A Work Project presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

Using virtual environments in contingent valuation to elicit willingness to pay for an oil spill combat program

Catarina Maia Nascimento Carvalho

Work project carried out under the supervision of:

Maria Antonieta da Cunha-e-Sá

Supported by Nova SBE Environmental Economics Knowledge Center (MarES team):

Ana Carina Vieira da Silva, Ana Faria and Margarida Ortigão

22-05-2020

Abstract

Oil spill threats off the Portuguese coast are real, though they have not recently occurred. In particular, areas such as the Arrábida Natural Park are highly vulnerable due to development pressures. A contingent valuation survey is used to elicit willingness to pay (WTP) for an oil spill combat program in the area. While from a methodological perspective, (preliminary) results suggest that WTPs are not sensitive to the way information is conveyed when simulating the spill (low vs. high-immersive virtual environment), inferring about the WTPs across heterogeneous agents for reducing the damages caused is valuable information for policy purposes.

Keywords: Oil spills, Virtual Reality, Contingent Valuation, Arrábida Natural Park.

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1. Introduction

The Arrábida Natural Park (ANP), located in the district of Setúbal, Portugal, is a marine and coastal protected area created in 1976 and enlarged in 1998 to include the marine area around Cape Espichel (Figure 1 in the Appendix). The marine park has a total area of 52 km^2, including complex and diverse marine habitats, and more than 1400 marine species (*Instituto de Conservação da Natureza e das Florestas*, 2020). Furthermore, its ecological diversity offers a wide range of ecosystem services.

Ecosystem services (ES) are the direct and indirect contributions of ecosystems to human well-being and they support directly or indirectly our survival and quality of life (Biodiversity Information System for Europe, 2020). Lopes and Videira (2016) followed

a stakeholder collaborative approach to identify different ES provided in the ANP. They identified different types of ES related to the marine area, namely provisioning services such as “handicraft” (scales, shells, fossils) and “fish”, regulation services such as “coastal zone protection”, support services such as “habitat provision” and “Atlantic marine biodiversity”, and cultural services such as “landscape”, “beaches” and “nature sports”. ES are diverse and whilst some of them are market goods (e.g. fishing), many others, mainly regulation and cultural ES, are nonmarketed due to their public good features. Therefore, either prices do not exist or do not reflect the social cost of economic activity leading to overuse and degradation. In this case, nonmarket valuation techniques are required. The appropriate technique depends on the type of value to be assessed. Since the ANP is a protected area, it is likely that individuals assign to the area not only use value, but also non-use value (value assigned to goods/services by an individual who has never or expects never to make use of it), for instance, related to the simple current existence of certain species or areas or their conservation for future generations’ bequest.

Should an environmental disaster, such as an oil spill, occur in the ANP, the value of the ecosystem services provided there could be severely impacted. Luckily, Portugal has not been severely affected by major oil spills, with some exceptions registered in the north of Portugal, following the Jacob Maersk (1975) in the port of Leixões, and the Prestige (2002) accidents along the north-western Spanish coast. Yet, oil spill risk in the ANP should not be disregarded, as it is commonly exposed to illegal oil disposals. This risk is aggravated by its proximity to important ports in the country, namely, those of Setúbal, Sines (main entrance of fossil fuels in Portugal) and Lisbon, and the growth of maritime and port activities. This risk was materialized in 2006, when the trawler “Mar Salgado” sank near the port of Setúbal, resulting in an oil spill that ended up being contained.

The present study uses data collected through a contingent valuation (CV) survey mixed with a low-immersive or a high-immersive virtual reality experiment, conceived and designed by Nova SBE Environmental Economics Knowledge Center, in partnership with the Faculty of Psychology of Universidade Lusófona. The subject of the valuation experiment is an oil spill combat program in Arrábida Natural Park.

This thesis aims to understand the potential effect of the level of immersion of the visualization experience in the willingness to pay for an oil spill combat program in ANP.

This paper is organized as follows: section 2 presents a literature review of the use of contingent valuation in oil spill-related studies and the use of virtual reality in experimental economics and environmental studies; section 3 covers the survey and experimental design; section 4 describes data collection, sample descriptive statistics and variables to be considered in the model; section 5 explains the econometric methodology employed; section 6 discusses the results; and section 7 presents the main conclusions of the study, focusing on its limitations and recommendations for further research.

2. Literature Review

As mentioned before, stated-preference valuation methods, such as choice experiments (CE) or contingent valuation (CV) are often required when eliciting preferences on environmental goods and services. In CE, individuals are presented with a choice set that reflects a trade-off between some sort of price and specific attributes of a good. In the context of oil spills, CE have been used to elicit preferences for different attributes of oil spill combat strategies in the German North sea coast (Liu et al., 2009) or to elicit willingness to accept for levels of risk and potential impacts of a spill in the Amazon river (Casey et al., 2008). Note that Casey et al. (2008) have evaluated willingness to accept (WTA) for losses rather than willingness to pay (WTP) for gains, based on the assumption

that individuals have the right to a clean environment. However, most studies employ WTP as it produces more conservative estimates of the value assigned to a good. As this study does not focus on valuing or understanding the role of specific attributes of the oil spill combat program, but on the value of oil spill damages in ANP through the valuation of a credible combat program as a whole and its determinants, the contingent valuation method is employed.

CV consists in directly eliciting preferences through carefully designed and administered sample surveys (Johnston et al, 2017). The surveys contain a description of the good and describe a hypothetical market, so responses are contingent on the setting. CV has been used to value damages from oil spills in different geographical areas. For instance, it has been used by Carson et al. (2003) to assess the environmental damages of the 1989 Exxon Valdez oil spill (Alaska); by van Biervliet et al. (2005) to assess WTP to prevent an oil spill in the Belgian Coast; by Ahtianien (2007) to elicit WTP for improvements in the oil spill response capacity and reduction of harm of potential oil spills in the Gulf of Finland; by Loureiro et al. (2009) to assess the environmental damages of the 2002 Prestige oil spill (Spain) in a use value versus a non-use value perspective; by Navrud et al. (2017) to value ecosystem services losses due to oil spills in the Norwegian Coast; or by Bishop et al. (2017) to assess damages from the 2010 BP Deepwater Horizon oil spill in the Gulf of Mexico. However, no published studies were found valuing damages from potential oil spills on the Portuguese coast. Moreover, none of these studies used virtual environments (or virtual reality) to describe the hypothetical setting being assessed.

Virtual reality (VR) experiments simulate real environments, providing contexts in which participants are immersed under the control of the experimenter, being framed field experiments (Innocenti, 2017). These experiments can be split into two main categories according to the level of immersion induced: high-immersive virtual environments

(HIVE) and low-immersive virtual environments (LIVE). In HIVE, movements in the virtual environment are (partially) controlled through body motion and subjects use head-mounted displays to visualize the environment, being barely aware of their surroundings. In LIVE, there is no motion capture and the visualization experience is computer screen-based, so there are numerous signals of the physical world.

