An Economic Valuation of Urban Cleanliness in Zaragoza, Spain: The Role of Time and Monetary Contributions

"Not everything that is faced can be changed,
but nothing can be changed until it is faced." - **James Baldwin**

Abstract

Urban cleanliness constitutes a public environmental good that can not be marketed and is subject to non-rival and non-excludable conditions that hinder its full provision. Despite long-standing efforts, litter continues to impose negative externalities on the environment, society, and the economy, and has become a ubiquitous element in cities. In this regard, cost-benefit analysis provides the optimal toolkit for economists to rigorously assess waste policy interventions' viability, maximizing their effectiveness and efficiency. While the financial costs of removing litter from urban spaces are straightforward to calculate, the economic benefits of such a public policy are uncertain and difficult to measure. Stated preferences, and in particular contingent valuation and behavioral methodologies provide a sound framework to explore and measure the benefits associated with the commons. The present work integrates both methods in a survey scheme to measure time and monetary contributions via a payment card elicitation format. Time is measured with a concept branded as "willingness to cleanup" and monetary contributions are expressed in terms of "willingness to pay". The former is gauged by participation in public cleanups and the latter is appraised by the payment of a green tax. The study's objective is then to measure the non-commercial benefits of litter removal, explore the determinants of willingness to contribute, and propose policy measures to face the litter challenge in cities. Results suggest that pro-environmental attitudes and behavior, education, and age are the main driving factors of time and monetary contributions by citizens in the sample. Respondents in the sample are willing to contribute a total of 7.65 hours per year and are willing to pay 19.48 euros per year in taxes to eradicate litter from urban spaces. The lion's share of participants would contribute to face the litter predicament and aggregated welfare estimates suggest that time contributions provide better economic outcomes than monetary payments.

Master:	Spatial, Transport, and Environmental Economics
Faculty:	School of Business and Economics
Institution:	Vrije Universiteit Amsterdam
Author:	Antonio Arruebo Lacleta
Supervisor:	Dr. Erik Ansink
Date:	January 5, 2023

Acknowledgments

Dear reader,

The present document constitutes the work that marks the completion of the master's degree in Spatial, Transport, and Environmental Economics at Vrije Universiteit Amsterdam. This research represents the bridge between my past activist life in organizing public cleanups and the academic and scientific knowledge that is necessary for catalyzing meaningful change. Concerning the master, it has been an arduous odyssey, plagued with perils, and that leaves an everlasting mark on me. I have come to learn that, regardless of the circumstances, industry, discipline, and mental strength make hardships easier to overcome. What is under your control can and should be mastered and resolved. What is not, shall not torment you further. Additionally, I can confidently state that I learned how to think like an environmental economist, which gives me an effective edge for the rest of my life.

I would like to thank my family and friends for their unconditional support throughout this stage of my life. I feel grateful for having such caring and loving people by my side, that ease any kind of pain. Furthermore, I would like to express my sincere gratitude to my thesis supervisor dr. Erik Ansink for his expert academic mentorship, splendid ideas, and research ambition. The depths achieved in the present study would have never been reached without him. Lastly, thanks to all the participants that responded and disseminated the survey, enabling me to obtain a relatively large sample in a very short period. Thank you all!

To be continued,

Antonio Arruebo Lacleta.

Amsterdam, January 2023

List of Figures

List of Tables

*: Note that commas are utilized as decimal markers.
**: Note that points are utilized as decimal markers.

Contents

1 Introduction

1.1 Background

Even though municipal waste only constitutes an approximate figure between 7% and 10% of the total waste produced in the European Union, the management of municipal waste streams is amongst the most intricate and hard-to-solve issues in European cities (Directive (EU) 2018/851 on waste, 2018). "The challenges of municipal waste management result from its highly complex and mixed composition, direct proximity of the generated waste to citizens, a very high public visibility, and its impact on the environment and human health" (Directive (EU) 2018/851 on waste, 2018). The legislative framework enunciated in this directive on waste provides an excellent legal foundation for establishing an "efficient collection scheme, an effective sorting system and a proper tracing of waste streams, the active engagement of citizens and businesses, an infrastructure adjusted to the specific waste composition, and an elaborate financing system" (Directive (EU) 2018/851 on waste, 2018). Yet, urban cleanliness is still not provided in cities. Improper disposal and accumulation of waste continue to represent an ominous urban challenge. Litter is currently ubiquitous and has become an inherent component of our society in urban populated areas. Litter occurs any time someone - public and private agents, and/or citizens - makes the egregious decision of improperly disposing waste by simply abandoning it in public areas or the environment ("The Litter Challenge"). The main reason why litter continues to constitute a threat to the environment, the economy, and society stems from the fact that urban cleanliness is an environmental public good, subject to non-rivalrous and non-excludable conditions. The former condition essentially means that urban cleanliness does not dwindle as more respectful citizens enjoy it, while the latter entails that access restriction to urban cleanliness is unfeasible.

