This is a post-peer-review, pre-copyedit version of an article published in Environmental Modeling & Assessment. The final authenticated version is available online at: https://doi.org/10.1007/s10666-021-09773-7

Valuations of Transport Nuisances and Cognitive Biases: A Survey Laboratory Experiment in the Pyrenees Region

Laurent Denant-Boemont[a], Javier Faulin[1b], Sabrina Hammiche[a], and Adrian Serrano-Hernandez[b]

[a] *University of Rennes 1 and CREM CNRS, 7 place Hoche, 35000 Rennes, France*

[b] *Public University of Navarre, Institute of Smart Cities, Department of Statistics, Computer Science, and Mathematics, Campus Arrosadia, 31006 Pamplona, Spain*

Abstract

We designed a survey that aims at estimating individual willingness-to-pay to reduce noise and air pollution arising from transportation activity near the Pyrenees in Navarre (Spain). Our participants cope with a series of contingent valuation questions and also with an economic experiment with real incentives about the same topic. Our goal is to identify several methodological problems in the valuation process coming from hypothetical bias, correlation effect and sequence effect when series of responses are requested. Our main results are that hypothetical bias is significant, because the willingness-to-pay is greater when the survey is hypothetical compared to when there is real monetary incentive. Likewise, the correlation effect also observes the same behaviour since the willingness-to-pay for pollution mitigation is close to the one established for noise reduction. Finally, we have obtained mixed evidence for the sequence effect, being present only in the contingent valuation survey part.

Keywords: Transportation; Willingness-To-Pay; Transport Externality; Pollution; Cognitive Bias; Laboratory Economic Experiment

[1] Corresponding author

1. Introduction

Transportation activity is a key condition of economic growth being more and more valuable for households and firms (Eddington, 2006). However, this activity has also long been recognized to generate massive and negative impacts, being the most frequently cited ones: traffic congestion, air pollution, Greenhouse Gas (GHG) emissions, traffic noise, and traffic fatalities/injuries, among others (see, e.g., for Europe EC 2019). These negative impacts have economic and social costs that need to be considered and compared to economic and social benefits of transportation activity in order to implement adequate regulations and public policies. As a consequence of this situation, any significant improvement of transport facilities implies the performance of a cost-benefit analysis associated to that change. Thus, the most widely used methodology at the international level is the cost-benefit analysis (OECD, 2018). A key component of this methodology lies in the possibility of defining shadow prices for non-market effects (Dreze and Stern, 1987) in order to quantify the net change in economic welfare arising from a public policy (Vickerman, 2007; Braun et al., 2016). These shadow prices might be measured thanks to willingness-to-pay (hereafter, WTP) stated by respondents, as for instance in the contingent valuation method frequently used for environmental impacts (Hanemann, 1994; Dupont, 2003).

Our study aims at measuring WTP in a very specific scenario, such as the Autonomous Community of Navarre (Northern Spain); which is a qualified witness of the heavy traffic, mainly freight trucks, crossing the Pyrenees border bound to Central Europe crossing France. The Navarrese region, our study scenario, had been identified to have a very good air quality (Aldabe et al., 2011, Aldabe et al., 2012, Rivas et al., 2019), one of the highest in Spain, being, moreover, a leader in health care system standards, along with Madrid and Basque Country. This is partly explained by the lack of big factories or facilities with pollutant emissions in the region, and consequently the damages to air quality mainly come from massive freight transportation that are due to the geographical location of the region next to the Pyrenees. It makes Navarre an excellent candidate to be tested for individual attitudes towards freight transportation emissions.

Contingent valuation techniques (Denant-Boemont and Hammiche, 2019), firstly proposed by Ciriacy-Wantrup in 1947 as a way to elicit an estimated market valuation for a non-market merchandise (Diamond and Hausman, 1994) are a useful tool to measure subjective well-being in a region. In Lera-Lopez et al.'s (2013, 2014), these techniques provided meaningful results of the Willingness-To-Pay (WTP) related to avoiding noise and air pollution due to transportation in the Navarrese region. These contributions are complementary to the ones done by Istamto et al. (2014a,b) who did a broader study involving many countries and regions in Europe. However, being contingent valuation methodology very popular, it also led to important debates and controversies (Whitehead and Blomquist 2006).

Among others, the respondents in a contingent valuation survey might be affected by different cognitive biases that, finally, could have raised some concerns regarding the capability of the method to produce valid and reliable estimates of WTP values (OECD, 2018). Empirical evidence had been provided about some of these biases: a) *Overstatements:* the hypothetical WTP values could be overstated in comparison to real WTP under experimental settings (Cummings et al 1995, Cummings et al 1997); *Correlation effect among the studied variables* (Kahneman & Ritov 1994): Stated WTPs for a particular public good are an indicator of intensity of a general attitude towards public goods b) *Sequencing effects:* WTP estimations for a particular program in a sequential valuation may depend on its position in the survey (Randall and Hoehn 1996). Hence, the previous potential weaknesses of the contingent valuation methodology are currently key questions to consider in our study (McFadden & Train 2017).

Thus, we contribute in this paper to the aforementioned debate by using a sequential contingent valuation survey to two environmental goods along with a real economic experiment with incentives. Consequently, we follow the Denant-Boèmont et al.'s (2018) methodology in order to keep under control the hypothetical bias, comparing contingent valuation survey results to laboratory economic experiment outcomes. We significantly extend that empirical strategy in order to control other biases, apart from the WTP overstatement. Accordingly, we have combined two different environmental nuisances with two survey instruments (hypothetical survey and laboratory experiment), along with two possible valuation sequences where participants are requested to choose among contributions of quasi-public goods. Furthermore, as we reversed the valuation sequence for half of our respondents, we can observe different behavioural biases: *Hypothetical bias* (Cummings et al 1995), *Correlation effect* (Kahneman & Ritov 1994) and, finally, *Sequencing effect* (Dupont 2003, Hanemann 1994, Carson and Mitchell 1995). Now, we can anticipate that we have obtained empirical evidences both for hypothetical bias and correlation effect, but we did not find meaningful results about the sequencing effect.

The whole paper is organized in five sections. Section 2 describes the methodological details of our procedure along with its theoretical background, whereas Section 3 presents the experimental design we have developed with its geographical scope. Section 4 depicts the main findings and results of our work. Lastly, the final paper discussion and the main conclusions of our study along with some research lines recommendations are presented in Section 5.

2. Literature Review and Theoretical Background

After the general introduction to the environmental pollution problem generated by road transportation (Sawik et al. 2017; Faulin et al., 2019), where we have described some of the main WTP biases, now, we depict some of the main bibliographic sources. These references provide the theoretical background that gives support to this research paper methodology.

2.1. Literature review

First of all, we would like to focus our attention of the Contingent Valuation methodology (Mitchell and Carson, 1988; Hanemann, 1994; Dupont, 2003) which constitutes the cornerstone of the methodology we are using in this paper along with the Experimental Economics (Abeler and Nosenzo, 2015). Therefore, when we are searching for a valuation sequence which involves different environmental goods (Carson and Mitchell, 1995), two different real effects can be observed:

1) The *sequencing* or *ordering effect*: The valuation stated by a survey respondent may depend on particular position of the question when the valuation occurred in the sequence. That is to say, that the valuation given by a person when s/he is asked about a particular issue in the first position in the sequence, is different than the valuation given being asked in a different place in the sequence (Tolley et al., 1986).

2) The *embedding effect* (Kahneman and Knetsch 1992; Carson and Mitchell 1995): the valuation stated by a survey respondent might depend on whether a particular commodity is nested or not in a more general commodity (e.g. the WTP to mitigate air pollution in Pamplona and the WTP to mitigate air pollution in the same amount in Navarre).

Furthermore, Hanemann (1994) and Dupont (2003) pointed out that the *embedding effect* combines, in fact, three different concepts. The first one is that the WTP varies with change in the scale or scope of the considered public good (*scope effect*). The second one is the *sequencing effect*, coming from diminishing the marginal utility and the substitution effects, where the WTP for a particular good valued in the first position may be greater or equal to the same good valued in a second, third, fourth,… positions (Carson et al., 1992). The third concept is associated to the *sub-additivity effect*: the WTP for a composite change in a group of public goods may be less than the addition of all the WTPs for each individual public good. In our study, sub-additivity or scope effects are not considered, as we focus our attention on the sequencing effect.

Furthermore, Randall and Hoehn (1996) showed that biases in valuation mainly happen when a project is valuated on its own being part of a more general policy agenda. They showed that the sum of the partially valued projects on the agenda, exceeds the simultaneous valuation, and claimed that the only way to get the all projects total value is a one-shot holistic valuation or a sequence approach methodology. Concerning the embedding effect (Halvorsen, 1996; Poe, 2016), Diamond and Hausman

(1994) suggested an adding-up test, where the value of each object taken separately should be compared to the value of all objects considered in a conjoint way. Moreover, Halvorsen (1996) made use of a contingent valuation survey design to test the embedding and the sequencing/ordering effects, focusing his attention on the information providing to respondents. Likewise, he found a strong evidence for the latter, but not for the former.

Additionally, considering only the ordering effect, an important lesson to be learned from Halvorsen (1996) is that all the nuisance information is given by parts sequentially to some respondents, and simultaneously to others. This protocol of revealing information to the respondents is the basis of the ordering effect observed in many tests. Moreover, as ordering effects may be due to a sequential revelation of information, and as there is no change within the valuation sequence for a particular nuisance, we are going to implement a *pure* ordering effect test, where nuisances are presented in reversed order from one sample to the other.

2.2. Theoretical background

Our microeconomic background is based on Carson and Mitchell (1995) and Burrows et al (2017), who defined rigorously the embedding effect and the empirical tests that could be done to identify it. A microeconomic formulation for *sequencing or ordering effects*, partly inspired by Whitehead and Blomquist, (2006), would be the one presented in the following paragraph.