High-immersive virtual reality "is a promising new tool in the experimental economics toolkit" (Mol, 2019, p. 155). This technology allows for experimental control (context is entirely embodied), experimental realism, automatic logging of responses (objective measures from detailed movement tracking), visualization of complex questions and "impossible"/unethical experiments. However, such technology introduces some challenges as well, which need to be considered by practitioners. Among these, Mol (2019) points out software and hardware costs, specialist skills, simulator sickness, different levels of familiarity with the technology across participants, naturalistic avatars and lab time (experiments tend to take more time due to equipment constraints).

VR has been employed in environmental willingness to pay (WTP) elicitation studies and it has been compared to other alternative ways of conveying information. Iftekhar et al. (2019) conducted a choice experiment aiming at eliciting WTP for a Singapore water program and compared the effect of presenting information verbally and with photos against virtual reality, finding that VR enhances respondents' WTP. Bateman et al. (2009) has used VR in a split-sample choice experiment to elicit preferences for changes in coastal land use, finding that using VR rather than standard presentation of attributes (tables of numerical and categorical data) increases precision and reduces the WTP/WTA gap. In non-environmental research, VR has been found to increase the realism of consumer choices compared to 2D images (van Herpen et al., 2016) and to increase understanding and choice consistency when compared to text and video (Farooq et al.,

2018). However, no literature has been found linking the degree of immersion of virtual reality (HIVE versus LIVE) to willingness to pay in an environmental context, particularly in an oil spill setting.

3. Survey and Experimental Design

The goals of the chosen design are to provide an appropriate stated preferences setup, to ensure content validity, to make the survey understandable, to make the scenario and the choice mechanism realistic, so that respondents take it seriously, avoiding hypothetical bias, to produce conservative estimates and to avoid selectivity bias.

The survey was administered to respondents in Portuguese and was designed following current best practices (Johnston et al., 2017). Moreover, to obtain the final version of the survey, three focus groups were conducted in a first stage, followed by a pilot version of the study (survey and HIVE experiment) with 30 participants.

Initially, respondents were asked to prioritize 3 environmental problems out of 9 relevant problems, including oil spills, and which oil spills they recalled. The next section of the questionnaire assessed the link/familiarity of participants with the study area. Participants were first asked if they had ever visited the study area. Then, they were presented with different recreational activities offered in the ANP and asked to score them on a 3-point Likert scale, according to how often they practice them ("never", "punctually", "regularly", "does not know/does not answer"). The same structure was used to ask respondents how often they visited each of the 10 Arrábida beaches.

Next, participants watched a video that provided a complete overview of the setting, containing information regarding oil spills and the combat program being evaluated. The video explained how the severity of impacts of oil spills depend on features of both the spill itself and the coast, also highlighting that these accidents are not frequent along the

Portuguese coast even though the Prestige (2002) and the Jakob Maersk (1975) affected the northern Portuguese coast, whereas the Mar Salgado (2006) and illegal oil discharges have affected the study area. Then, the ANP was described, presenting the risks to which it is exposed. Still during the video visualization, a description of the combat program was presented, and participants were asked about their individual WTPs for this program. The competent authority (*Direção Geral da Autoridade Marítima*, DGAM) was mentioned to be in charge of implementing the program. For realism purposes, the video also mentioned that the implementation of the oil spill program requires a 53,7 million euro investment and that its specific measures include workers' training, oil spill response drills, acquisition of new combat equipment and surveillance reinforcement (use of drones to monitor the area and avoid illegal discharges). Note that the program encompasses not only features of a prevention program, reducing the risk of an event, but also of a program that aims to reduce the damages caused to the area in case of an event.

The payment vehicle is also an important component of contingent valuation studies, as it can generate protest responses (respondents do not reveal their true valuation as a protest against the payment vehicle presented). Respondents were informed that they could voluntarily contribute to the program described when finishing the questionnaire and that their contributions would be collected through a direct debit authorization. This information creates a stronger sense of realism, reducing the possibility of hypothetical bias that may lead to an overstatement of WTP responses (in hypothetical settings participants tend to state values that in real life they would not give). Moreover, they were informed that in case the authorities were not able to collect the minimum amount of aggregate contributions to implement the program (53,7 million euros), contributors would be fully reimbursed. This avoids protest responses against corruption or lack of

trust in authorities/program implementation and reduces free-riding incentives (Poe et al., 2002), common in voluntary contribution mechanisms.

After watching the video, respondents are asked to classify on a 3-point Likert scale how concerned they are with respect to 6 different potential impacts of oil spills in ANP (bathing and sports interdiction, fishing interdiction, damages on animals and habitats, damages to tourism, damages on the landscape and damages on food consumption).

Then, the participants are randomly assigned the LIVE or the HIVE version of the experiment. In the HIVE version, participants use head-mounted displays to interact with the virtual environment through partial body motion (head and body rotation), while clicking on a remote control to simulate walking. In the LIVE version, which is computer-screen based, headphones and a standard videogame control are used to navigate through the virtual environment. A pre-experiment was conducted in both versions, serving as a tutorial to teach participants how to navigate through the virtual environment and to make participants overcome the novelty effect of the VR experiment and focus more on the visualization (and sound) experience itself. The subsequent oil spill simulation was split in 3 stages. In the first stage (90 seconds), participants would find themselves in Creiro beach (one of the most iconic beaches in Arrábida). Subjects could visualize different elements, namely a restaurant, the mountains, a small island (*Pedra da Anicha*), birds flying, the sea and a vessel at mid-distance. In the second scenario (30 seconds), everything was kept except that the vessel had caught on fire. In the final scenario (90 seconds), the impacts of the oil spill were shown – the sea was completely black and there were dying animals (dolphins and birds) at the seashore. Note that, in terms of images and sounds, the pre-experiment and the main experiment were exactly the same in both versions. An illustration of each scenario can be found in Figures 2, 3 and 4 in the Appendix, respectively.

The WTP question is framed as a mandatory open-ended question, since in voluntary contributions, this format has been found to predict actual contributions better than dichotomous choice designs (Poe et al., 2002). Respondents are asked to state a number that represents how much they are willing to voluntarily contribute to the program (in euros) described in the previous video. They were told that they could choose not to contribute. Before answering, respondents were reminded to think about their disposable income and the opportunity cost of the amount spent , to consider that the oil spill could impact the whole ANP area and not only the Creiro beach they visualized, that when ending the survey they would have the opportunity to pay the amount chosen, through a direct debit authorization, and that there is a minimum amount to be raised for the program to be implemented.

Some follow-up questions were included. Participants had to answer to the Portuguese version of the Igroup Presence Questionnaire (IPQ, Vasconcelos-Raposo et al., 2016) to evaluate the sense of presence during the virtual experiment (HIVE or LIVE).

Next, respondents were required to answer follow-up questions on budget constraint, fairness and institutional quality issues. Particularly, they were asked to which degree of certainty they would or would not contribute to the program and their agreement with possible reasons for their stated WTP on a 5-point Likert scale. Possible reasons include: financial possibility to contribute, belief that national actions towards oil spills should be improved, desire to prevent oil spill impacts in the ANP, to prevent risks to the environment or limitations to the recreational use of the area, belief that the costs of the program should be borne by residents in the area, by oil companies or by taxation, belief that the contributions would actually be applied in the presented program and belief in the effectiveness of the program. This aids in identifying what drives possible protest responses.