Common environmental goods, as Hardin (1968) enunciated, are cursed by what he coined as the "tragedy of the commons". A citizen, more or less consciously, finds that the cost of littering urban spaces is less than the cost of properly discharging residues, and derives a higher utility by polluting (Hardin, 1968). Since litter demeanor is not socially condemned and rarely punished by law enforcement officers, the incentive to not litter is very low for the average citizen. This conclusion is reached by a substantial share of the citizens sharing the same urban spaces, and the tragedy is served. In this regard, Hardin (1968) stated that "it is true that we are locked into a system of fouling our own nest, so long as we behave only as independent, rational, free enterprises" (Hardin, 1968; p. 1245). Consequently, if social and economic arrangements in the form of mutual coercion do not create the necessary accountability conditions, the common property is doomed to degrade (Hardin, 1968). These conditions and mechanisms generate what is referred to as "the litter predicament", which constitutes the problem faced in the present work. Despite the gargantuan magnitude of the problem, "the fight against litter need not be a losing battle" (Huffman et al., 1995, p. 154). While transformations of the waste management systems that operate in municipalities are a complex endeavor, societal environmental education would break the litter cycle after several generations, litter removal from streets, parks, and rivers fundamentally relies on little financial resources, human provision of labor, and collective determination, and has proven to have an immediate effect (Huffman et al., 1995).

Research must precede policy and in this regard, a sound cost-benefit analysis (CBA) is a fundamental tool that should be utilized to design and evaluate waste policy interventions. On the one hand, the costs of litter removal can be accurately measured and appear to be low, but on the other hand, the economic benefits of carrying out such an ambitious endeavor are

difficult to discern. As stated before, urban cleanliness is a public good that can not be marketed and priced, and therefore, the economic benefits remain unknown to policymakers and citizens. Not accounting for non-market benefits would outcome an incomplete and flawed assessment. With consequentiality and incentive compatibility at the heart of the investigation, this study both proposes policy, explores driving factors, and measures disposition to contribute with the help of the contingent valuation (CV) and contingent behavior (CB) methodologies.

1.2 Relevance of the research

Firstly, waste disposal and accumulation in urban areas impose harmful impacts on a wide range of dimensions, constituting a health and safety menace to humans, a grave threat to wildlife, a serious contributor to global warming and environmental degradation, an aesthetic scourge, and an economic burden (Eastman et al., 2013). Regarding health and safety hazards, broken bottles, for example, are the main cause of laceration, and the majority of injuries provoked by glass cuts take place on the streets (Martin and Makary, 1998). Discarded cigarette butts are considered to be the single most discarded residue worldwide, and pose a serious toxic waste problem (Novotny et al., 2009). Food envelopes and food waste are associated with the presence of disease vectors and pathogenic bacteria, in addition, to providing breeding habitats for roaches, mosquitoes, and rats (Muñoz-Cadena et al., 2012). Waste pollution in cities, not only represents a health and safety hazard but is also aesthetically unappealing, notoriously reducing the quality of public amenities such as parks or waterways (Carmi, 2019). Moreover, litter has an adverse economic impact due to the costs associated with litter collection and the economic losses derived from existing waste in public spaces; it reduces residential property values; diminishes tourism activity, and when litter reaches the ocean, it affects the marine economy and coastal communities (Huffman et al., 1995). Finally, litter represents a serious threat from an environmental protection perspective because waste pollution puts it at risk and kills wildlife (Carmi, 2019). Litter and, in particular, plastics, transported by the water and the wind, easily enter rivers and ultimately oceans, severely harming the land, coastal and marine ecosystems (Carmi, 2019). This multi-dimensional and overarching challenge poses a grave danger to our society and the environment and avoidance of all the previously enumerated negative impacts provides sufficient justification for conducting policy-relevant research on the matter. Incentive-compatible and consequential stated preference (SP) techniques, provide the optimal framework because it enables the construction of a hypothetical scenario where the local authorities are committed to solving the litter predicament and citizens' are petitioned to state their willingness to contribute.