Let us assume that an individual *k* having as utility function $U(x, a, b)$, with x the level of consumption for a private good, and let it be a, b two quasi-public environmental goods. Similarly, let us assume two quasi-public goods (the quality of ambient noise and the quality of local air, respectively), with initial levels denoted a and b, and increments for both levels defined as α and β. In a similar way to Carson and Mitchell (1995), $\{\alpha_k^s, \beta_k^s\}$ will denote the increments to be valued by the respondent/subsample $k = \{1,2\}$ at position $s = \{1,2\}$. We suppose that these environmental goods are *normal* goods (Hicksian substitutes), in particular, we assume that we have non-nested goods (not goods for which one is the subset of the other). Without any loss of generality, we can assume that the valuation sequence V to be $V(\alpha_k^1, \beta_k^2)$ where 1 denotes the increment of the good to be valuated first and 2 the other good increment to be valued in second position *within* the valuation sequence. The WTP for the respondent k for the first increment is given by

$$v(p, a_k, b_k, y) = v\left(p, a_k + \alpha_k^1, b_k, y - wtp(\alpha_k^1)\right), \tag{1}$$

where v is the indirect utility function, p the vector of prices for private goods and y respondent k's income. After the first valuation task, the WTP for second increment is

$$v\left(p, a_k + \alpha_k^1, b_k, y - wtp(\alpha_k^1)\right) = v\left(p, a_k + \alpha_k^1, b_k + \beta_k^2, y - wtp(\alpha_k^1) - wtp(\beta_k^2)\right). \tag{2}$$

Our experimental design implies corresponds to

Table 1: Experimental tests for sequencing effect

Sample position	1 (or A)	2 (or B)
1	$\alpha_1^1,$	$\beta_2^1,$
2	$\beta_1^2,$	$\alpha_2^2.$

The external tests for sequencing effect would be, following Carson and Mitchell (1995)[2]

$$wtp(\alpha_1^1) \geq wtp(\alpha_2^2), \quad (3)$$

and

$$wtp(\beta_2^1) \geq wtp(\beta_1^2). \quad (4)$$

The following experimental design implements precisely this sequencing effect test.

3. Experimental Design and its Geographical Approach

This section is going to be devoted to the design of the Economic Experiment for eliciting the values of the WTP we have mentioned in the previous paragraphs. The purpose of this experiment is the obtaining of a set of economic valuations considering two survey frameworks: the first one, a standard *Contingent Valuation Method (CVM)*, and, secondly, an *Economic Experiment with real monetary incentives*. In order to obtain sincere valuation in the economic experiment, we use an incentive-compatible procedure described in Denant-Boemont et al. (2018), and based upon Horowitz (2008) and Messer et al (2010), that we describe briefly in this section. Moreover, our experiment has been geographically located in Navarre (Spain), and we explain this scenario in the following paragraph.

Thus, Navarre is located in the Western border between Spain and France, being one of the seven European regions having Pyrenees included in their geographical scope. The study area was selected because of its prominence as a natural boundary between Spain and France, and the importance of the road freight traffic crossing the Pyrenees in that area: More than 141,820 vehicles, almost 25% of which are freight trucks, cross daily the Pyrenees (Spanish-French Observatory of Pyrenees Traffic, 2018). In fact, the areas with a heavier traffic are sited in Catalonia (La Junquera), and the Basque Country and Navarre (Irún-Behovia) (Figure 1). Hence, we have developed our experiment, explained hereinafter, in this Pyrenees area.

[2] We do not perform adding-up test since our design is not suitable for that.

Figure 1. Geographical scope: Importance of Navarre in the Freight Transportation crossing the Pyrenees. Adapted from the Spanish-French Observatory of Pyrenees Traffic (2018) report.

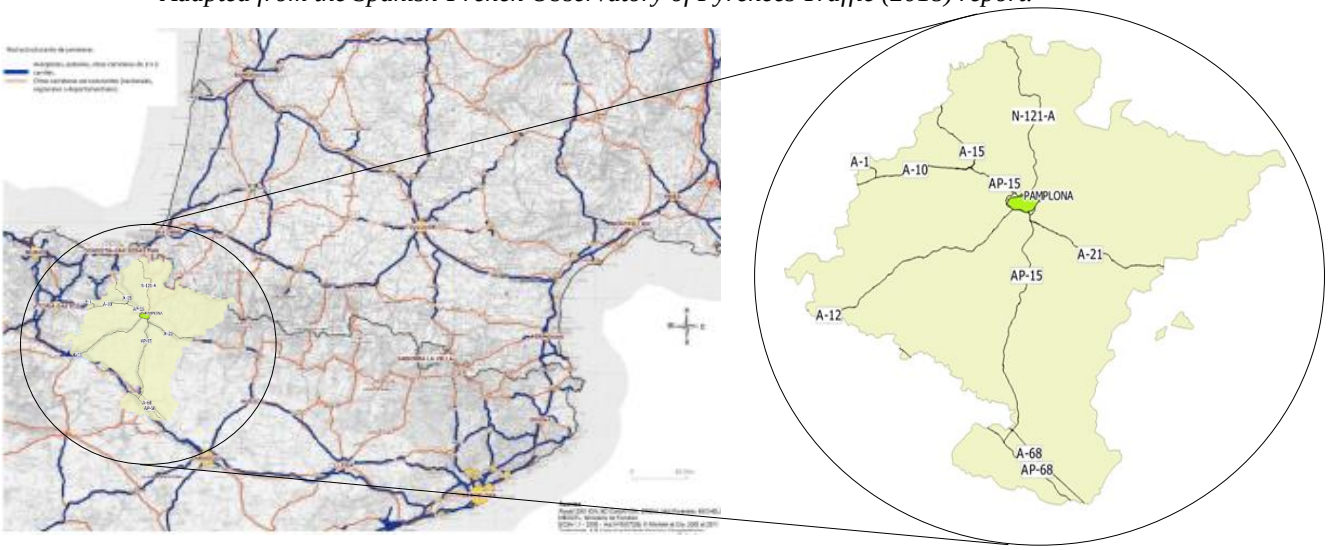


Fig. 1.a. Map of the main roads (in brown) and motorways (in blue) crossing the Pyrenees border between Spain and France

Fig.1.b. Map of Navarre showing the main transportation routes for trucks.

Hence, our fieldwork description can be done as follows. The experiment has been split up into two groups of people, having 25 participants each one, and was developed in Pamplona (Navarre). These two groups of people were randomly selected from candidates who volunteered according to some newspapers and social networks requests. Furthermore, the participants in each group have to fill in two types of sessions established on a survey sequence: on the one hand, a Contingent Valuation Survey (CVS) session, considering two kinds of nuisances (noise and pollution) related to freight traffic; on the other hand, a Real-Incentivized Experiment Survey (RIES) session, about their individual WTP to mitigate the aforementioned nuisances of noise and pollution (see the Appendix A about the instructions for CVS and RIES sessions). Moreover, all participants were aware that the whole survey is organized in four parts or sections, but no additional precision was settled about the nature of each part at the beginning of the session.

Likewise, before taking part in the two sessions of the survey, all the participants in both groups have had a presentation of two organizations related to the mitigation of transport noise and air pollution arisen from transportation. The first organization was Red NELS (https://www.navarra.es/home_es/Temas/Medio+Ambiente/Sostenibilidad/Red+NELS.htm), a local institution in Navarre connected to city halls in the region, which explained all their actions, locally focused, to control and compensate the transportation externalities, previously mentioned. The second organization was Greenpeace (https://es.greenpeace.org/es/) which, having a worldwide action scope, also presented a group of global actions they are carrying out with the purpose of environmental protection against transport noise and pollution. Each presentation lasted 20 minutes and their order of appearance was randomly determined: Red NELS was selected as the first presenter and, consequently, Greenpeace was the second one.

Afterwards, some survey sequence activities were designed for all participants Thus, the 50% of the participants (Group A), randomly selected, were assigned to a given room to complete the CVS session, and the other 50% (Group B) were assigned to a computer room to complete the RIES session. After finishing the first part of the sequence on behalf of each group (either CVS or RIES sessions), participants have to complete the second session with the part done by the other group during their initial session. This change of roles between the two groups is done by swapping their respective rooms were the activities were taken in the first session. During this rooms change, participants were not allowed to communicate in any way.

Additionally, the CVS session (also denoted as Hypothetical part) was successively conducted on transport noise and pollution, being the order reversed depending of the subsample (see Table 1). In contrast, the RIES session includes a monetary endowment of 60 euros per respondent and nuisance (noise and air pollution) and each participant had to decide how much money they want to donate to one of the aforementioned associations which did the presentations (Red NELS or Greenpeace). Having the purpose of ensuring incentive compatibility, we used the procedure designed by Horowitz et al. (1999), Horowitz (2008) and Messer et al. (2010), called *group format elicitation procedure*. According to that methodology, each participant belonging to a group of size N is enquired to state her/his WTP to obtain a certain level of reduction for a specific damage (noise and air pollution in our case). Thus, our survey respondents would indicate how much of a given endowment €E s/he would be ready to allocate to reduce damage actions.

Besides, all WTP statements were anonymous and simultaneously collected through a computer in a first round. In the second round, the revealed WTPs are ranked from the lowest value €b_i to the highest one, €B_i where i =1,…,N. Then, let it be €x a random number between €b_i and €B_i, uniformly generated. Therefore, the following rule is applied to settle individual payoffs. If more than 50% of the individual statements that lie between €b_i and €B_i are greater or equal to €x, then each group member gets privately €$(E-x)$ and, as a consequence, €Nx are allocated to environmental damage reduction. Conversely, if less than 50% of individual statements (comparison with the median bid) are greater or equal to €x, then each group member gets privately €E and nothing is allocated to environmental damage reduction. We show that this procedure ensure truth revealing regarding actual WTPs for each respondent in Denant-Boemont et al. (2018). The money allocated to environmental damage reduction (either noise or air pollution) would be assigned to the institutions Red NELS or Greenpeace according to the choice of each respondent.