As the survey could be considered long, the degree of attention of the respondents is monitored at this point using multiple-choice questions about the information provided in the video (study area, payment vehicle and the authority responsible for the implementation of the program). At this point, as no further questions will be raised on the stated WTP amount, respondents are also informed that no contributions will be collected by the end of the survey.

Finally, respondents filled in a socioeconomic and demographic questionnaire, where they were asked about different 6 pro-environmental behaviours (recycling, buying environmentally-friendly products, watching shows about the environment, participating in nature conservation actions, donations to environmental organizations or campaign) on a 5-point scale from "never" to "always" and whether they belonged to an environmental organization, age, residence area, education, employment and earnings. In the end, they were asked to evaluate the difficulty, clarity and interest of the whole questionnaire.

4. Data

4.1. Data Collection

The data used in this study consists of survey answers and it was collected in-person, from the end of November 2019 to the beginning of March 2020 in two locations: Nova SBE Campus, in Carcavelos, and Universidade Lusófona Campus, in Campo Grande, Lisbon. Participants were recruited through advertisements via several Nova SBE channels (e-mail lists and social networks), flyers distributed randomly close to both campuses and through Facebook ads. The advertisements did not give away any details on the subject of the study to avoid self-selection bias (disproportional number of sign-ups from environmentally-concerned individuals), they only provided information on how to sign-up, where the survey was conducted, duration of the experiment (around 40 minutes) and the monetary incentive amount (10€ supermarket voucher).

When signing-up, participants would fill in a small questionnaire with a few socioeconomic questions (age, whether or not the person earns income and/or participates in the household budget decisions, employment status) and contact information (e-mail and/or phone number). The responses would then be filtered to select participants in order to guarantee sample heterogeneity and representativeness. Participants were required to speak Portuguese, to have lived in Portugal for at least one year, to be over 18 and to have income of their own or to participate in their household's budget management.

Differently from the pilot version of the study, where questions were asked by the interviewer, participants answered the CV survey on a computer, facilitating the understanding of the questions, with minimum interaction with the experiment facilitator that was in a separate room or in the opposite side of the room for most of the experiment to reduce the chance of interviewer bias. A letter of support signed by DGAM, stating their agreement and support, was also provided as well as a direct debit collection authorization form which participants would fill in at the end of the survey with their proposed contribution. This information was made available to increase the credibility of the study, and, therefore, for policy consequentiality (Zawojska et al., 2019)

4.2. Sample Descriptive Statistics

The descriptive statistics can be found in Table 1, showing heterogeneity in the whole sample and homogeneity between the participants of the HIVE and the LIVE version. The study has a total of 127 participants so far, from which 63 participants were assigned the HIVE version and 64 the LIVE version.

Table 1: Sample descriptive statistics

	Sample (n=127)	**HIVE (n=63)**	**LIVE (n=64)**	**t-test (p-value)**
Member of an Environmental Organization	17 (13.4%)	6 (9.5%)	11 (17.2%)	0.2070
Age (average)	33.35	33.49	33.22	0.8964
Gender (female)	78 (61.4%)	38 (60.3%)	40 (62.5%)	0.8025
Income (euros)				
No income of their own	25 (19.7%)	10 (15.9%)	15 (23.4%)	0.2869
[0;600]	18 (14.2%)	9 (14.3%)	9 (14.1%)	0.9715
]600;1000]	27 (21.3%)	17 (27%)	10 (15.6%)	0.1203
]1000;2000]	44 (34.6%)	19 (30.2%)	25 (39.1%)	0.2952
]2000;1500]	13 (10.2%)	8 (12.7%)	5 (7.8%)	0.3686
Professional Status				
Student	35 (27.6%)	17 (27.0%)	18 (28.1%)	0.8867
Employed	77 (54.3%)	39 (61.9%)	38 (59.4%)	0.7726
Unemployed	10 (7.9%)	5 (7.9%)	5 (7.8%)	0.9795
Retired	5 (3.9%)	2 (3.2%)	3 (4.7%)	0.6637
Education				
1^{st} to 9^{th} grade	2 (1.6%)	0	2 (3.1%)	0.1589
High School or equivalent (12^{th} grade)	15 (11.8%)	8 (12.7%)	7 (10.9%)	0.7609
Bachelor's degree or equivalent	56 (44.1%)	23 (36.5%)	33 (51.6%)	0.8888
Master's degree or above	54 (42.5%)	32 (50.8%)	22 (34.4%)	0.0622

A similar average age and a similar share of female participants are observed in both versions of the experiment. Regarding the share of participants that are members of an environmental organization of any sort, there is no reason to believe it is an excessive share, given the way the study was advertised. Despite the effort put into the prescreening phase to avoid a large concentration of observations from students, almost a third of the participants are students. Education levels seem evenly distributed among the two treatments. The larger differences observed are between bachelor's degree and master's degree level, but no big difference is found when considering higher education as a whole. Income-levels show some differences when comparing the subsamples of both types of the experiment, but it does not seem troubling since all income-levels are represented in both subsamples.

To confirm homogeneity between the HIVE and the LIVE samples, t-tests on the null hypothesis of equal means were performed on each variable above and none revealed a p-value below 5%, leading to the non-rejection of the hypothesis that the means are different across samples and, therefore, reassuring homogeneity.

4.3. Variables Description

Aiming at studying the determinants of WTP for the program, 7 explanatory variables are considered to model *WTP*, the dependent variable that corresponds to the amount bided (in euros) by each participant. Given the sample size, it is important to avoid overfitting by including too many variables. Therefore, the final set of variables presented in Table 3 in the Appendix were chosen to be included in the model because they were considered key to explain the respondents´ willingness to pay, according to the literature, to answer the research question and/or due to their statistical significance, ensuring simultaneously good model fitness.

Age (age of the individual in years) and *Gender* (dummy variable coded as 1 if the individual is a female and 0 otherwise) were included as they are the main individual characteristics, typically included as demographic control variables, and due to their significant results. WTP is expected to be higher in women and in younger individuals, according to the literature.

Since deciding on willingness to pay is an economic and financial decision, the inclusion of *income* (monthly disposable income, in euros) is typically included in similar studies. It is expected to show a positive significant effect on WTP. *Income* is modelled as the center of each income class to reflect the different class sizes used in the questionnaire.

Protest responses are usually associated with moral and ethical considerations. In environmental studies, it is a common belief that some individuals might state zero WTP

as a protest, since environmental protection is perceived as a government's responsibility (Alvarez-Farizo, 1999). In this sense, the variable *unfair* is a dummy variable that incorporates fairness motives as a proxy to protest responses that might explain the decision to bid zero or a positive value. If *unfair* equals 1, it means that the respondent agrees with at least one of the fairness-related questions: residents should pay for the program, oil companies should pay for it or the costs of the program should be paid through taxation. In sum, they agree it is unfair to have to bear the costs of the program. If the variable shows to be significant, it would indicate that a significant portion of zero WTP answers does not reveal the individual's true valuation of the good, but rather some form of protest related to fairness considerations.