Secondly, the literature review of this study has found no previous socioeconomic research devoted to the assessment of the non-commercial benefits of eliminating litter accumulation in cities, based on public perceptions, attitudes, and socio-demographic characteristics. Generally, the litter predicament has been studied with the use of CV methods in the context of coastal areas and marine ecosystems and represents a well-established contingent literature. Concerning time contributions, a few publications exist that explore the citizens' disposition to provide volunteer labor for the removal of litter, and the factors driving the decision to participate in cleanups by the use of CB schemes. Thus, while the contingent money market is cemented on the prominent and widely studied field of CV, the contingent time market had to be designed and analyzed based on the CV method but also grounded on intuitive reasoning. As stated before, the presence of litter in parks and urban rivers has become a permanent additional feature of these areas, to the point that it is difficult to visualize a future without it. However, and even though littering represents a formidable opponent, future desirable outcomes can be realized if conceived in the present with precise confidence. This is what is truly interesting and useful from Stated Preference techniques. They provide a well-established research toolkit that,

if properly used, gives the possibility to create a clear image of a potential future in the mind of the participant, which is now compelled to respond realistically in a way that is representative of their true intentions and subject to budget constraints. The keystone is to design a hypothetical setting of certain realization via a credible mechanism of change. Lastly, it should be highlighted that even though, the strategies proposed in the study are hypothetical, and designed to measure time and money contributions, they constitute two sound strategies for combatting the litter predicament in the short and medium term. Due to the importance of consequentiality and incentive compatibility, these two measures could be implemented at scale. Additionally, respondents' heterogeneous characteristics are exploited in the present research to discover the driving factors of a positive WTC and WTP. Insight into the determinants of time and monetary contributions is fundamental for understanding what drives pro-environmental behavior to better design future waste policies.

Thirdly, the measurement of non-market benefits should not be accounted for in isolation, but rather as an essential component of a comprehensive and sound Cost-Benefit Analysis (CBA) in waste policymaking. Research should play a fundamental role in the decision-making processes of policy design to assess and evaluate the impacts of a proposed environmental change (Annema et al., 2007). Consequently, this paper represents one of the several studies that should be conducted if a more comprehensive and effective policy is to be enunciated and applied. The challenge posed by waste pollution cannot be resolved by only removing existing and future residues, but it has to be tackled at the root. Ergo, a much wider, holistic, and multi-layered plan is required in Zaragoza. In spite of this, as expressed before, measuring the non-commercial benefits of urban cleanliness and disposition to volunteer in tackling the problem, is an essential task for effectively communicating with policymakers and citizens about the importance of the matter at hand. Concluding, there exist enough arguments for researching creative and collective solutions to the litter predicament. This work focuses on researching citizens' disposition to pay and/or cleanup to eradicate litter accumulation in Zaragoza's public spaces.

1.3 Research Questions

As stated before, both CV and CB methodologies are employed and integrated into one single survey to obtain data on willingness to pay (WTP) and willingness to cleanup (WTC) in the city of Zaragoza. By computing maximum likelihood estimations, research efforts are channeled toward answering the following research questions:

- What are the non-market benefits of removing litter in the city of Zaragoza?
- What is the disposition of citizens to participate in public and systematic clean-ups of parks and rivers?
- What are the main determinants of time and monetary contributions to the eradication of litter accumulation in Zaragoza?
- What policy measure renders the highest welfare outcome?

1.4 Structure

This thesis is formed by five main chapters. Firstly, Chapter 2 provides a review of the literature on CV and CB methods and the research applications to the litter predicament. Furthermore, the literature on time and monetary contributions to charity is explored. Secondly, Chapter 3 presents the econometric models employed for the analysis of results. Thirdly, Chapter 4 outlines the major characteristics of the data utilized in the present work and the data collection

process. Fourthly, results retrieved in the multivariate regression analysis are unveiled and discussed. The conclusions extracted from the entire investigation are included in Chapter 5. Lastly, the bibliography utilized in the study is presented in Chapter 7. Before full engagement, it should be noted that recommendations for future research are scattered across the entire paper and will be found in their respective sections of interest. This way, recommendations are presented more organically as they were written in this fashion.