Accordingly, all Group A participants complete firstly the hypothetical part of the CVS session and after the RIES part. During each part, Group A was exposed first to a noise valuation procedure and after to a pollution valuation one. Similarly, Group B developed that sequence in a reverse way (firstly the RIES part and afterwards, the CVS session), and during each part they were exposed to a pollution valuation procedure firstly, and a noise valuation protocol, afterwards. Table 2 summarizes the sequence for each group (see the Appendix A about the instructions for CVS and RIES sessions).

Table 2. Experimental Design and subsequent parts of the experimental sequence

Group / Part	A	B
1	Presentations by Red NELS & Greenpeace	
2	Hypothetical Noise CVS survey	Real-Incentivized Pollution Experiment
3	Hypothetical Pollution CVS survey	Real-Incentivized Noise Experiment
4	Real-Incentivized Noise Experiment	Hypothetical Pollution CVS survey
5	Real-Incentivized Pollution Experiment	Hypothetical Noise CVS survey

Likewise, our design enables us to control for various biases that are well-known in the behavioral literature:

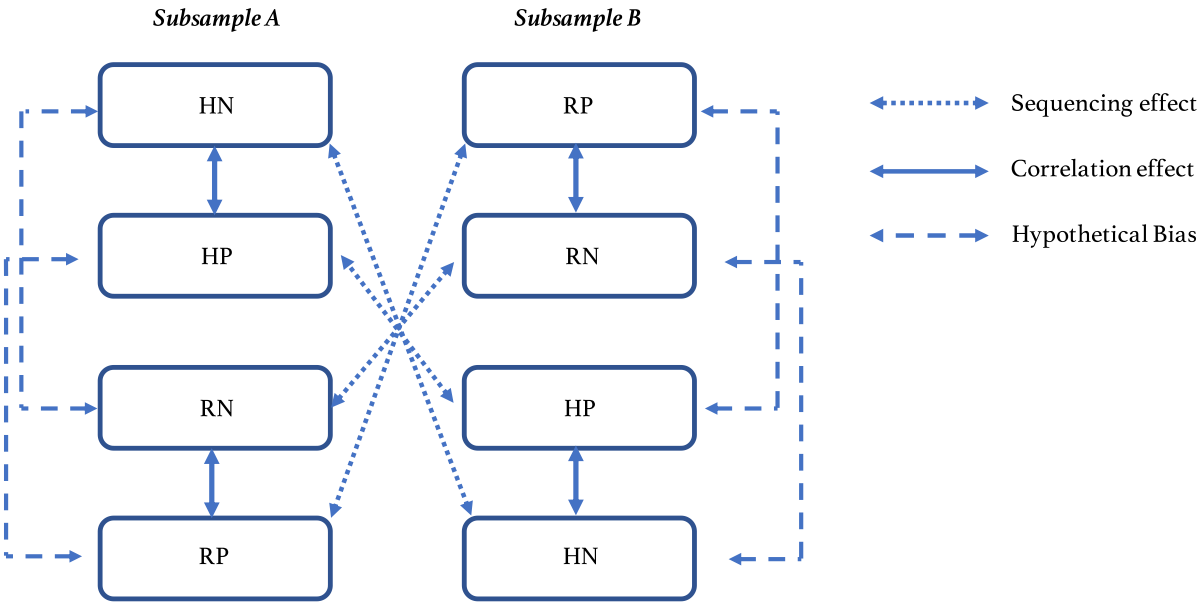
i. The first one is *hypothetical bias* (Denant-Boemont et al. 2018; Carson and Hanemann, 2005; Foster and Burrows 2017), that could be tested by comparing WTP for the same commodity under hypothetical or real-incentives frameworks (CVS versus RIES) for a particular subsample (respondent).
ii. The second one is a *correlation effect*, that is to say a statistical relationship between economic valuations for different goods that are not close substitutes, by comparing under the same framework (CVS or RIES) for a given subsample and for each respondent, WTP stated for noise to the WTP stated for pollution,
iii. The third one is a *sequencing effect*, where the economic valuation for a (non-nested) good depends on the position where this valuation occurs in the valuation sequence. The comparison is, therefore, between subsamples, when WTP under a particular framework and for a specific good is stated in first or second position during the valuation sequence.

More precisely, a *correlation effect* means that the outcome of valuation in a given sequence for a particular impact may be statistically related to the outcome of valuation for another impact. In our design, it implies that valuations reported in noise (respectively in pollution) might impact the valuations reported in pollution (respectively in noise). *Primacy effect* would be a particular case of sequencing effect where the valuation for the second outcome that come in the experimental sequence would be lower than the valuation stated for the first outcome. *Recency effect* would be the reverse one, the last valuation being higher than any former one. Our empirical strategy for observing the different possible cognitive biases is therefore based on the combination of a between-subject design and of a within-subject design (see Figure 2).

Considering the within-subject design, it enables us to compare valuations given by a same respondent at different steps of the valuation sequence. Being confronted successively to pollution and noise (or the reverse) under the same survey method (CVS or RIES) enables us to control for *correlation effect*, and having to consider the same environmental good under two different survey methods for hypothetical

bias. The between-subject characteristic of our design implies that the sequence valuation ordering differs from one subsample to the other, enabling us to control for *sequencing effect*.

Figure 2: Empirical strategy for non-nested environmental goods



NB: H is for Hypothetical, R for Real-Incentivized, P for air pollution and N for noise.

4. Results

This section is devoted to the description of our experiments results with the presentation of the cognitive biases identified in our experiments. Firstly, we have performed a descriptive analysis in order to highlight the homogeneity within the two subsamples. Later, we have presented the results for the hypothetical bias, and, finally the results corresponding to the sequencing bias. Furthermore, we define the notions of zero-protester respondent and payer respondent, as the survey participants who, respectively, declare to pay zero for being against the payment procedure, and declare to pay any amount of money greater than zero. These concepts are very well-known in Contingent Valuation methodology and their definitions can be found, for instance, in Lera-Lopez et al. (2014) or Denant-Boemont et al. (2018).

4.1. Subsamples homogeneity

Having the purpose of establishing the role of the cognitive bias in a suitable way, we have to analyze the respondents' groups in order to ensure that two taken subsamples were homogeneous in terms of socioeconomic attributes (age, income, education level, and other behavioral factors) as well as the concern and perception of the noise and air pollution problems. It is important to remember that the

participants were randomly assigned to each group. As it can be observed in the Table 3, both groups A and B show similar patterns, especially in the noise/pollution concerns and education level. These observational traits along with the building structures of the subsamples imply that the hypothesis of subsamples homogeneity is sound and can be assumed without contradiction observed in the data.

Table 3. Descriptive analysis among sample groups

	Group A	Group B	Full Sample
Age			
From 18 to 24	48%	56%	52%
From 24 to 50	48%	40%	44%
Older than 50	4%	4%	4%
Income			
Low	4%	20%	12%
Mid-low	44%	28%	36%
Mid-high	32%	20%	26%
High	20%	32%	26%
Education level			
None	0%	0%	0%
Basic	4%	4%	4%
Medium	24%	32%	28%
University	72%	64%	68%
Other factors			
Man (yes)	64%	32%	48%
Smoking (yes)	24%	28%	26%
Doing sport (Yes)	84%	76%	80%
Owing a car (yes)	32%	44%	38%
Noise concern			
Not at all worried	52%	48%	50%
Something Worried	16%	24%	20%
Moderately concerned	24%	20%	22%
Pretty worried	0%	8%	4%
Very worried	8%	0%	4%
Air pollution concern			
Not at all worried	28%	32%	30%
Something Worried	48%	28%	38%
Moderately concerned	12%	20%	16%
Pretty worried	8%	20%	14%
Very worried	4%	0%	2%

4.2. Hypothetical bias

Finding hypothetical biases is expected in experiments such as we have built here. We first perform some Wilcoxon Mann-Whitney signed rank (non-parametrical) bilateral tests, where we compare WTP stated under hypothetical or real-incentives settings for each nuisance, per subsample and only considering respondents that are not zero-protesters. We also carry out the same tests for respondents that do not state a zero WTP for both settings (that is stating a zero for a given nuisance under hypothetical and real-incentives). Results are shown in Table 4.

Table 4. Results obtained from the application of the Wilcoxon-Mann-Whitney test to the subsamples A and B

Who?	Comparison	Noise Subsample A	Pollution Subsample A	Noise Subsample B	Pollution Subsample B
Non-zero protesters		Z=1.874 p=0.0609*	Z=1.522 P=0.1280	Z=0.432 P=0.6656	Z=0.882 P=0.3777
Non-double zero respondents		Z=2.100 P=0.0358**	Z=1.863 P=0.0625*	Z=0.952 P=0.3411	Z=1.852 P=0.064*

Observing the results given in Table 4, there is a clear evidence for the existence of hypothetical bias (taking a 10% significant level, for instance) in all scenarios except the case of Noise Subsample B. Should we perform unilateral tests, due to the fact that the literature about hypothetical bias suggests lower valuations when real incentives are considered, then the hypothetical bias would be significant in all cases. Clear evidence for hypothetical bias appeared in our sample.

4.3. Correlation effects between CVS and RIES parts

We now turn to study correlation effects between the WTPs observed in both sessions of our survey. As there is a significant number of zero responses in our survey, we used Tobit regressions for contrasting the results of the CVS and RIES parts. Likewise, we have also run these regressions on each session (Session A or Session B). Table 5 reports the corresponding coefficients for the Tobit regression models. In Table 5, all the coefficients are significant at 1% level, with exception of the regression run for the CVS part in session A. We have estimated conditional marginal effects at means, with the purpose of make a good interpretation of the Tobit coefficients.

The analysis of Table 6 suggests that an increase of one unit of Hypothetical WTP for pollution will produce, on average, an increase of 0.76 unit of subsequent Hypothetical WTP for noise in subsample A.