Contributions (frequency of contributions to environmental NGOs, in a 5-point Likert-scale) is used as a proxy to pro-environmental attitudes. Other pro-environmental behaviours could be included, but they revealed no statistically significant results. A higher level of contributions is expected to positively impact WTP, not only as it signals a pro-environmental attitude, but also as it indicates that the individual is more receptive to the form of contribution presented (voluntary contribution).

The variable *Beaches* (frequency of visits to ANP beaches as the sum of ten 3-point Likert items related to beach use) is used as a proxy to familiarity with the study area. Subjects that go to Arrábida more often will enjoy the ecosystem services offered in the area to a greater extent and will be more aware of the damages of a potential oil spill in the area. Participants who frequently visit ANP's beaches are hence expected to state higher WTP.

HIVE is the main variable of interest as it accounts for the treatment effect. In accordance to the previously described literature findings, being exposed to the HIVE version rather than the LIVE version is expected to increase willingness to pay as it increases the sense of realism of the information regarding the potential impacts of an oil spill in ANP.

5. Methodology

Since the valuation question is framed as an open-ended question, a concentration of answers is expected around the zero-lower bound, in case individuals do not want to make a voluntary contribution to the program (in case they do, they can choose any strictly positive value).

The question that arises relates to the origin of the zero answers. There are three main types of zeros. The first type of zeros and the most common belief in economic theory is considered a pure corner solution in which an individual maximizes their utility subject to a budget constraint and the optimal solution is to contribute zero. It is implicit that the oil spill combat program is a good that will increase the individual's utility and a change in circumstances (individual's budget constraint), whether it be income or price of the good (the latter not applicable in this setting), would eventually lead to a different decision. A second reason relates to the possibility that the program described has no positive impact on someone's utility, a reasonable possibility as 7 participants stated they had the financial possibility to contribute, but decided not to, being a matter of preferences rather than a budget constraint problem. The individual is expected to never contribute to the program, no matter what the circumstances (price or income) are. Then, the decision to contribute must come from a change in preferences. Thirdly, the individual might bid zero as a protest response, mainly because they do not know how to value the good, they do not agree with the means of payment or they do not agree they should pay for it, as is common in public goods, not representing their true valuation of the good.

Wooldridge (2010) presents three alternatives to deal with corner solutions. Note that, as discussed by the author, using a model such as OLS would lead to inconsistent parameter estimates as it ignores censoring. The first alternative that seems to be the most common in contingent valuation literature with open-ended questions and that is closely related to

the first type of zeros presented is the standard Tobit model. It seems a straightforward choice; however, heteroskedasticity and nonnormality would lead to inconsistent parameter estimates. Plotting a histogram of WTP (figure 5 in the Appendix) raises suspicions of nonnormality, confirmed by a simple conditional moment test (Table 5 in the Appendix) that leads to the rejection of the null hypothesis of normal errors when a Tobit model is estimated considering the variables described in the previous chapter (Table 4 in the Appendix). Cameron and Trivedi (2009) suggest the use of tobit with lognormal data to overcome this problem, but it raises the obvious concern of how to deal with an undefined logarithm of zero, that can only be solved through some type of data manipulation. The second alternative proposed by Wooldridge (2010) concerns hurdle models, which include the possibility to specify a lognormal hurdle model (LHM), helping to get closer to normality. Hurdle models have the advantage of being more flexible than tobit models as they allow the participation decision (whether to contribute) and the amount decision (how much to contribute, conditional on deciding to contribute) to be modelled by different mechanisms. On the other hand, they impose a conditional independence assumption, meaning that the participation and the amount decisions are independent. Finally, the author proposes the use of an exponential type II tobit (ET2T) model. This model nests the lognormal hurdle model, combining the desirable features: it uses an exponential specification, helping with nonnormality, it allows for flexibility regarding the mechanism behind the participation and the amount decisions and it allows for the errors of both equations to be correlated, as it seems natural to think that unobserved factors might influence both decisions. However, it requires the identification of at least one strong exclusion restriction (variables included in the participation equation but not in the amount equation). Among the three described models, it is also the less commonly found in contingent valuation to model corner solution models, being more

frequently employed to model sample selection and protest responses (the third type of zeros described), under the famous Heckman's model.

The previous three alternatives all have strengths and drawbacks. However, the wish to assess possible different impacts of each variable in the participation and in the amount decisions, respectively, calls for the flexibility of the second and the third alternatives presented. For instance, one could hypothesize that a person that contributes more frequently to environmental causes will contribute to this program but might contribute smaller amounts as they contribute more frequently. The failure to identify a strong exclusion restriction accompanied by a not so large sample size leads to an estimation of very large standard errors of the correlation coefficient between the errors of the participation and the amount equation and to the rejection that this coefficient is statistically different from zero, which might not be true, thus arising from a possible failure of model identification. Therefore, the lognormal hurdle model will be applied, under the assumption of conditional independence. Later, robustness checks considering the two other alternatives will be performed.

In a lognormal hurdle model, the dependent variable (y) is generated as:

$$y = w \cdot y^* = 1[\mathbf{z}'\boldsymbol{\gamma} + \boldsymbol{v} > 0] \cdot \exp(\mathbf{x}'\boldsymbol{\beta} + \boldsymbol{u}), \boldsymbol{u}|\mathbf{x} \sim \text{Normal}(0, \sigma^2) \qquad (1)$$

where w is a binary variable that reflects the decision to contribute (w=1) or not (w=0) to the program and it is fully observed, whereas y* (nonnegative, continuous random variable) is only observed when w=1. The set of explanatory variables included in the participation equation (w) is represented by **z**, whereas the explanatory variables of the amount equation (y*) are represented by the vector **x**. Recall that **z** and **x** may or may not be equal. The conditional independence assumption will imply that $\boldsymbol{u}$ (error term of the amount equation) and $\boldsymbol{v}$ (error term of the participation decision) are independent and

follow a normal distribution with mean zero and variance σ^2. Parameters $\boldsymbol{\gamma}$ (participation equation) and $\boldsymbol{\beta}$ (amount equation) are estimated based on maximum likelihood estimation of the following log-likelihood function:

$$\log L(\theta) = 1[y_i = 0]\log[1 - \boldsymbol{\Phi}(\mathbf{z_i}'\boldsymbol{\gamma})] + 1[y_i > 0]\log[\boldsymbol{\Phi}(\mathbf{z_i}'\boldsymbol{\gamma})] + \\ + 1[y_i > 0]\{\log(\boldsymbol{\phi}[(\log(y_i) - \mathbf{x_i}'\boldsymbol{\beta})/\boldsymbol{\sigma}]) - \log(\boldsymbol{\sigma}) - \log(y_i) \quad (2)$$

where $\boldsymbol{\Phi}$ corresponds to the standard normal cumulative distribution function and $\boldsymbol{\phi}$ to the standard normal density function. In terms of parameter estimation, it is equivalent to estimating a probit of w_i on $\mathbf{x_i}$ to obtain $\boldsymbol{\gamma}$ and estimating an OLS regression of $\log(y_i)$ on $\mathbf{x_i}$, conditional on $y_i > 0$, to obtain $\boldsymbol{\beta}$.

