Noise level (US\$/NL per month)	23.68 (20.43–39.25)	25.15 (18.56–37.63)	33.98 (23.74–45.96)
Travel time to work (US\$/h)	3.00 (2.68–5.85)	3.67 (2.55–5.63)	5.18 (3.62–7.06)

the case, calls are made to include among the model assessment check-list, if the microeconomic conditions implied by the postulated indirect utility function are violated (Cherchi and Ortúzar, 2003); also the use of special distribution functions to avoid the possibility of incorrect signs has been advocated (Hensher, 2005; Hess et al., 2005).

4.4. Subjective versus objective (dB) perceptions of noise level

The final objective is estimation of an objective monetary value for noise levels reductions. To do this, the 10-point scale subjective values are related to the decibel scale measurements taken at the 96 dwellings where this is possible. One problem is that the dB(A) measures are generally fairly high whilst the ranges was not wide (i.e. from 37 to just under 61 dB) implying that many respondents with a low objective noise level report a high grade as their subjective level. So a simple linear regression did not achieve a reasonable fit even when separate regressions were estimated for each building. Only one building, in fact, gave a more satisfactory fit but its respondents were too few for further analyses (Galilea, 2002).

To improve the estimation the extra information provided by the households at the interview stage is incorporated to achieve *ceteris paribus* conditions. In particular, the results of two questions are used: whether they were aware that their dwelling had a significant noise level and, if they thought that the level of noise was an important attribute when searching for a place to live. Thus, a multiple regression was estimated using the 10-point scale subjective grades as the dependent variable, and the decibel scale plus two dummies representing *Awareness* (one, if the household was aware that the dwelling had a significant noise level) and *Importance* (one, if the household thought that noise level was an important attribute) as independent variables. The results of the regression are shown in Table 10; they appear quite reasonable. With these results, we are able to transform the estimated parameter for the noise level, by multiplying it by the coefficient for dB(A).

The transformed noise level parameters allowed us to derive new subjective values of noise level (SVN) for each model (Table 11). These represent the willingness-to-pay, in US$, associated to decreasing the noise level inside a dwelling in 1 dB(A) per month. Although the values appear reasonable and consistent, a caveat related to their use in social project evaluation is that there are other terms and elements that should form part of the total WTP for reducing noise level; for example, the health costs incurred as an effect on human health because of noise.

The importance of establishing a relationship between noise perception and dB(A) is clear for cost-benefit analysis. For instance, if noise levels in a central business district are high and as a consequence most households and offices start incorporating double glazing as a personal protection it could be more efficient to coordinate a central double glazing programme to attack the noise problem, for example, but this should be evaluated. Such a programme could even attract some other agents which, otherwise, could not afford the cost of double glazing.

Table 10
Multiple regression results for the 10-point scale

Attributes	Parameter	(*t*-test)
dB(A)	0.0893	(6.1)
Awareness	2.1295	(3.6)
Importance	1.3184	(2.3)
Multiple correlation coefficient	0.5124	
Sample size	96	

Table 11
Subjective values for noise level reduction in dB

Model	SVN (US$/dB(A) per month)		
	Lower bound	Mean	Upper bound
MNL-1	1.81	2.10	2.43
MNL-2	1.80	2.09	2.42
MNL-3	1.80	2.09	2.42
MNL-4	1.82	2.12	3.51
ML-1	1.66	2.25	3.36
ML-4	2.12	3.03	4.10

As a caveat, we must recall that the relationship established above has to be considered tentative in so far as we did not conduct any kind of external validity of our results. Therefore, it is debatable whether it could be adopted immediately by the environmental authority. More work should follow in this direction.

5. Conclusions

The application of SP techniques to a complex problem as valuing noise level reductions in Chile shows great promise for the use of this methodology in other countries. Two results emerged from the design stage of the survey: the identification of residential location as an appropriate experimental framework, and the formulation of a variable metric for noise level (although only related to family perceptions) that was clearly understood by the participants in the exercise.

We also found that the statistical design was able to represent the respondents' preferences for the variables included in the experiment. This is supported not only by the good general fit of the estimated models, but also by the anticipated parameter signs and reasonable significance *t*-tests. Equally, the subjective values of time obtained from the various models estimated turned out to be consistent with prior studies. This is, in our opinion, a clear indication that respondents understood the experiment which included two new variables in relation to previous experiences: noise level and sun orientation. So in spite of its complexity, the SP experiment was able to capture individual preferences adequately.