Table 5. Results (first step) obtained from Tobit regressions contrasting CVS and RIES parts

	Session A		Session B	
Explained Variable	wtpH(Poll)	wtpR(Poll)	wtpH(Noise)	wtpR(Noise)
Explanatory Variable				
wtpH(Poll)			+2.360***	
			(0.797)	
wtpH(Noise)	+0.039			
	(0.514)			
wtpR(Poll)				+1.012***
				(0.137)
wtpR(Noise)		+0.936***		
		(0.172)		
Left censored obs	12	11	12	6
Right Censored obs	0	0	0	0
Pseudo R2	0.000	0.156	0.055	0.169

Table 6. Results (second step) obtained from Tobit regressions contrasting CVS and RIES parts

	Session A		Session B	
Explained variable	wtpH(Poll)	wtpR(Poll)	wtpH(Noise)	wtpR(Noise)
Dy/dx	+0.014	+0.454***	+0.760***	+0.657***
	(0.179)	(0.094)	(0.249)	(0.116)

4.4. Sequencing bias

Now, we focus our attention of the sequencing bias of our survey. Thus, the analysis of the aforementioned survey bias is performed based on the following steps. First of all, Tables 8 and 9 show the average and median valuations for each commodity, i.e., noise and air pollution, in the hypothetical and real setting, respectively. These Tables 8 and 9 also provide that information when the samples are restricted to non-protesters (*ZP removed* column) and to payers (*Payers* column). Notice that the zero protesters are identified using additional questions, when the respondent gives a zero WTP for any good. Thus, that zero observation, once is identified as a protest, is removed from the sample forming the *ZP removed* subsample. Additionally, the *Payers* subsample contains only observations with a WTP value strictly greater than zero for each good. Finally, Tables 8 and 9 also depict information for the values obtained in the two groups in the experiment according to Table 1. To this respect, we also provide the results for the full sample (*FS*), i.e. when both groups are considered.

Secondly, Tables 9 and 10 provide the same kind of average and median valuations, but considering a change in the order of goods valuations. Thus, we can see the first valuation in the two groups, without paying attention in the type of good that was to be valuated, and for each evaluation setting, i.e. real and hypothetical one. Moreover, Tables 9 and 10 also consider the same subsamples described above for the *ZP removed* and the *Payers*. Additionally, the specific group values are also shown in those tables.

Likewise, Figures 2 to 5 show the graphical results for the aforementioned ordering analysis. Accordingly, Figures 2 and 3 describe the ordering effects using the three subsamples we have defined, i.e. FS, ZP, and PAY, in both the real and the hypothetical settings. This comparison is performed, initially, with a distribution description with boxplots analysis in Figure 2 and with barplots for the WTP averages in Figure 3. The inter-group analysis is further revealed in the Figures 4 and 5, which show the barplots for the average WTP in the hypothetical (Figures 4) and real (Figures 5) settings for the orders valuations maintaining the aforementioned distinctions for the FS, ZP and PAY subsamples. The point of these figures is that they show the different valuations given per each group depending on the order the good is presented, i.e. first and second good, and the type of good presented, i.e. air or noise pollution.

If we analyze in detail Tables 8-10 and Figures 2-5, we can pinpoint a group of interesting insights.

i. Firstly, we can observe a clear hypothetical bias, as described in subsection 4.2. Namely, the values given for valuation the goods are significantly higher in the contingent or hypothetical scenario rather than in the real one. On average, these hypothetical values double the real ones. The in-depth study on the hypothetical bias is available in Denant-Boemont et al. (2018).
ii. Secondly, we observe a sequencing effect or recency effect in the hypothetical settings whilst there is not such an effect in the real setting. Specifically, in our experiment the ordering effects arise in hypothetical scenarios but are not observed in the real-incentivized settings. This sequencing effect is even more evident in Figure 3, and, particularly, comparing the Figures 4 and 5 where the recency effect disappears as real valuation is performed. Actually, considering groups, that recency effect is even clearer in Group B as the second valuation given in that group was always greater. For instance, let us focus our attention on the hypothetical full sample in Table 8. For Group A, it is more valuable the pollution variable whereas for Group B it is more valuable the noise variable. This situation is also kept in the case of real valuation in Table 9, but with a lower effect. The point is that, for both groups, the good in second position always was more valuable, no matter if it was noise or air pollution; as we can observed in Tables 10 and 11.
iii. Thirdly, the medians remain unaltered in almost all cases as can be observed in Figure 2. Actually, non-parametric tests give no differences for order comparisons, as we present in the next paragraph.

To sum up, this sequencing effect is very smooth in our data, probably due to small sample size. Having the purpose of obtaining statistical evidence, and due to the fact that the number of observations is less than 30, we perform some bilateral non parametric comparison tests. These tests consist firstly in a

Mann-Whitney rank sum test and, secondly in a Two-Sample Robust Rank Order test[3], where WTPs for respondents in subsample A are compared to the ones stated by subsample B. If we consider only respondents that are non-zero protesters, the results are given in Table 7. None of these tests succeed to reject the null hypothesis of equality in WTPs.

Table 7. Results obtained from Mann-Whitney rank sum and Two-Sample Robust Rank Order tests to contrast the WTPs equality

Test	Comparison	Hypothetical Noise	Hypothetical Air Pollution	Real-Incentives Noise	Real-Incentives Air Pollution
Mann-Whitney		Z=-0.611, p=0.541	Z=0.147, p=0.883	Z=0.850, p=0.395	Z=0.267, p=0.789
Robust Rank Order		p=0.280	p=0.444	p=0.201	p=0.401

[3] The Wilcoxon-Mann-Whitney test which assumes that samples are drawn from the same population, which implies that under the null hypothesis, they have not only the same first moment, *but also the same higher order moments such for both hypotheses*. The Robust Rank Order test is a correction of this shortcoming (see Feltovich 2003).

Table 8. Average and median analysis of air and noise pollution in the hypothetical valuation. Goods Hypothetical Valuation

	Full Sample						*ZP removed*						*Payers*					
	Noise	Median	n	Pollution	Median	n	Noise	Median	n	Pollution	Median	n	Noise	Median	n	Pollution	Median	n
FS	24.02 (48.81)	7.50	50	19.84 (34.35)	5.00	50	27.30 (51.22)	10.00	44	26.11 (37.34)	10.00	38	42.90 (59.03)	20.00	28	34.21 (39.44)	30.00	29
A	21.96 (30.71)	10.00	25	22.29 (42.74)	5.00	25	23.87 (31.33)	10.00	23	29.33 (47.09)	10.00	19	36.61 (32.29)	20.00	15	39.80 (51.20)	30.00	14
B	26.08 (62.55)	5.00	25	17.40 (23.87)	5.00	25	31.05 (67.32)	12.00	21	22.89 (25.04)	10.00	19	50.15 (80.72)	20.00	13	29.00 (24.83)	30.00	15

FS: Full Sample; A: Subsample A; B:Subsample B

Table 9. Average and median analysis of air and noise pollution in the real valuation. Goods Real Valuation

	Full Sample						*ZP removed*						*Payers*					
	Noise	Median	n	Pollution	Median	n	Noise	Median	n	Pollution	Median	n	Noise	Median	n	Pollution	Median	n
FS	9.54 (13.31)	5.00	50	9.58 (12.30)	5.00	50	10.25 (13.97)	5.00	44	11.37 (13.28)	5.00	38	14.91 (14.06)	10.00	32	14.52 (12.56)	10.00	33
A	7.60 (11.56)	1.00	25	7.92 (10.31)	1.00	25	8.22 (11.87)	2.00	23	10.32 (10.79)	10.00	19	14.62 (12.50)	10.00	13	14.14 (10.08)	14.50	14
B	11.48 (14.85)	5.00	25	11.24 (14.02)	5.00	25	12.48 (15.95)	5.00	21	12.42 (15.61)	5.00	19	15.11 (15.37)	10.00	19	14.79 (14.38)	7.00	19

FS: Full Sample; A: Subsample A; B: Subsample B

Table 10. Average and median analysis of first and second valuations in the hypothetical setting. Ordering Hypothetical Valuation

	Full Sample						*ZP removed*						*Payers*					
	First	Median	n	Second	Median	n	First	Median	n	Second	Median	n	First	Median	n	Second	Median	n
FS	19.68 (27.32)	7.00	50	24.18 (53.05)	5.00	50	23.43 (28.32)	10.00	42	30.23 (57.87)	11.00	40	32.80 (28.57)	22.00	30	44.79 (65.92)	20.00	27
A	21.96 (30.71)	10.00	25	22.29 (42.74)	5.00	25	23.87 (31.33)	10.00	23	29.33 (47.09)	10.00	19	36.61 (32.29)	20.00	15	39.80 (51.20)	30.00	14
B	17.40 (23.87)	5.00	25	26.08 (62.55)	5.00	25	22.89 (25.04)	10.00	19	31.05 (67.32)	12.00	21	29.00 (24.83)	30.00	15	50.15 (80.72)	20.00	13

FS: Full Sample; A: Subsample A; B: Subsample B

Table 11. Average and median analysis of first and second valuations in the real setting. Ordering Real Valuation

	Full Sample						*ZP removed*						*Payers*					
	First	Median	n	Second	Median	n	First	Median	n	Second	Median	n	First	Median	n	Second	Median	n
FS	9.42 (12.85)	5.00	50	9.70 (12.78)	5.00	50	10.12 (13.67)	4.50	42	11.45 (13.62)	5.50	40	14.72 (13.44)	10.00	32	14.70 (13.20)	10.00	33
A	7.60 (11.56)	1.00	25	7.92 (10.31)	1.00	25	8.22 (11.87)	2.00	23	10.32 (10.79)	10.00	19	14.62 (12.50)	10.00	13	14.14 (10.08)	14.50	14
B	11.24 (14.02)	5.00	25	11.48 (14.85)	5.00	25	12.42 (15.61)	5.00	19	12.48 (15.95)	5.00	21	14.79 (14.38)	7.00	19	15.11 (15.37)	10.00	19

FS: Full Sample; A: Subsample A; B: Subsample B

Figure 2. Boxplots for the WTP values in real (R, at the top) and hypothetical (H, at the bottom) settings for the first and second valuations made with full sample (FS) and the 'zero protesters removed' (ZP) and 'payers' (PAY) subsamples.