6. Results and Discussion

6.1. Determinants of Willingness to Pay

Table 2 presents the coefficients and standard errors estimated by the lognormal hurdle model. Notice that the dummy variable *unfair* is not included in the amount equation as it is related to zero bids. Its effect is reflected in the participation equation, as explained previously. The model is overall statistically significant at the 1% confidence level (Prob > chi2 is 0.0009). Summary statistics can be found in Table 6 in the Appendix.

Table 2: Results of the lognormal hurdle model

	Participation Equation	**Amount Equation (lnWTP)**
	w	***y****
HIVE	**-0.05796**	**-0.24242**
	(0.24646)	(0.21524)
Age	**-0.04133 *****	**0.01132**
	(0.01220)	(0.01136)
Gender	**0. 57799 ****	**0.11963**
	(0.26790)	(0.238501)
Income	**0.00030 ***	**0.00030 ****
	(0.00017)	(0.00015)
Contributions	**0.53581 *****	**0.11218**
	(0.15951)	(0.10590)
Beaches	**0.00636**	**-0.00351**
	(0.04218)	(0.03815)

Unfair	**-0.58429**	
	(0.46065)	
Constant	**0.37589**	**1.45782 ****
	(0.69426)	(0.59187)
Number of observations	127	
Censored observations	51	
Wald chi2(13)	34.68	
Prob > chi2	0.0009	
Pseudo R-squared	0.0653	
Log Likelihood	-334.615	

Standard errors in parentheses. *, **, and *** denote significance at 10%, 5%, and at 1%, respectively.

Note that despite the fact that the two equations are independent, the marginal effect of each variable on the willingness to pay is not linear and cannot be directly interpreted from the coefficients above. For instance, a variable present in both equations will impact *y* through *w* and *y**, possibly in opposite directions. However, there is no particular interest in evaluating the average marginal effect for the purpose of this paper, as the main interest lies in the significance and direction of the *HIVE* effect and not on its magnitude.

Regarding *age*, growing older seems to significantly decrease the probability of being willing to contribute at the 1% confidence level, whereas being a woman seems to significantly increase the probability to answer a positive bid, at the 5% confidence level. Neither age nor gender proved to be statistically significant in explaining the amount of the contribution to the program, conditional on deciding to contribute.

Individual's income level is not a statistically significant predictor of the participation decision at the standard confidence level of 5%. On the other hand, it is the only variable to present some explanatory power regarding the amount decision, conditional on deciding to participate (significant at the 5% confidence level). Its predicted impact on willingness to pay for the program is positive, in accordance with economic theory.

Fairness concerns are not a relevant determinant of the decision to either contribute or not to the program, meaning that even though the respondent agrees that someone else should

bear the costs of the program (residents of the area, the government or oil spill companies) it does not significantly dictate a zero willingness to pay and individuals would still decide to contribute to the program.

As expected, the more frequently someone monetarily contributes to environmental organizations, the higher the probability they would be willing to pay for the program. This might be a hint that individuals care about environmental protection as a whole, and not particularly about oil spills. It might not have to do with the environmental problem at stake but with a desire for environmental protection of any type. In fact, when presented with a list of 9 environmental problems and asked to choose 3 problems that required priority action in Portugal, only 8 participants (6,3%) chose oil spills as one of them, being the least chosen option among the 9 choice possibilities. On the other hand, the most chosen problem relates to the increase in the frequency and intensity of wildfires, picked by 94 respondents (74%), followed by the accumulation of litter in the ocean and/or land (88 respondents, 69,3%). The frequency of monetary contributions to environmental causes shows no significant effect on the amount equation, refuting the previously stated hypothesis that it could have a negative impact on the value of the contribution.

The frequency of visits to beaches in Arrábida Natural Park does not have any statistically significant impact on the probability of contributing to the program nor on the value of contributions, conditional on deciding to contribute, which indicates that willingness to pay is independent of familiarity and degree of use at the study area. This result can be interpreted as a hint on the relevance of non-use values linked to the area, which is also found in other studies on environmental goods and services (Loureiro et al. 2009)

HIVE accounts for the treatment effect, which has to do with the technology employed to create different levels of immersion in the virtual environment. Being assigned the HIVE

version rather than the LIVE version shows no statistically significant impact on the decision to contribute nor on the amount decision. This result might be explained by the similarity of the visualization experience. The design of the virtual environment was exactly the same, except for the equipment used during the experience, which creates higher immersion in the HIVE version. Hence, the contrast between both versions might not be large enough to produce significant results with the current sample size.

6.2. Robustness Checks

Table 9 in the Appendix presents the results of the ET2T model (Wooldridge, 2010), of the lognormal tobit (LT) approach proposed by Cameron and Trivedi (2009) that replaces the logarithm of zero by a value that undercuts the minimum value of the uncensored variable. Both models are overall statistically significant at the 1% confidence level.

When comparing to the results of the LHM, the ET2T and LT models produce no significant changes. *Unfair* and *beaches* remain statistically insignificant both in the ET2T and the LT models, *gender* loses significance in the LT model but remains significant at the 5% level in the participation equation of the ET2T and *income* remains significant at the 5% level in the LT specification and in the amount equation of the ET2T, gaining significance in the participation equation. In the LT model only *age* and *income* are statistically significant, with negative and positive sign, respectively, as in the LHM. The variable *contributions* loses its significance in the LT model, but keeps its significance in the participation equation and gains significance in the amount equation of the ET2T model. *HIVE*, the main variable of interest, remains statistically insignificant, indicating that the model chosen performs well in answering this paper's research question. The correlation coefficient (ρ) of the ET2T model is not statistically significant, supporting the assumption of conditional independence. However, as mentioned, its result cannot be trusted due to lack of a strong exclusion restriction.

6.3. Assessment of the Survey and Experiment Design Quality

Another way to assess the validity of the results is to evaluate the quality of the survey and the experiment in different dimensions. Participants had the chance to directly evaluate the questionnaire. Regarding clearness of the questions presented, 61.4% of the subjects classified them as clear and 38.6% as very clear. Concerning the degree of difficulty, 52% classified it as easy, 44.9% as very easy and 3.1% as difficult. The questionnaire was also classified as interesting by 51.2% of the participants, as very interesting from 47.2% of them and of little interest by 1.6% of them. The evaluations are very positive from the participant experience point of view.

Even though the experiment takes a long time to complete (average of 40 minutes), participants showed a good level of attention, with 74% of the subjects answering correctly 3 out of 3 questions related to features of the setting described (study area, payment vehicle and responsible authority), 23.6% getting 2 correct answers and the remaining 2.4% answering 1 question correctly.

It is also important to understand if the features of the survey designed to make it as realistic and credible as possible had the desired effect from the participants point of view. In this matter, 90 participants (70.87%) agreed that the program described would be effective in combating the impacts of oil spills and 105 participants (82.68%) agreed that the information collected in this questionnaire would be taken into account by national authorities to decide on the implementation of the oil spill combat program.