2 Literature Review

The purpose of this literature review is to discuss relevant work concerning SP methods and the study of the relationship between time and money contributions to charitable activities. In the first section, contingent valuation and contingent behavioral models are reviewed, shedding a light on the literature body related to biases and how to minimize them. In particular, hypothetical bias, which seems to encompass all of them, is further analyzed in-depth. In the second section, the analysis focuses on the valuation of non-commercial benefits and behavioral valuation studies in the context of litter removal. Lastly, the third section examines the literature body on labor and monetary contributions to charity causes. This literature review provides the substrate of the investigation by guiding the conduct of the present work while "standing on the shoulders of the giants". Conclusions and thoughts are included at the end of every section to identify and briefly discuss knowledge gaps and future research recommendations.

2.1 Direct and Indirect Elicitation Methods

Individuals may be willing to make a monetary tradeoff to secure an increase in the environmental quality of a public good, even if that person or maybe anyone else will derive any pleasure or direct reward from it (Carson, 2012). This economic exchange is labeled in economics as "passive use value", although it can be named also"stewardship value" and "existence value" (Carson, 2012). When studying the total economic value of public goods, passive values cannot be revealed by decisions in the marketplace, because there is no market for that public good in the first place (Carson, 2012). The concept of passive use is a fundamental pillar for estimating the value of a change in a public good, where use and non-use (passive) values coexist (Freeman, 2003). Seems evident that, without market information, a change in the quality of a public good cannot be directly priced, which means that other strategies must be developed to capture the economic tradeoff at hand (Carson, 2012). Over more than 6 decades, in recognition of the importance of these non-commercial values, several non-market valuation strategies have been crafted and refined under the umbrella of one of the most prominent, yet controversial, research fields in welfare and environmental economics (Carson et al., 2001).

Traditionally, methods for the economic valuation of public and environmental amenities have been taxonomized into two main categories: indirect and direct procedures (Adamowicz et al., 1994). Indirect methods essentially utilize actual market and non-market choices made by individuals to constitute revealed preferences over goods (Adamowicz et al., 1994). Direct methods, conversely, estimate economic values based on hypothetical intended market and non-market behavior (Adamowicz et al., 1994). Two main economic valuation methodologies stand out as the most widely utilized over the years: the stated preference and the revealed preference (RP) methods. The former (SP) methodology harvests data on consumer preferences by direct elicitation, whereas the latter (RP) method deduces preferences from market demand choices (McFadden, 2017). While both methodologies provide estimates of the economic value of non-market benefits using stated responses of participants to survey schemes, these two approaches differ substantially (Johnston et al., 2017). Both valuation techniques have edges and downsides, but ultimately, the research setting determines the suitability of one approach or the other (Adamowicz et al., 1994). In this regard, indirect techniques can not be applied in settings that have no record of past experiences, and elicitation of economic values would require high-risk simulation and extrapolation of past to future behavior (Adamowicz et al., 1994). On the contrary, SP methods are specifically designed for this endeavor, providing

researchers with the optimal research technique for measuring quality changes in public goods and the environment (Adamowicz et al., 1994).