In terms of results, individual households do not necessarily have linear utility functions. Not only are several interaction terms significant but the introduction of additional variables (socio-economic and related to noise level) affect the coefficients of the main-effects variables. These latter variables indicate the significance of age and income in the valuation of travel time. Further, the flexible and powerful Mixed Logit model easily outperformed the simple MNL; this may be because it accepts the presence of random taste variations among individuals (which indeed appeared as an effect) or just because it allows consistent treatment of the problem of repeated observations by each individual, which is a feature of stated preference and panel data.

In relation to the estimated values for reducing noise levels, and given the caveats, a conservative value of US$2.12 per decibel per month emerges as corresponding to the lower bound of the confidence interval associated to the best ML model (and also to the point estimate of the various MNL functions). This value (and most values estimated) appears to be reasonable when compared (although the comparison is *per force* not strict) with the real costs[12] associated to reducing noise by physical means (i.e. double glazing).

Acknowledgements

The final version of this paper was written during a six-week stay of the second author at the Institute for Transport and Logistics Studies, University of Sydney. We wish to thank Ana María

[12] The cost of double glazing is roughly US$130 per square meter. If, for example, the noise level was near 70 dB(A) double glazing would reduce it to approximately 40 dB(A). However, there would also be a reduction of approximately 40% in central heating costs.

Ivelic for help at the survey design stage, Luis Ignacio Rizzi for his invaluable help throughout the work and Mauricio Sillano for his pointers in relation to Mixed Logit estimation. Thanks are also due to Luis Cifuentes, David Hensher, Sergio Jara-Diaz and Huw Williams for their advice. Ken Button and two anonymous referees were also instrumental in shaping the paper. The research was supported by the National Fund for Scientific and Technological Development (FONDECYT) through Project 1000616. We are also grateful for the dedicated work of our interviewers and to those responsible for the noise level measurement campaign.

References

Ampt, E.S., 2003. Respondent burden. In: Stopher, P.R., Jones, P.M. (Eds.), Transport Survey Quality and Innovation. Pergamon, Amsterdam.

Armstrong, P.M., Garrido, R.A., Ortúzar, J. de D., 2000. Confidence intervals to bound the value of time. Transportation Research E 37, 143–161.

Arsenio, E., 2002. The Valuation of Environmental Externalities: A Stated Preference Case Study on Traffic Noise in Lisbon. Ph.D. Thesis, Institute for Transport Studies, University of Leeds.

Arsenio, E., Bristow, A.L., Wardman, M., 2000. An innovative stated preference computer survey model for valuing noise impacts from road traffic. In: Proceedings of Seminar F, European Transport Conference, Cambridge.

Azqueta, D., 1994. Valoración Económica de la Calidad Ambiental. McGraw-Hill, Madrid (in Spanish).

Bjorner, T.B., 2004. Combining socio-acoustic and contingent valuation surveys to value noise reduction. Transportation Research D 9, 341–356.

Brownstone, D., 2001. Discrete choice modelling for transportation. In: Hensher, D.A. (Ed.), Travel Behaviour Research: the Leading Edge. Pergamon, Oxford.

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

Cherchi, E., Ortúzar, J. de D., 2003. Alternative specific variables in non-linear utilities: influence of correlation, homoscedasticity and taste variations. In: Proceedings 10th International Conference on Travel Behaviour Research, Lucerne.

Daniels, R., Hensher, D.A., 2000. Valuation of environmental impacts of transport projects. Journal of Transport Economics and Policy 34, 189–214.

Feitelson, E., Hurd, R., Mudge, R., 1996. The impact of airport noise on willingness to pay for residence. Transportation Research D 1, 1–14.

Freeman, A.M., 1993. The Measurement of Environmental and Resource Values: Theory and Methods. Resources for the Future, Washington, DC.

Galilea, P., 2002. Preferencias Declaradas en la Valoración del Nivel de Ruido en un Contexto de Elección Residencial. M.Sc. Thesis, Department of Transport Engineering, Pontificia Universidad Católica de Chile, Santiago (in Spanish).