Regarding the VR experiment, one way to validate that the technology used in the HIVE version creates in fact a higher level of immersion (HIVE) is to check whether participants of the HIVE version experienced a higher sense of presence in the virtual environment than LIVE participants. The degree of presence can be measured from the answers to the

Igroup Presence Questionnaire (Vasconcelos-Raposo et al., 2016). The answers to the 14 questions are aggregated into a single composite measure of spatial presence, involvement and experienced realism called (global) presence (Table 7 and Equations 3, 4, 5 and 6 in the Appendix). A one-sided t-test on the null hypothesis that presence is equal for the HIVE version and for the LIVE version against the alternative hypothesis that presence is higher for the HIVE version was performed. With a p-value of 1,6% (Table 8 in the Appendix), the null hypothesis is rejected with a 95% confidence level, suggesting that the HIVE version leads in fact to a higher sense of presence.

7. Concluding Remarks

The information collected by this survey and the results obtained are very important in pointing out the limitations of the current design and directions for further research.

The first recommendation that seems vital is to increase sample size. With each observation taking an average of 40 minutes to be collected, implementing the experiment is a resource and time-consuming task, imposing restrictions on the sample size collected, currently aggravated by the Covid-19 outbreak, which interrupted data collection. Despite this limitation, a larger sample size would not only increase the robustness of the results obtained, but it would also allow to estimate, at a national level, the financial and economic viability of a program as described on a cost-benefit perspective, having important policy implications. Though the purpose of this study is more methodological, namely by investigating the consequences of conveying information to respondents in different ways (low vs high immersive virtual environment), the economic valuation of the losses associated to such accidents is an important instrument for environmental damage assessment which can then be integrated into a decision-making tool to be used for policy purposes. The purpose of valuation is to translate the environmental losses into monetary values which can then be used to assist policy makers in assessing policy

choices or trade-offs concerning different management options (*e.g.*, prevention or restoration), also helping spotting priority areas of intervention. Moreover, by identifying the populations affected by the accident and by estimating the corresponding losses in monetary terms, it is also valuable information related to the appropriate amount of compensation to injured parties when liability is at stake. In fact, liability can be an important incentive-based instrument for preventing oil spills and a sustainable approach for restoring coastal resources (Grigalunas et al., 1996).

The current design uses an open-ended valuation question rather than a dichotomous choice, which is the most common design in the literature. The main advantage of a dichotomous choice is that it resembles real life choices, as individuals are used to be presented with a price and deciding whether or not to pay for it. This way of framing the question provides easier and more realistic judgement. Answers to open-ended questions might be more conservative, but it does not mean they are more realistic, as subjects often do not know how to value the good. The answers collected with the current valuation question design are of great importance, as they provide insightful information on possible bids to be presented to the respondents in order to extract maximum WTP responses as accurately as possible. Note that the goal is to find subjects' maximum willingness to pay in order to estimate a demand function. Of course, this would require randomly assigning different bids to the respondents, which would once again require a larger sample size due to subsampling.

Due to the fact that being assigned a high-immersive version of the virtual reality experiment rather than a low-immersive one did not reveal any significant impact on WTP, it is also worth revisiting how the visualization experience and the way the information regarding the effects of an oil spill is conveyed in this contingent valuation survey. The lack of significant results with the current sample and sample size might be

related to the strong similarity between the treatment (HIVE) and the control version (LIVE), as the video and sound components were exactly the same in both versions, with the only difference coming from the technology used to interact with the virtual environment and corresponding levels of immersion. It does not mean that the visualization experience does not have any impact on willingness to pay. Instead, it means that besides increasing sample size, a different control group should be used, more clearly contrasting with the HIVE version, for instance, using video, 2D images or text to illustrate the potential impacts of an oil spill. In fact, there could be more than 2 different treatments, but it would always depend on the sample size, which is somehow constrained, as discussed. On the other hand, it indicates that it might not pay off for researchers to incur in significant equipment costs, by acquiring the necessary expensive equipment to create HIVE experiments. If the visualization experience is enough to elicit WTP, a low-immersive computer-screen based experiment might be enough.

Finally, it would be interesting to compare the effect of the visualization experience when assessing willingness to pay to prevent the impacts of different environmental problems, even though they cannot be fully compared since contingent valuation is fully dependent on the setting. For instance, maybe the impact of the visualization experience becomes less and less significant as the problems are more present in subjects' everyday lives and memories. For instance, salience associated with the occurrence of extreme events can best explain decision-making (Gallagher, 2014). In the case of an oil spill it seems likely that the visualization experience assumes a more important role since it is cognitively harder to visualize as oil spills have been seldom experienced in Portugal.

References

Ahtiainen, H. (2007). "The willingness to pay for reducing the harm from future oil spills in the Gulf of Finland – an application of the contingent valuation method."

Alvarez-Farizo, B. (1999). "Estimating the benefits of agri-environmental policy: Econometric issues in open-ended contingent valuation studies." *Journal of Environmental Planning and Management*, 42(1): 23-43.

Bateman, I. J., Day, B. H., Jones, A. P., and Jude, S. (2009). "Reducing gain–loss asymmetry: A virtual reality choice experiment valuing land use change." *Journal of Environmental Economics and Management*, 58: 106-118.

Biodiversity Information System for Europe. (2020). "Ecosystem services – Biodiversity Information System for Europe." Accessed January 22. https://biodiversity.europa.eu/topics/ecosystem-services.

Bishop, R. C., Boyle, K. J., Carson, R. T., Chapman, D., Hanemann, W. M., Kanninen, B., ..., and Paterson, R. (2017). "Putting a value on injuries to natural assets: The BP oil spill." *Science*, 356(6335), 253-254.

Cameron, A. C., and Trivedi, P. (2009). *Microeconometrics Using Stata*. Texas: Stata Press.

Carson, R. T., Mitchell, R. C., Hanemann, M., Kopp, R. J., Presser, S., and Rudd, P. A. (2003). "Contingent valuation and lost passive use: Damages from the Exxon Valdez oil spill." *Environmental and Resource Economics*. 25: 257-286.

Casey, J. F., Kahn, J. R., and Rivas, A. A. F. (2008). "Willingness to accept compensation for the environmental risks of oil transport on the Amazon: A choice modeling experiment." *Ecological Economics*, 67: 552-559.

Farooq, B., Cherchi, E., and Sobhani, A. (2018). "Virtual immersive reality for stated preference travel behavior experiments: A case study of autonomous vehicles on urban roads." *Transportation Research Record: Journal of the Transportation Research Board*, 2672(50): 35-45

Gallagher, J. (2014). "Learning about an infrequent event: evidence from flood insurance take-up in the United States." *American Economic Journal: Applied Economics*, 206-233.

Grigalunas, T. A., Anderson, R. C., Brown Jr., G. M., Congar, R., Made, N., and Sorensen, P. E. (1996). "Estimating the cost of oil spills: lessons from the Amoco Cadiz incident." *Marine Resource Economics,* 2(3): 239–262.

Iftekhar, S., Buurman, J., Lee, T. K., He, Q., and Chen, E. (2019). "Non-market value of Singapore's ABC Waters Program." *Water Research,* 157.

Innocenti, A. (2017). "Virtual reality experiments in economics." *Journal of Behavioral and Experimental Economics,* 69: 71-77.