2.1.1 Stated Preference Methods

In SP, a plethora of approaches exist for measuring economic value. The three most common SP methods that fulfill the elicitation task are contingent valuation (CV), choice experiments (CE), and contingent behavior (CB) (Xie et al., 2022). The CV method involves respondents being questioned whether they would support, or not, a proposed change (single binary choice) and at what price (Johnston et al., 2017). For example, Abas et al. (2021), utilize a CV methodology to estimate the mean WTP of residents in the rural area of Kelantan for improved management of solid waste. In other words, residents were straightforwardly asked to state the maximum amount they wished to pay for improved management of municipal waste (Abas et al., 2021). The CE method has respondents expressing their consumption taste among two or more multi-attribute options (Johnston et al., 2017). For instance, Schuhmann et al. (2016), treat a beach as a multi-dimensional system that provides an array of benefits that are valued by individuals. The authors elicit preferences by a survey design that involves a series of attributes (e.g. beach litter) and combine them with a set of levels (e.g. number of residues per 25m2 ranging from 0 to 15) (Schuhmann et al., 2016). Then, they integrate every possible combination of attributes and levels and create six scenarios depicting every possible setting (Schuhmann et al., 2016). Finally, respondents are asked to choose between them (Schuhmann et al., 2016). , in CV formats, the change to be valued affects the item under study as a whole, whereas, in CE formats, the change to be valued only affects specific attributes of the item (Johnston et al., 2017). Lastly, the CB method asks for intended behavior in quantities or frequencies given a proposed change (Xie et al., 2022). For instance, Whitehead (2005) incorporated behavioral change in the WTP scheme by jointly estimating CV, revealed preferences, and contingent behavioral data. Respondents are asked for their WTP to undergo a water quality improvement in the Neuse River, North Carolina; then, the authors gather revealed behavior trip data by asking respondents for the number of trips taken before the potential water quality improvement (during the past 12 months); and finally, responders are petitioned to state the number of trips they would take if water quality actually improved (during the next 12 months) (Whitehead, 2005). Concluding, the focal point of CV/CE and CB is different: CV and CE concentrate on monetary valuation (how much are you willing to pay) while the CB approach concentrates on behavior (e.g. how many trips one would take) (Xie et al., 2022).

2.1.2 Validity Assessment of CV and CB Methods

SP methods are considered to be the only available approach to estimate use and non-use values, in addition to intended behavior, associated with changes that fall outside the marketplace (Johnston et al., 2017). However, CV and CB methods are also highly controversial, giving birth to an intense and long-lived debate over their validity (Carson et al., 2003). Historically, economic theory has been grounded on the assumption that choices in the market are the key to understanding the motives underlying consumer behavior, whereas consumption or behavior decisions stated in a survey setting should be subject to systematic skepticism (Kling et al., 2012). This long-standing skepticism towards SP methods gave rise to a vast and robust published literature investigating the validity and reliability of valuation responses (Johnston et al., 2017). In terms of a sound validity assessment, three main aspects have to be assessed: content, criterion, construct, and convergent validity (Boyle et al., 1985). Content validity reflects the suitability of the valuation methodology and "adherence to best-practice guidance concerning design and implementation" (Johnston et al., 2017; p.371);

secondly, assessing for criterion validity involves comparing results from SP tasks with presumably true value estimations that involve real money payments (Boyle et al., 1985). Thus, one parallel study needs to be carried out to set it as the "criterion" (Xie et al., 2022); thirdly, construct validity is concerned with whether results are in line with prior expectations based on theory, intuition, and past empirical evidence (Bishop and Boyle). Those pre-established expectations are converted into hypotheses that can be examined afterward according to current data and statistical results (Boyle et al., 1985); lastly, convergent validity is considered to be an extension of construct validity and involves comparing SP results with elicitations from other valuation methods such as RP (Boyle et al., 1985). Because content validity refers to the use of current state-of-the-art procedures to "design and implement a survey, the content of the survey instrument itself, data analysis, and study reporting" (Johnston et al., 2017; p.372), it has been the main focus of the study to maximize the content validity of the study to compose a survey methodology that is conducive to the elicitation of true values. In addition, due to the existence of prior SP literature applied to litter, construct validity is also evaluated, by contrasting previous studies' results with the present ones.

2.1.3 Principal Concerns Over CV and CB Methods

The main concern over content validity in SP methods is that surveys have a hypothetical nature, which is, as explained before, a simultaneous blessing and curse for sound economic valuation of public and environmental goods. Focusing on the problematic side, respondents have no previous market experience and are unable to accurately answer a hypothetical question, regardless of whether it involves a monetary or a time transaction (Hausman, 2012). This means that, if the intended behavior revealed during a survey does not correlate with the posterior actual behavior, then, results from the behavioral intentions data set are not useful for policy. Nonetheless, there exists a wide range of minimization techniques for decreasing the likelihood of unrealistic responses in surveys which will be extensively covered in the present work.