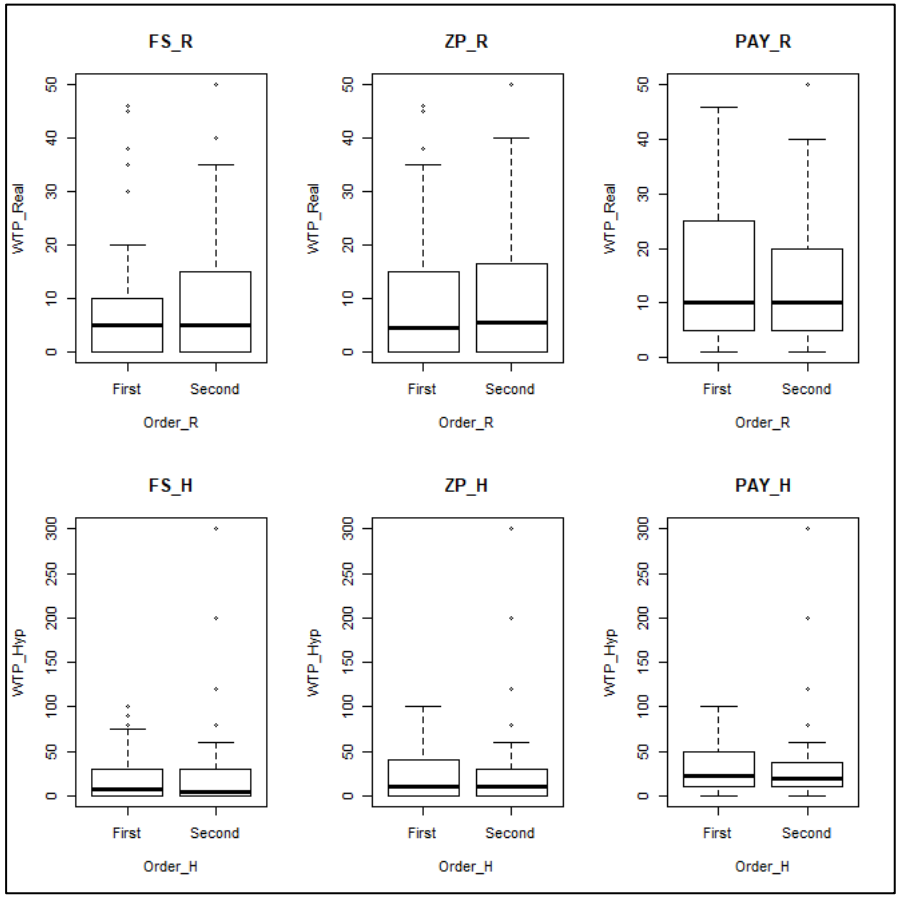


Figure 3. Barplots for the WTP means in real (R, at the top) and hypothetical (H, at the bottom) settings for the first and second valuations made with full sample (FS) and the 'zero protesters removed' (ZP) and 'payers' (PAY) subsamples.

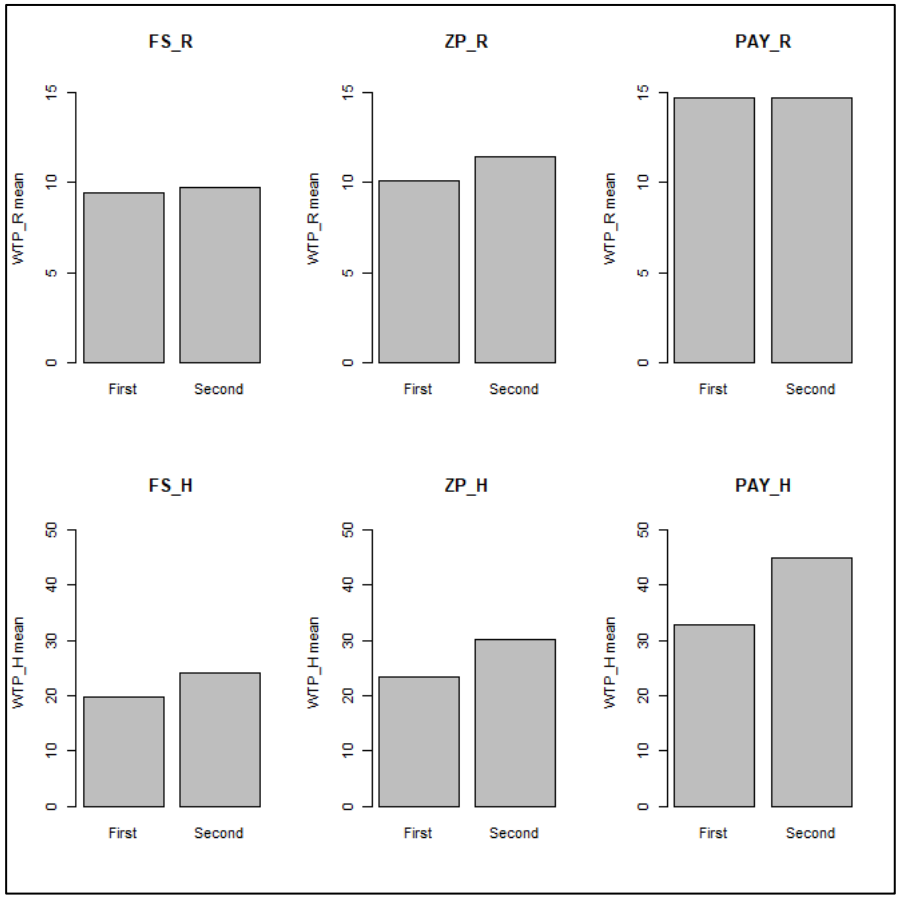


Figure 4. Barplots for the WTP means in the hypothetical settings for the first and second valuation made with full sample (FS) and the 'zero protesters removed' (ZP) and 'payers' (PAY) subsamples.

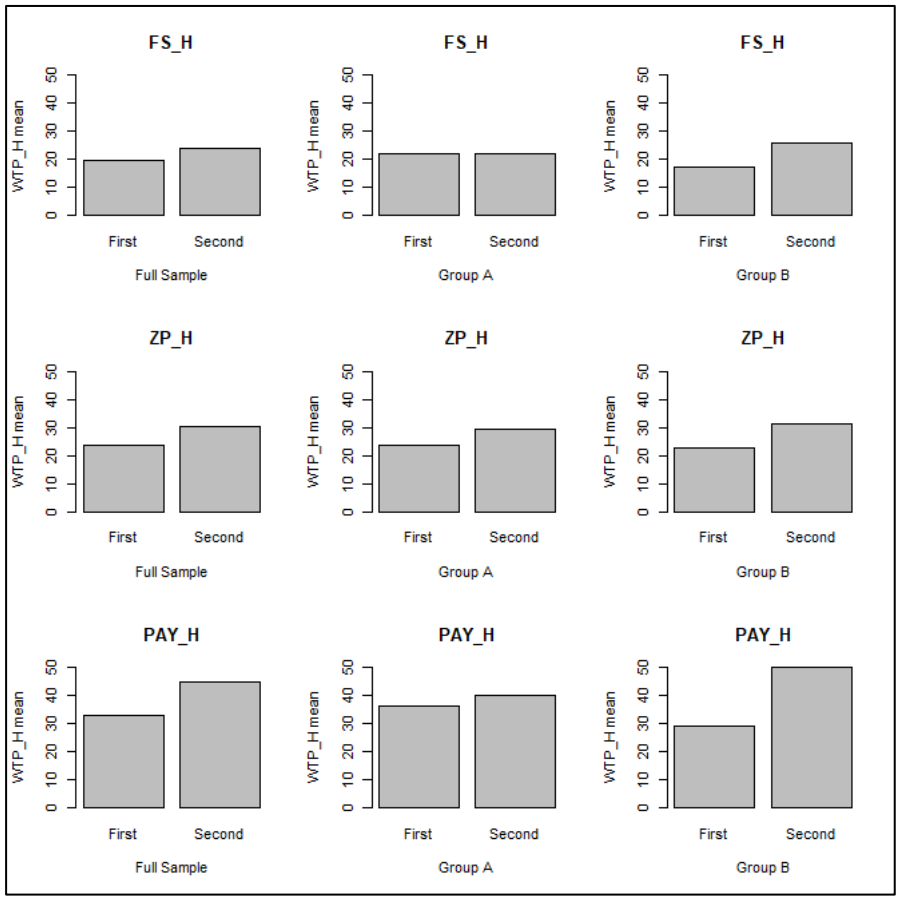
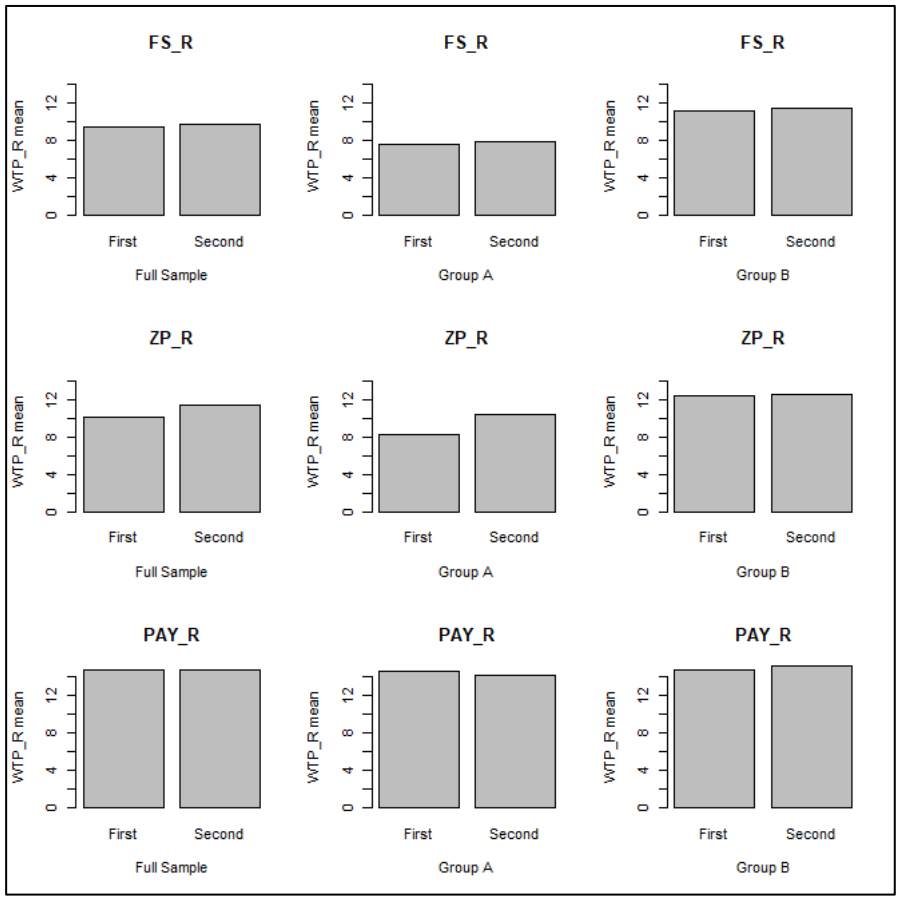


Figure 5. Barplots for the WTP means in the real settings for the first and second valuation made with full sample (FS) and the 'zero protesters removed' (ZP) and 'payers' (PAY) subsamples.