Instituto de Conservação da Natureza e das Florestas. (2020). "Parque Natural da Arrábida – ICNF." Accessed January 22. http://www2.icnf.pt/portal/ap/p-nat/pnar.

Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., ... & Tourangeau, R. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2): 319-405.

Liu, X., Wirtz, K., Kannen, A., and Kraft, D. (2009). "Willingness to pay among households to prevent coastal resources from polluting by oil spills: A pilot survey." *Marine Pollution Bulletin*, 58: 1514-1521.

Lopes, R., and Videira, N. (2016). “A collaborative approach for scoping ecosystem services with stakeholders: The case of Arrábida Natural Park.” *Environmental Management*, 58: 323-342.

Loureiro, M. L., Loomis, J. B., and Vásquez, M. X. (2009). “Economic valuation of environmental damages due to the *Prestige* oil spill in Spain.” *Environ Resource Econ*, 44: 537-553.

Mol, J. M. (2019). “Goggles in the lab: Economic experiments in immersive virtual environments.” *Journal of Behavioral and Experimental Economics*, 79: 155-164.

Navrud, S., Lindhjem, H., and Magnussen, K. (2017). “Valuing marine ecosystem services loss from oil spills for use in cost-benefit analysis of preventive measures.” In *Handbook on the economics and management of sustainable oceans*, ed. Paulo A. L. D. Nunes, Lisa E. Svensson and Anil Markandya, 124-137. Cheltenham: Edward Elgar Publishing.

Poe, G. L., Clark, J. E., Rondeau, D., and Schulze, W. D. (2002). “Provision point mechanisms and field validity tests of contingent valuation.” *Environmental and Resource Economics*, 23: 105-131.

Van Biervliet, K., Le Roy, D., and Nunes, P. A. L. D. (2005). “A contingent valuation study of an accidental oil spill along the Belgian coast.” *Marine Resource Damage Assessment*, 8: 165-207.

Van Herpen, E., van den Broek, E., van Trijp, H. C. M., and Yu, T. (2016). “Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store.” *Appetite*. 107: 196-207.

Vasconcelos-Raposo, J., Bessa, M., Melo, M., Barbosa, L., Rodrigues, R., Teixeira, C., Cabral, L., and Sousa, A. (2016). "Adaptation and validation of the iGroup presence questionnaire (IPQ) in a Portuguese sample." *Presence*, 25(3): 191-203.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge: The MIT Press.

Zawojska, E., Bartczak, A., and Czajkowski, M. (2019). "Disentangling the effects of policy and payment consequentiality and risk attitudes on stated preferences." *Journal of Environmental Economics and Management*, 93: 63-84.

8. Appendix

8.1. Figures

Figure 1: Map of Arrábida Natural Park including the marine park

Figure 2: Scenario 1 of the VR experiment

Figure 3: Scenario 2 of the VR experiment

Figure 4: Scenario 3 of the VR experiment

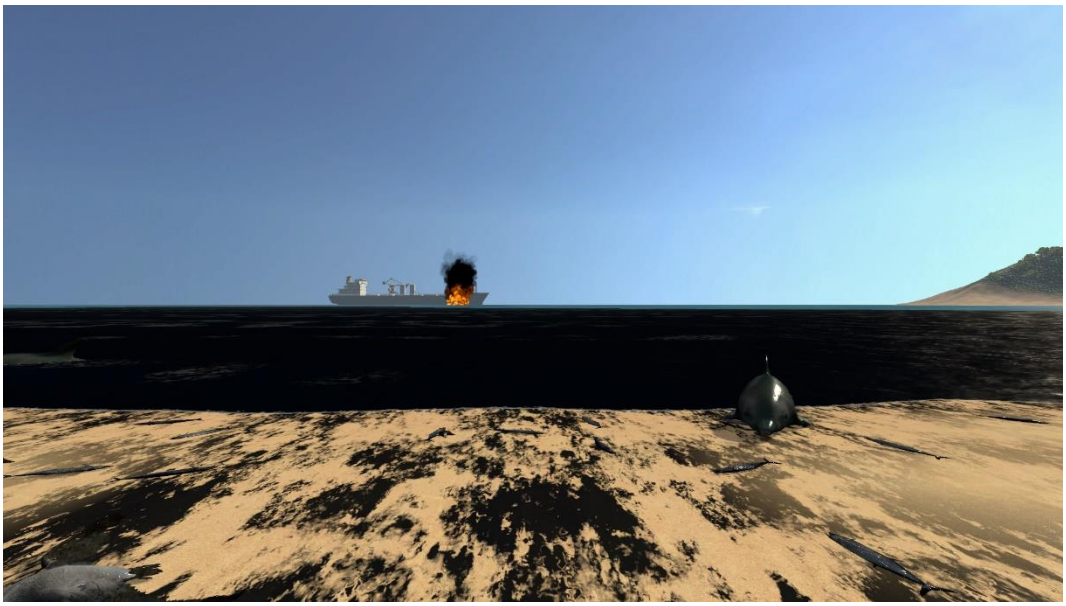
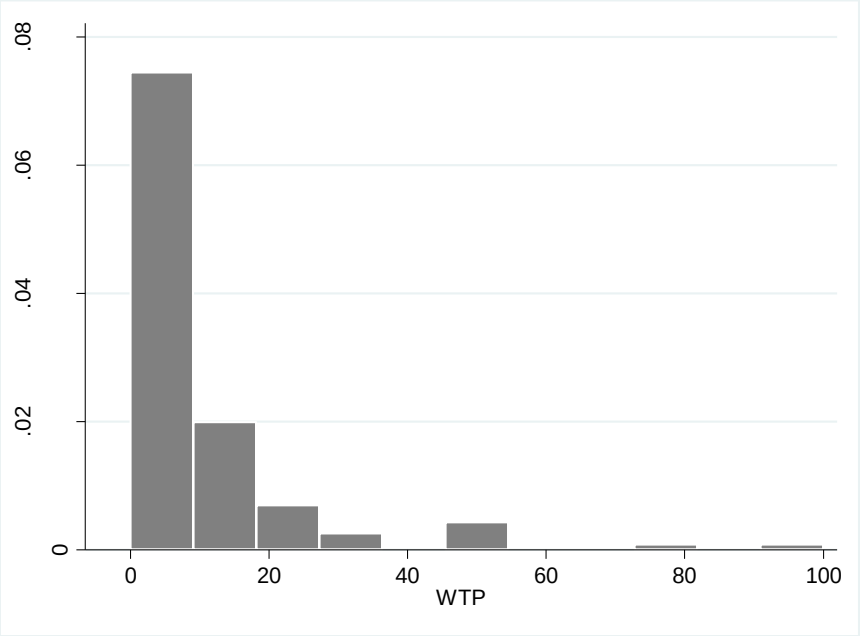


Figure 5: Histogram of *WTP*



8.2. Tables

Table 3: Variables description

Variable	**Description**
WTP	Dependent variable, corresponding to the amount bided (in euros) by each participant.
Gender	A dummy variable coded as 1 if the individual is a female and 0 otherwise.
Age	The age of the individual in years.
Income	Individual's monthly disposable income (euros) coded as the centre of the corresponding category: no income of their own as 0,]0;600] as 300,]600;1000] as 800,]1000;2000] as 1500 and]2000;3500] as 2750.
Unfair	A dummy variable coded as 1 if the individual agreed with at least one of the three fairness-related questions (if the costs of the program should be borne by the residents, oil companies or through taxation).
Contributions	A measure of how frequently the individual makes monetary contributions to environmental organizations, with 1 as "never", 2 as "rarely", 3 as "sometimes", 4 as "regularly" and 5 as "always".
Beaches	A measure of how frequently the individual goes to ANP beaches, built as the sum of the ten 3-point Likert items related to beach use.
HIVE	A dummy variable coded as 1 if the individual was assigned the HIVE version of the experiment and 0 otherwise (LIVE version).