Secondly, another major challenge to the CV method and its construct validity arises from what is known as "scope and embedding effects". Kahneman and Knetsch (1992) were the pioneers in studying what they called the "embedding effect": "the same good is assigned a lower value if WTP for it is inferred from WTP for a more inclusive good rather than if the particular good is evaluated on its own" (Kahneman and Knetsch, 1992; p.58). The authors found that residents in Toronto were only willing to pay a slightly higher amount for preventing a drop in fish stocks in all Ontario lakes than to protect fish in a small area of the province (Kahneman and Knetsch, 1992). Embedding is connected to the "scope effect", which is the more comprehensive hypothesis that willingness to pay for environmental goods should vary with their size (Hausman, 2012). Another striking example of embedding and scope effects was found by Desvousges et al. (1993). The authors discovered that the mean WTP to prevent 2,000 migratory birds from dying in oil-filled ponds was approximately equal to that of preventing the death of 20,000 or 200,000 birds (Desvousges et al., 1993). These results pose serious concerns about the rationality of responses in CV surveys due to the inconsistency with economic theory predictions which state that individuals' WTP for the desired good should increase as the goods' quantity increases (Carson et al., 2001). According to Carson et al. (2001), the primary cause of insensitivity problems is badly designed surveys and second-rate administration procedures. The author argues that all evidence found of scope insensitivity is subject to biased results arising from low-quality research attempts. For instance, the author argues that Desvouges et al's (1993) study on preventing birds from dying in oil ponds suffers from a poorly survey design and an inexpert administration procedure, done in a North Caroline shopping mall. Carson et al. (2001) conclude that none of the experiments on scope

and embedding effects are constructed by state-of-the-art CV procedures where participants fully comprehend the good undervaluation and the welfare consequences of the proposed change.

Thirdly, another important response bias arises when the researcher utilizes a discrete choice question with a follow-up which increases the bid if the respondent agrees on the payment of the first bid, or is diminished if the respondent does not agree on the payment of the initial bid. This approach is dubbed the "double-bounded model" (Flachaire and Hollard, 2006). The key drawback of the double-bounded model is that the participant's second response may be conditioned by the first bid offered (Flachaire and Hollard, 2006). This is known as the "starting point bias" and has been empirically proven to be positive and present in CV surveys (see Boyle et al., 1985). Many studies have searched for ways to effectively tackle the starting point bias (see Boyle et al., 1985; Holmes and Kramer, 1995; Boyle et al., 1997; and, Flachaire and Hollard, 2006), but the selection of a payment card (PC) scheme appears to be the most efficient and effective approach for eliminating it (Mitchell and Carson, 1981). For instance, a PC format has been successfully utilized by Bouma and Koetse (2019) in their study of the existing gap between stated and revealed donations and the effect of behavioral factors on hypothetical bias in the CV methodology.

Fourthly, there is a potential for information bias, which originates when the respondent does not understand nor believe the information provided in the survey, and as a consequence, she is not able to provide an accurate response (Johnston et al., 2017). Johnston et al., (2017) recommend that SP questionnaires should explicitly explain the baseline conditions, the payment vehicle, and the environmental change under valuation with precision. Baseline conditions and proposed changes should be clearly explained and understood by the respondent, while the payment vehicle must be evidently presented in terms of monetary amounts, periodicity (e.g. monthly or annually), whether the payment is obligatory or voluntary, who is entitled to pay (e.g. households or individuals), the total period of the payment (e.g. 5 years), and the payment scheme (e.g. taxes) (Johnston et al., 2017). Finally, Johnston et al. (2017) advise testing the credibility and coherence of the survey-design elements selected via e.g. focus groups, to develop a questionnaire that presents information effectively and steadily.

Importantly, respondents may also respond strategically. There are two main types of strategic behavior: "free-riding" and "over-pledging" (Venkatachalam, 2004). Free-riding behavior materializes if an individual states a lower WTP for a public good, with the hope that others pay higher amounts that cover up the costs related to the provision of the public good, and as a consequence, she would not have to pay (Venkatachalam, 2004). Conversely, an individual may over-pledge by overstating his/her WTP for a public good to secure its provision, given that the referendum is only centered on the decision of supplying a public good or not, not in the actual contributions stated by participants (Venkatachalam, 2004). Very few studies have attempted to study the strategic bias in CV surveys, but the majority of them conclude that it is not a serious threat to CV research (Venkatachalam, 2004). According to Johnston et al. (2017) and Mitchell and Carson (1989) design procedures in place for minimizing hypothetical bias, simultaneously deal with strategic bias. Thereby, the fundamental objective of survey schemes should be to present participants with an incentive-compatible scenario that involves real-world consequential decisions (Johnston et al., 2017).