5. Discussion

The possible impacts of behavioral cognitive biases in Stated Preference (SP) studies – like Contingent Valuation-CV surveys – on value elicitation had been extensively discussed in the literature (see, e.g., Johnston et al (2017) for a survey). Thus, methodological choices are to be adapted to these possible impacts, in particular regarding the framing of response option for the chosen question format. Johnston et al. (2017) discuss 5 areas for methodological choices regarding value elicitation: (1) attribute or non-attribute methods; (2) the choice of welfare measure to be elicited (3) the framing of response options for the question format; (4) the choice of payment vehicle and (5) the use of auxiliary questions to evaluate validity.

To sum up our results, we have obtained clear statistical evidence for hypothetical bias and correlation effect, and more mixed results about the sequencing effect: under the hypothetical survey, the latter economic valuation tend to be greater than the former, but this result does not appear in the economic experiment. It is important to notice that, most often, hypothetical bias in the literature had been identified when comparing Stated Preferences data to Revealed Preference ones (see Fifer et al. (2014), Krcal et al. (2019) for evidence regarding the valuation of time). One strength of our survey is to compare SP data obtained under a hypothetical framework to others resulting from a real incentivized one, which guarantees a certain degree of robustness for our results.

Furthermore, we had, in our design, the same response format (a continuous contingent valuation survey) whatever real incentives exist or not for each nuisance. By doing this, we rule out possible differences in incentives properties due to the format (Carlson and Groves, 2007). The format chosen helped also to control for sequence effects that may arise with successive binary choices (Carson and Mitchell, 1995; Day et al., 2012), as respondents were confronted to single choices consisting in stating WTPs. Moreover, as we repeat the same choice procedure in a within-subject design to assess differences between hypothetical valuations and real-incentivized ones, we control also for hypothetical bias. We also ensure to have an incentive compatible mechanism format to elicit individual WTPs by implementing an auction-based procedure at the group level (Horowitz, 2008).

We also obtain a correlation effect, i.e., a statistical relationship between valuations for different goods. As underlined by Bateman et al (2008), an anchoring bias, i.e., a psychological process that occurs when a reported or revealed valuation for a given respondent is correlated with some prior numerical cue, could explain that result. There had been numerous and consistent evidence for anchoring bias, which seems to depend strongly from elicitation format. For instance, Champenois et al (2018) showed that in the case of a contingent valuation survey, a circular payment card reduces anchoring compared to the standard one (closed-ended formats). Our results showed that, even with an open-ended format as ours, individual bids to obtain the nuisance reduction, that should avoid anchoring coming from numerical

cue, we have obtained some endogenous anchoring bias. This situation is due as the first bid stated by our respondents for a given nuisance anchors latter valuations for other nuisances.

In contrast, the empirical evidence for sequencing effect is only observed under hypothetical scenarios, which differs from most part of the literature. For instance, Page and Page (2010), using RP data from the American Idol contest, show empirical evidence for sequence effects: a primacy effect and a recency effect. The former denotes the fact that participants to valuation or choice studies remember better the beginning (*primacy*) of a sequence of choices when the latter denotes the fact that participants remain better the end of a sequence of choices (*recency*). This combines in a U-shape curve of memory. Payne et al (2000), using a series of CV surveys for environmental goods, observe a strong primacy effect when, at most, we obtain the reverse (a recency effect). Day et al (2012) also show the presence of strong order effects in discrete choice experiments applied to water consumption depending on the format of the sequence survey. They observed order effect evidence for *stepwise disclosure format*: that is to say that respondents are informed at the end of a step that a new step occurs, and thus the total sequence is unknown to respondents. Likewise, there have been almost no evidence for the *advanced disclosure format*, when the full sequence is revealed to respondents at the beginning, for which valuation remains stable. We are not aware of any other transportation study as ours that would focus either on anchoring or on sequential effect and consequently it is a necessity to gather more data to build stronger evidence.

6. Conclusion

We have presented in this paper the outcomes we obtained developing an experimental economics methodology to make estimations of the WTPs for air pollution mitigation and noise abatement in the road transportation crossing the Pyrenees in the Navarrese region (Spain). Denant-Boemont et al.'s (2018) achieved some rich results concerning noise WTP from the same transportation scenario. Now, we consider that the main contributions of this paper to literature are connected to the description and valuation of transport nuisances in general, and in the Pyrenees region, in particular, using a Contingent Valuation Survey (CVS) and a Real-Incentivized Experiment Survey (RIES). These methodologies try to elicit the WTP, from a hypothetical and real points of views, of a group of surveyed people in Navarre (Spain) concerning their perceptions of the air and noise pollutions caused by road transportation. This conjoint analysis of both nuisances has also allowed us to study the cognitive biases associated to our particular surveys. In particular, we found evidence for:

i. ***Hypothetical bias***. We obtained a clear evidence for the existence of hypothetical bias in most of the scenarios defined in the surveys. Usually, this bias assumes lower WTP values when real incentives are present in the scenario. In particular, we found a mean bias ratio, i.e. the ratio for the hypothetical valuation divided by the real one, of 2.52 and 2.66 for noise and air pollution, respectively. This is in line with the meta-analysis described in Foster and Burrows (2017).

ii. ***Correlated CVS and RIES results.*** Similarly, we found correlation effects between the results of the CVS and RIES parts in most of the scenarios. Actually, for each monetary unit of increase in the Hypothetical WTP for pollution will produce, on average, an increase of 0.76 unit of subsequent Hypothetical WTP for noise.

iii. ***Sequencing bias***. Likewise, we attained some evidence of the influence of the order in the surveys parts to reach different results in our analysis: hypothetical settings versus real settings, either considering the full samples or truncated samples (zero-protesters or zero-values removed). This is, by far, the most important bias which has been revealed in our study.

The economic valuation of external costs of transports is deeply rooted in the microeconomic representation of welfare changes at the individual level. This framework assume individuals are selfish and perfectly rational, assumptions that, to say the least, had been challenged by a great area of works belonging to behavioural economics. In the field of transportation, Garcia-Serra et al. (2015) underlined important studies that focused on the implications of behavioural biases for designing policies, but none tackled in detail with environmental policy issues related to transport. Garcia-Serra et al (2015) tried to fill this gap by checking the most important biases that are to be considered when environmental transport policies are to be designed. However, the biases that are considered, focus its interest on traveller behavior (route choices or mode choices), which is a different topic from the one here presented, where local citizens have to valuate transport nuisances.

The recent handbook of external costs 2019 (EC 2019) estimated that the total external cost (excluding congestion costs) of transportation activities in EU28 amounts 716 billion euros in 2016, that is 4.8% of EU28 GDP, 83% coming from road transport. Noise and Air Pollution costs represent respectively 7% and 14% of this total amount, which, if only road mode were considered, amounts to roughly 125 billion euros of road costs for noise and air pollution.

In 2011, the White Paper on European Transport Policy stated that it would proceed with the internalization of external costs for all transport modes (EC, 2011), and to date, at most only road transport had been addressed. Santos (2017), studying 22 countries showed that internalization is far from being achieved, under-taxation of road transport being the most frequent case. But as Parry et al. (2014) underlined, the exercise of "getting prices right" involves a multitude of caveats, especially when legitimate questions about data reliability and methods remain when quantifying environmental damage. We hope that our paper, by challenging contingent valuation method for assessing noise and pollution environmental damages, brings some clarification in the debate of economic valuation of transportation externalities.

Acknowledgements

We want to thank Elven Priour for programming the Z-TREE software that was used in the survey. This work has been partially supported by the financial support of the Erasmus+ Program (2019-I-ES01-KA103-062602), the Spanish Ministry of Economy and Competitiveness (TRA2013-48180-C3-P) and the Spanish Ministry of Science, Innovation, and Universities (RED2018-102642-T). We also want to acknowledge the support received from the CAN Foundation in Navarre, Spain (Grant ID 903616 100010434 under the agreement LCF/PR/PR15/51100007).

References

Abeler, J., Nosenzo, D., (2015). Self-selection into laboratory experiments: Pro-social motives versus monetary incentives. *Experimental Economics*, 18(2), 195-214. Doi:10.1007/s10683-014-9397-9

Aldabe, J., Elustondo, D., Parra, A., Foan, L., Simon, V., Santamaría, J. M., & Santamaría, C. (2012). Polycyclic Aromatic Hydrocarbons (PAHs) Sampled in Aerosol Phase at Different Sites of the Western Pyrenees In Navarra (Spain). *Environmental Engineering & Management Journal (EEMJ)*, 11(6) 1049-1058.

Aldabe, J., Elustondo, D., Santamaría, C., Lasheras, E., Pandolfi, M., Alastuey, A., X. Querol & Santamaría, J. M. (2011). Chemical characterisation and source apportionment of PM2. 5 and PM10 at rural, urban and traffic sites in Navarra (North of Spain). *Atmospheric Research*, 102(1-2), 191-205.