Table 4: Results of the standard tobit model

	Standard Tobit
HIVE	**-4.61402**
	(4.00877)
Age	**-0.34304 ***
	(0.19907)
Gender	**5.97448**
	(4.45855)
Income	**0.00438**
	(0.00275)
Contributions	**6.70394 *****
	(2.05890)
Beaches	**0.10973**
	(0.69164)
Unfair	**1.29552**
	(6.85006)
Constant	**-9.81563**
	(11.64088)
Number of observations	127
Censored observations	51
LR chi2	17.75
Prob > chi2	0.0131
Pseudo R-squared	0.0234
Log Likelihood	-370.09477

Standard errors in parentheses. *, **, and *** denote significance at 10%, 5%, and at 1%, respectively.

Table 5: Conditional moment test against the null of normal errors

CM	Prob > chi2
25.995	0.00000

Table 6: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
WTP	127	8.43307	15.13882	0	100
HIVE	127	0.49606	0.50196	0	1
Age	127	33.35433	11.76255	19	68
Gender	127	0.61417	0.48871	0	1
Income	127	1013.787	824.652	0	1
Contributions	127	2.11024	0.99386	1	5
Beaches	127	12.9685	3.07302	10	28
Unfair	127	089764	0.30432	0	1

Table 7: IPQ items table

ID	IPQ item	English question	English anchors
1	INV1	How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?	"extremely aware", "moderately aware" or "not aware at all"
2	REAL3	How real did the virtual world seem to you?	"about as real as an imagined world" to "indistinguishable from the real world"
3	SP4	I had a sense of acting in the virtual space, rather than operating something from outside.	"fully disagree" to "fully agree"
4	REAL2	How much did your experience in the virtual environment seem consistent with your real-world experience?	"not consistent", "moderately consistent" or "very consistent"
5	REAL1	How real did the virtual world seem to you?	"completely real" to "not real at all"
6	SP3	I did not feel present in the virtual space.	"did not feel" to "felt present"
7	INV2	I was not aware of my real environment.	"fully disagree" to "fully agree"
8	G1	In the computer-generated world, I had a sense of "being there".	"not at all" to "very much"
9	SP1	Somehow, I felt that the virtual world surrounded me.	"fully disagree" to "fully agree"
10	SP5	I felt present in the virtual space.	"fully disagree" to "fully agree"
11	INV3	I still paid attention to the real environment.	"fully disagree" to "fully agree"
12	REAL4	The virtual world seemed more realistic than the real world.	"fully disagree" to "fully agree"
13	SP2	I felt like I was just perceiving pictures.	"fully disagree" to "fully agree"
14	INV4	I was completely captivated by the virtual world.	"fully disagree" to "fully agree"

Note: the IPQ questions presented above correspond to the English version of the questionnaire and can be found in http://www.igroup.org/pq/ipq/download.php?fbclid=IwAR0e6tXUD1txppiPTgNEQ6UNTl-FgoPK4p4KuWlhz7HR2B1FEqxLLzpVvjE. However, participants answered the Portuguese IPQ version from Vasconcelos-Raposo et al. (2016).

Table 8: *Global Presence* - Two-sample t-test of unequal variances

Group	**Obs**	**Mean**	**Std. Err.**	**Std. Dev.**	**[95% Conf. Interval]**	
LIVE	64	2.97569	0.06203	0.49623	2.85174	3.09965
HIVE	63	3.15212	0.05281	0.41919	3.04654	3.25769
Combined	127	3.06321	0.04138	0.46629	2.98133	3.14509
Diff		-0.17642	0.08147		-0.33769	-0.01515

diff = mean (*LIVE*) – mean (*HIVE*) t = -2.1656

H_0: diff = 0 Satterthwaite's degrees of freedom = 122.197

H_a: diff < 0	H_a: diff ≠ 0	H_a: diff > 0
Pr(T<t) = 0.0161	Pr(|T| > |t|) = 0.0323	Pr(|T| < |t|) = 0.9839

Table 9: Robustness checks

	Exponential Type II Tobit		**Lognormal Tobit**
	(1)	**(2)**	
HIVE	**-0.07300**	**-0.27999**	**-0.32754**
	(0.23871)	(0.23999)	(0.34882)
Age	**-0. 04077 *** **	**-0.00815**	**-0.04710*** **
	(0.01191)	(0.01397)	(0.01738)
Gender	**0.67221 ** **	**0.03981**	**0.62240 * **
	(0.26572)	(0.26561)	(0.38719)
Income	**0.00031 ** **	**0.00038 ** **	**0.00052 ** **
	(0.00016)	(0.00017)	(0.00024)
Contributions	**0.50715 *** **	**0.30823 ** **	**0.63215**
	(0.16204)	(0.14209)	(0.17978)
Beaches	**0.01066**	**0.00322**	**0.02196**
	(0.04155)	(0.04241)	(0.06005)
Unfair	**-0.54046**		**-0.21767**
	(0.39537)		(0.59575)
Constant	**0.25008**	**0.85791**	**0.11528**
	(0.64565)	(0.67458)	(1.01117)
Number of Observations	127		127
Censored Observations	51		51
ρ	**0.82438** (0.16487)		-
σ	**1.17027** (0.15987)		-
λ	**0.92638** (0.30229)		
Wald chi2(13)	39.34		-
Wald chi2(7)	-		22.71
Prob > chi2	0.0002		0.0019
Pseudo R-squared	0.0635		0.0381
Log-Likelihood	-333.267		-353.125

Equation 1 denotes the participation equation and equation 2 denotes the amount equation of the ET2T model. Standard errors in parentheses. *, **, and *** denote significance at 10%, 5%, and at 1%, respectively.

8.3. Equations

Equation 3: *Spatial Presence*

$$Spatial\ Presence = \frac{(SP1 + SP4 + SP5 + G1) + (-1 \times SP2 + 6) + (-1 \times SP3 + 6)}{6} \quad (3)$$

Equation 4: *Involvement*

$$Involvement = \frac{(INV2 + INV4) + (-1 \times INV1 + 6) + (-1 \times INV3 + 6)}{4} \quad (4)$$

Equation 5: *Experienced Realism*

$$Experienced\ Realism = \frac{REAL1 + REAL2 + REAL3 + REAL4}{4} \quad (5)$$

Equation 6: *Global Presence*

$$Global\ Presence = \frac{Spatial\ Presence + Involvement + Experienced\ Realism}{3} \quad (6)$$