Hausman, J.A. (Ed.), 1993. Contingent Valuation: a Critical Assessment. North-Holland, Amsterdam.

Hensher, D.A., 2005. Imposing a globally signed condition on a generalised parameterisation of attributes in choice models to produce sign compliant willingness to pay distributions (mimeo).

Hess, S., Bierlaire, M., Polak, J.W., 2005. Estimation of value of travel-time savings using mixed logit models. Transportation Research A 39, 221–236.

Hojman, P., Ortúzar, J. de D., Rizzi, L.I., 2004. On the joint valuation of averting fatal victims and serious injuries in highway accidents. Journal of Safety Research, under review.

Hunt, J.D., McMillan, J.D.P., Abraham, J.E., 1994. Stated preference investigation of influences on attractiveness of residential locations. Transportation Research Record 1466, 17–35.

Iraguen, P., Ortúzar, J. de D., 2004. Willingness-to-pay for reducing accident risk in urban areas: an internet-based web page stated preference survey. Accident Analysis and Prevention 36, 513–524.

Jones, S., Hensher, D.A., 2003. Predicting financial distress using reported cash flows: an ordered mixed (random parameter) logit model. Working Paper, Faculty of Economics and Business, University of Sydney.

Louviere, J.J., 2002. Private communication.
Louviere, J.J., Hensher, D.A., Swait, J.D., 2000. Stated Choice Methods: Analysis and Application. Cambridge University Press, Cambridge.
Molin, E., Oppewal, H., Timmermans, H., 1999. Group-based versus individual-based conjoint preference models of residential location: a comparative test. Environment and Planning A 31, 1935–1947.
Navrud, S., 2002. The state of the art on economic evaluation of noise. Final Report to European Commission, Department of Economics and Social Sciences, Agriculture University of Norway. Available from: <http://europa.eu.int/comm/environment/noise/020414noisereport.pdf>.
Ortúzar, J. de D., Rodriguez, G., 2002. Valuing reductions in environmental pollution in a residential location context. Transportation Research D 7, 407–427.
Ortúzar, J. de D., Willumsen, L.G., 2001. Modelling Transport, third ed. John Wiley and Sons, Chichester.
Ortúzar, J. de D., Martinez, F.J., Varela, F.J., 2000. Stated preference in modelling accessibility. International Planning Studies 5, 65–85.
Perez, P., Martinez, F.J., Ortúzar, J. de D., 2003. Microeconomic formulation and estimation of a residential location choice model: implications for the value of time. Journal of Regional Science 42, 771–789.
Rizzi, L.I., Ortúzar, J. de D., 2003. Stated preference in the valuation of interurban road safety. Accident Analysis and Prevention 35, 9–22.
Saelensminde, 1999a. Stated choice valuation of urban traffic pollution. Transportation Research D 4, 13–27.
Saelensminde, 1999b. Valuation of non market goods for use in cost-benefit analyses: methodological issues. Ph.D. Thesis, Department of Economics and Social Sciences, Agricultural University of Norway.
Saelensminde, K., 2001. Inconsistent choices in stated choice data, use of the logit scaling approach to handle resulting variance increases. Transportation 28, 269–296.
Saelensminde, K., 2002. The impact of choice inconsistencies in stated choice data. Environmental and Resource Economics 23, 403–420.
SESMA, 1999. Estudio Base de Generación de Niveles de Ruido del Gran Santiago. Sub-Departamento de Entorno Saludable. Programa de Salud Metropolitano del Ambiente, Santiago (in Spanish).
Sillano, M., Ortúzar, J. de D., 2005. Willingness-to-pay estimation with mixed logit models: some new evidence. Environment and Planning A 37, 525–550.
Street, D.J., Bunch, D.S., Moore, B., 2001. Optimal designs for 2^k paired comparison experiments. Communications in Statistics—Theory and Methods 30, 2149–2171.
Train, K.E., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge.
Vainio, M., 2001. Comparison of hedonic prices and contingent valuation methods in an urban traffic noise context. 2001 International Congress and Exhibition on Noise Control Engineering, The Hague.
Wardman, M., Bristow, A.L., 2004. Traffic related noise and air quality valuations: evidence from stated preference residential choice models. Transportation Research D 9, 1–27.