Bateman I., Munro A., Rhodes B., Starmer C. and R. Sugden (2008). Anchoring and yea-saying with private goods: An experiment. in List J. (ed), U*sing Experimental Methods in Environmental and Resource Economics*, Edward Elgar. Chetelham, UK.

Braun, A., Schmeiser, H., and Schreiber, F. (2016). On consumer preferences and the willingness to pay for term life insurance. *European Journal of Operational Research*, 253(3), 761-776.

Burrows J., Newman R., Genser J. and J. Plewes (2017). Do contingent valuation estimates of willingness to pay for non-use environmental goods pass the scope test with adequacy? A review of the evidence from empirical studies in the literature. In McFadden D. & K. Train (ed). *Contingent Valuation of Environmental Goods – A comprehensive critique*. Edward Elgar. Chetelham, UK

Carson R.T. and R.C. Mitchell (1995). Sequencing and Nesting in Contingent Valuation Surveys. *Journal of Environmental Economics and Management*, 28, 155-173.

Carson R.T. and W.M. Hanemann (2005). Contingent Valuation. In K.G. Mler, J. R. Vincent (eds) *Handbook of Environmental Economics*, vol.2. Ch 17 pp. 821-936. Elsevier.

Carson, R.T. & Mitchell, R.C. & Hanemann, W.M. & Kopp, R.J. & Presser, S. & Ruud, P.A. (1992). A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill, MPRA Paper 6984, University Library of Munich, Germany.

Carson, R. T., and Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics* 37:181–210.

Champenois V., Chanel O. and K. Makhloufi (2018). Reducing the anchoring bias in multiple question CV surveys. *Journal of Choice Modelling*, 28, 1-9.

Cummings, R.G., Elliot, S., Harrison, G.W., Murphy, J. (1997). Are hypothetical referenda incentive compatible? *Journal of Political Economy* 105, 609–621.

Cummings, R.G., Harrison, G.W., Rutström, E.E. (1995). Homegrown values and hypothetical surveys: is the dichotomous choice approach incentive-compatible? *American Economic Review* 85, 260–266.

Ciriacy-Wantrup S. V. (1947). Capital returns from soil-conservation practices. *Journal Farm Economics* 29: 1181-96.

Day B., Bateman I., Carson R., Dupont D., Louviere J., Morimoto S, Scarpa R. and P. Wang (2012). Ordering effects and choice set awareness in repeat-response stated preference studies. *Journal of Environmental Economics and Management*, 63, 73-91.

Denant-Boemont L. and S. Hammiche (2019). Economic Measurement of Environmental Costs for Transportation Activity. Ch. 6, in Faulin, J., Grassman S., Juan A. and P. Hirsch (eds), *Sustainable Transportation and Smart Logistics – Decision Making Models and Solutions*, Elsevier, 154-66. Cambridge, MA, USA.

Denant-Boemont L., Faulin J., Hammiche S. and A. Serrano (2018). Managing transportation externalities in the Pyrenees region: Measuring the willingness-to-pay for road freight noise reduction using an experimental auction mechanism. *Journal of Cleaner Production*, 202, 631-41.

Diamond P.A. and J.A. Hausman (1994). Contingent Valuation: Is Some Number better than no Number? The *Journal of Economic Perspectives*, 8(4).45-64.

Dreze, J. and Stern, N., 1987. *The theory of cost-benefit analysis*. In: Handbook of Public Economics. Elsevier, pp. 909-989.

Dupont D.P. (2003). CVM Embedding Effects When There Are Active, Potentially Active and Passive Users of Environmental Goods. *Environmental and Resource Economics*, 25, 319-41.

Eddington, R., 2006. The Eddington Transport Study-The Case for Action. (Sir Rod Eddington‘s Advice to UK Government). Retrieved from. http://collections.europarchive.org/tna/20070129122531/http://www.hm-treasury.gov.uk/media/39A/41/eddington_execsum11206.pdf

European Commission (2011). White Paper “Roadmap to a Single European Transport Area – Towards a Competitive and Resource-Efficient Transport System”. COM (2011) 144 final, 28th, March.

European Commission (2019). Handbook on the External Costs of Transport. CE Delft, January 2019, https://ec.europa.eu/transport/sites/transport/files/studies/internalisation-handbook-isbn-978-92-79-96917-1.pdf. Last access: September 2020.

Faulin, J., Grasman, S., Juan, A., & Hirsch, P. (Eds.). (2019). *Sustainable Transportation and Smart Logistics: Decision-Making Models and Solutions*. Elsevier. Cambridge, MA.USA.

Feltovich N. (2003). Nonparametric Tests of Differences in Medians: Comparison of the Wilcoxon–Mann–Whitney and Robust Rank-Order Tests. *Experimental Economics*. 6, 273-97.

Fifer S., Rose, J. and S. Greaves (2014). Hypothetical bias in Stated Choice Experiments: Is it a problem? And if so, how do we deal with it? *Transportation Research Part A*, 61, 164-77.

Foster H. and J. Burrows (2017). Hypothetical Bias: A New Meta-Analysis. McFadden D. & K. Train (eds). *Contingent Valuation of Environmental Goods – A comprehensive critique*. Ch.10. Edward Elgar. Chetelham, UK. DOI: https://doi.org/10.4337/9781786434692.

Garcia-Sierra M., van den Bergh, J. and C. Miralles-Guasch (2015). Behavioural economics, travel behaviour and environmental-transport policy, *Transportation Research Part D*, 41, 288-305.

Halvorsen B. (1996). Ordering Effects in Contingent Valuation Surveys. *Environmental and Resource Economics*, 8, 485-99.

Hanemann, W.M. (1994). Valuing the environment through contingent valuation. *Journal of Economic Perspectives* 8 (4), 19–43.

Horowitz J. K. (2008). The use of a real-money experiment in a stated-preference Survey, in List J. (ed), U*sing Experimental Methods in Environmental and Resource Economics*, Edward Elgar. Chetelham, UK.

Horowitz J.K., McConnell K.E. and J. Quiggin (1999). A test of competing explanations of compensation demanded. *Economic Inquiry*, 37(4), 637-46.

Istamto, T., Houthuijs, D., Lebret, E. (2014a). Multi-country willingness to pay study on road-traffic environmental health effects: are people willing and able to provide a number? *Environmental Health*, 13(1), 13-35. Doi:10.1186/1476-069x-13-35

Istamto, T., Houthuijs, D., Lebret, E. (2014b). Willingness to pay to avoid health risks from road-traffic-related air pollution and noise across five countries. *Science of the Total Environment*, 497: 420-429. doi:10.1016/j.scitotenv.2014.07.110

Johnston, R. J. et al. (2017) Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2), pp. 319-405.

Kahneman D. and I. Ritov (1994). Determinants of Stated Willingness to Pay for Public Goods: A Study in the Headline Method. *Journal of Risk and Uncertainty*, 9, 5-38.

Kahneman, D., and J.L. Knetsch. (1992). Valuing Public Goods: The Purchase of Moral Satisfaction. *Journal of Environmental Economics and Management* 22(1): 57–70.

Krcal, O., Peer S., Stanek R. and B. Karlinova (2019). Real consequences matter: Why hypothetical biases in the valuation of time persist even in controlled lab experiments. *Economics of Transportation*, 20, 127-138.

Lera-López, F., Faulin, J. Sánchez, M. (2013). Willingness to pay to reduce environmental impacts from road transportation: a case study from the Spanish Pyrenees. *Journal of Applied Operational Research*, 5(4), 135-152.

Lera-López, F., Sánchez, M., Faulin, J., Cacciolatti, L. (2014). Rural environment stakeholders and policy making: Willingness to pay to reduce road transportation pollution impact in the Western Pyrenees. *Transportation Research Part D*, 32, 129-142. doi:10.1016/j.trd.2014.07.003

McFadden D. & K. Train (2017). *Contingent Valuation of Environmental Goods – A comprehensive critique*. Edward Elgar. Chetelham, UK

Messer K.D., Poe G.L., Rondeau D., Schulze W.D., and C.A. Vossler (2010). Social preferences and voting: An exploration using a novel preference revealing mechanism. *Journal of Public Economics*, 94, 308-17.

Mitchell, R. C. and R. T. Carson (1988). *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington: Resources for the Future.

Page L. and K. Page (2010). Last shall be first: A field study of biases in sequential performance evaluation on the Idol series. Journal of Economic Behavior and Organization, 73, 186-198.

Payne J.W., Schkade D.A. and W.H. Desvouges (2000). Valuation of Multiple Environmental Programs. *Journal of Risk and Uncertainty*, 21(1), 95-115.

Parry, I.W.H., Heine, D., Lis, E., Li, S. (2014). Getting Prices Right: From Principle to Practice. International Monetary Fund, Washington D.C. [Spreadsheets available from ⟨http://www.imf.org⟩/environment⟩].

OECD (2018). Contingent Valuation. In Cost-Benefit Analysis and the Environment – Further Developments and Policy Use. Ch. 4. OECD Publishing.

Poe G.L. (2016). Behavioral Anomalies in Contingent Values and Actual Choices. *Agricultural and Resource Economics Review*, 45(2), 246-69.

Randall, A., Hoehn, J.P. (1996). Embedding in market demand systems. *Journal of Environmental Economics and Management* 30, 369–380.

Rivas, E., Santiago, J. L., Lechón, Y., Martín, F., Ariño, A., Pons, J. J., & Santamaría, J. M. (2019). CFD modelling of air quality in Pamplona City (Spain): Assessment, stations spatial representativeness and health impacts valuation. *Science of the Total Environment*, 649, 1362-1380.

Sawik, B., Faulin, J., Perez-Bernabeu, E. (2017): Multi-Criteria Optimization for Fleet Size with Environmental Aspects. *Transportation Research Procedia*, 27, pp. 61-68.

Spanish-French Observatory of Pyrenees Traffic (2018) Document#8 May 2018. Spanish Government and French Government (in Spanish) https://www.mitma.gob.es/recursos_mfom/informe_otp8.pdf. Last access: September 2020.

Tolley, G.S. and Randall, A., Blomquist, G.C., Fabian, R.G., Fishelson, G., Frankel, A., Hoehn, J.P., Krumm, R., Mensah, E. (1986). Establishing and valuing the effects of improved visibility in the Eastern United States. Report to the Environmental Protection Agency No. Cooperative Agreement 807768-01-0, University of Chicago, Chicago.

Santos, G., (2017). Road-fuel taxes in Europe: do they internalize road transport externalities? *Transport Policy* 53, 120-134.

Vickerman R. (2007). Cost-benefit analysis and large-scale infrastructure projects: state of the art and challenges. *Environment and Planning B: Planning and Design*, vol.34, 598-610.

Whitehead J.C. and G.C. Blomquist (2006). The Use of Contingent Valuation in Cost-Benefit Analysis. Chapter in Alberini A.A. and Kahn J.R. (eds) Handbook on Contingent Valuation, Edward Elgar. Chetelham, UK

Appendix A: Survey instructions (translated from Spanish, not to be published, available on website)

MODEL H1

upna Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

ROAD TRAFFIC NOISE EVALUATION SURVEY-EXPERIMENT

Instructions

Welcome to this survey-experiment. The study you are going to take part has the only purpose the academic research and its results will be published in scientific journals. Therefore, no commercial exploitation of results will be done. The answers you are going to give will be anonymous, which means that nobody, including the researchers, will be able to know who answered this questionnaire

A. Noise Analysis

(FOR ALL PARTICIPANTS)
Public University of Navarra has estimated the level of noise to which you are exposed living as you do in the vicinity of the AP-7/N-II/C-17/C-16/C-14 roads at 70 decibels, which is equivalent to the noise made by a heavy truck or a vacuum cleaner. This recording will give you an idea of how loud this is [**HIGH NOISE RECORDING**]

Q.1. Compared with the recording you have just heard, the level of noise you hear from your home is…

a.	…much louder	1
b.	…a bit louder	2
c.	…the same / very similar	3
d.	…a bit softer	4
e.	… much softer	5

The Navarrese Government is considering measures such as the installation of acoustic panels for which they might ask all the citizens to pay a compulsory tax, similar to the one paid for garbage collection. Imagine that this measure would reduce the noise level from 70 to 50 decibels (like the noise made by a washing machine). You will now hear a brief recording of the current noise level followed by one of the reduced level **[COMBINED RECORDING - FRAGMENT OF HIGH NOISE LEVEL AND SHIFT TO A LOWER NOISE LEVEL]**

1) Noise Hypothetical treatment

Q.2. How much would you be willing to pay per year and household for a period of 5 years to achieve a **40%** reduction in the noise level affecting your home? Keep in mind that the money would be taken from your household budget, thereby limiting other payments.

Q.3. (if answer Q.2.=0) Given that you are unwilling to pay any sum at all, please indicate whether you agree (A) or disagree (D) with the following statements:

		A	**D**
a.	I am not aware of any traffic noise at home	1	2
b.	I cannot afford to pay because my income is too low	1	2
c.	I do not think my health is affected by traffic noise	1	2
d.	I would pay more for more effective traffic noise abatement	1	2
e.	The traffic noise is not my fault. People who cause the noise are who should pay for it.	1	2
f.	The taxes I am paying already should cover it	1	2
g.	Other reasons (please specify) _____	1	2

upna Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

ROAD TRAFFIC NOISE EVALUATION SURVEY-EXPERIMENT

2) Noise real treatment

With regard to noise, the two experiment collaborative entities (Red NELS and GreenPeace) perform activities related to noise abatement that later have positive implications for the citizens of Navarra. For instance:

1) Red NELS supports local entities to the implementation of projects of Local Action Plans (Agendas Locales 21), many of them related to mitigation, control, and awareness of noise reduction.
2) GreenPeace participates in noise environmental importance awareness emitted by vehicles, making goods distribution companies to design protocols focused on noise abatement.

We will provide you 60 euros which you could invest part of them in fight against noise, giving us the authorization to transfer a portion of this amount to one of the associations you choose. You will keep the remaining money through a bank transfer. In addition, the sum to be eventually received by the chosen associations will be determined as follows:

a) You are located in a 25 participants group. Each of these participants will indicate through a computer how much money of 60 euros will allocate to one of the above associations (let us call this individual amount x, in euros).

b) Anonymously, we will collect the 25 payment proposals. Later, the computer will sort them from the lowest to the greatest amount. Let us call xMin the lowest proposal and xMax the greatest one. You will not know the money proposed by the other participants, just as the other participants will not know how much you propose.

c) The computer will select randomly a number between 0 and 60 euros. All values between 0 and 60 have the same probability. Let us call Y this random number.

d) If most of the x values proposed by participants are greater than Y (for example, in a 25 people group, if 13 propose an amount x greater than or equal to Y), then the association that had been chosen will receive the amount of 25*Y. At the end of the experiment, each participant will receive a individual amount of (60-Y) euros.

e) If most of the x values proposed by the participants are lower than Y, then the selected association does not receive any money and each participant will receive the amount of 60 euros.

An example would be useful here. Suppose that participant 1 proposes to pay 5 euros to one of the associations, that participant 2 proposes to pay 10 euros, (etc.) and the participant 25 proposes to pay 55 euros. In this situation, the minimum would be 5 and the maximum 55. The computer will choose a random number between 0 and 60 euros.

i. Suppose the random number obtained is 25. If more than the 50% of the participants are willing to pay 25 euros or more, then the association would receive 25 * 25 euros = 625 euros. In this case, at the end, each participant would actually obtain 60 - 25 = 35 euros.

ii. Suppose the random number obtained is 35. If more than 50% of the participants choose to pay an amount lower than 35 euros, then the selected association would receive 0 euros and each participant would receive 60 -0 = 60 euros.

MODEL Q

upna Universidad Pública de Navarra Nafarroako Unibertsitate Publikoa

ROAD TRAFFIC NOISE EVALUATION SURVEY-EXPERIMENT

C. Socioeconomic questionnaire:

Q.4. Did you know **GreenPeace** before this experiment?

a.	Yes	1
b.	No	2

Q.5. Did you know **Red NELS** before this experiment?

a.	Yes	1
b.	No	2

Q.6. How would rate your **health** over the last **12 months**?

a.	Excellent	1
b.	Good	2
c.	Satisfactory	3
d.	Poor	4
e.	Very poor	5

Q.7. Now, please answer YES or NO to following statements about your **lifestyle habits**.

		YES	NO
a.	I smoke	1	2
b.	I have or have had hearing problems	1	2
c.	I do sport or take a walk at least 3 times as week	1	2
d.	I sleep between 7 and 8 hours a day	1	2
e.	I eat 4 or 5 pieces of fruit and/or vegetables every day	1	2
f.	I use a car for most trips	1	2

Q.8. Using the following scale, would you please indicate your degree of **concern** regarding the level of **noise** in the area where you live?

a.	Not at all concerned	1
b.	Slightly concerned	2
c.	Moderately concerned	3
d.	Very concerned	4
e.	Extremely concerned	5

Q.9. Using the following scale, would you please indicate your degree of **concern** regarding the level of **air pollution** in the area where you live?

a.	Not at all concerned	1
b.	Slightly concerned	2
c.	Moderately concerned	3
d.	Very concerned	4
e.	Extremely concerned	5

Q.10. How many people from each of the following groups **make up your household**?

a.	Children (up to 10 years of age)	
b.	Adolescents (11 to 18 years of age)	
c.	Adults	
d.	Pensioners	

Q.11. Are you a full-time student?

Q.12. What is your job? (Answer jointly with question 15)

Q.13. What is the current job of the main earner in your household? (Retired people, please state previous job, widows/widowers please state job of the deceased person).

Self-employed:	**Q.14**	**Q.15**
Self-employed in business, retail, or industry	1	1
Crop farmer, Livestock farmer…….	2	2
Liberal professional……...	3	3
Self-employed tradesman……	4	4
Employee:		
Director, Manager……..	5	5
High-level tenured post, High level technician	6	6
High-ranking civil servant…….	7	7
Middle-ranking civil servant……..	8	8
Mid-level tenured post, Middle manager……	9	9
Commercial agent / representative…….	10	10
Non-degree civil service post……..	11	11
Administrative worker, Office clerk.. …….	12	12
Salesperson, retail sales assistant..... …….	13	13

Skilled worker…….……	14	14
Unskilled worker, Labourer……..	15	15
Apprentice, Junior……..	16	16
No occupation:		
Unemployed……	17	17
Student……	18	-
Retired / Pensioner……	19	-
Homeworker……	20	-

Q.14. Now, taking into account all current sources, where would you say your total average monthly **household income** falls on the following scale?

a. More than 4.000 euros
b. Levels in between (Go to **Q.15**)
c. Less than 1000 euros

Q.15. More specifically, how would you estimate your **household income** on the following scale?

a. 2.801 to 4.000 euros per month
b. 1.701 euros to 2.800 euros per month
c. 1001 euros to 1.700 euros per month
d. Do not know / No answer

Q.16. If you have chosen not to answer questions 16 and 17, are you willing to indicate which **social class** you belong to?

Upper	
Upper-middle	
Middle-middle	
Lower-middle	
Lower	

Q.17. Do you mind telling us your **gender**?

a. Male 1
b. Female 2

Q.18. Do you mind telling us your **age**?

Q.19. And your **education** level?

a. None/primary 1

b.	Lower secondary	2
c.	Upper secondary or vocational	3
d.	University	4
e.	Other (please specify) _____	5

Q.20. What is your sensitivity to **local** economic and political issues?

a.	Not interested at all	1
b.	A little interested	2
c.	Something interested	3
d.	Very interested	4
e.	I do not know	5

Q.21. What is your sensitivity to **global** economic and political issues?

a.	Not interested at all	1
b.	A little interested	2
c.	Something interested	3
d.	Very interested	4
e.	I do not know	5

Q.22. Do you consider yourself sensitive to **ecological issues**?

a.	Yes	1
b.	No	2
c.	I do not know	3