Finally, a hardwired component of CV survey schemes is the existence of a substantial proportion of respondents who are not willing to pay to secure an improvement in a particular environmental public good (Jorgensen and Syme, 2000). Among them, some may state a zero bid because they may feel they can not afford to pay and/or they simply do not value the good that is subject to valuation (Jorgensen and Syme, 2000). These responses are considered to be

true-zero bids. Conversely, some respondents may also state a zero bid because they oppose some aspect of the valuation methodology (Jorgensen and Syme, 2000). These zero bids are commonly dubbed as "protests". For instance, Edwards and Anderson (1987, p.168) define protest bids as "valuations that are intended to express displeasure with some part of the contingent market rather than to reveal true preferences". Generally, contextual elements in survey schemes such as the payment vehicle are particularly problematic and give rise to protest behavior. A typical payment vehicle is the use of a tax as an instrument to elicit an individual's WTP for environmental goods. Taxes are troublesome because some participants may hold strong beliefs against heavy taxation by governments and may decide to state a zero bid as a protest. As Mitchell and Carson (1989, p.166) write, protestors "refuse to play the game economists want them to play in a CV study" and wittingly or unwittingly decide to not reveal their true monetary valuation of the environmental good. To make matters worse, recent empirical research has discovered that protestors may also decide to participate in the market, under/overstating their true valuation as a sign of protest (Brouwer and Martín-Ortega, 2012). Strazzera et al. (2003) propose a two-step methodology for dealing with protest bids: first, create a survey scheme that enables accurate distinction between authentic zero bids and protest responses; second, utilize a statistical model that can adequately treat each type of zero and positive bid. The first step is rather straightforward by including a follow-up question to zero bidders, to understand if the individual truly places a zero value on the environmental good or not (Strazzera et al., 2003). On the contrary, controlling for positive protest responses is more complicated because it hugely increases the complexity of the survey for the respondent (Brouwer and Martín-Ortega, 2012). The second step is even more challenging and has been widely discussed over the years, although econometric modeling of protest responses is still a contemporary field of research (Strazzera et al., 2003). In this regard, frequently protest bids are conveniently excluded from the analysis, but as proved by Calia and Strazzera (2001), this may not be an appropriate strategy if protest responses create a sample selection bias. Removing these observations would simply assume that these responses are randomly distributed across the sample (Brouwer and Martín-Ortega, 2012). In addition to that, censoring protestors disenfranchises them from market valuation (Gowdy, 2004). Nevertheless, despite the intrinsic complexity of the statistic treatment of protest bids, recent literature provides numerous effective econometrical approaches without the necessity of eliminating them or treating them as zero bids (see Strazzera et al., 2003; Chen and Qi, 2018; Brouwer and Martín-Ortega, 2012; and Meyerhoff and Liebe, 2006). These statistical procedures have been taken into account in this study and will be discussed and applied.

Shockingly, all previously discussed biases are only identified and measured in CV methods, but not in CB schemes. Thus, it remains unknown the source and magnitude of the impact of biases in contingent behavioral market surveys. Based on Xie et al. (2022), a parallel strategy has been applied to both methodologies because the source and impact of biases may be similar between them. This strategy will be further discussed in the coming sections. Concluding, there exists a wide array of biases originating from the fact that respondents lack experience with non-market choices and provide unrealistic responses (Johnston et al., 2017). All these biases should be avoided via survey design and when possible, surveys should be designed to investigate the origin and magnitude of these effects (Johnston et al., 2017). Overall, as stated before, the major weakness (and strength) of SP methods is their hypothetical nature, which is the main focus of this literature review.

2.1.4 Hypothetical Bias in CV and CB Methods

Essentially, hypothetical bias can be determined by the difference between stated and revealed values (Murphy et al., 2005). Differences account for the fact that surveys do not generally

involve an experiment in a real setting that involves people's actual money (Loomis, 2011). As stated above, CB shares with CV, the potential concerns over its validity since both methodologies are grounded on the construction of a hypothetical valuation scenario (Xie et al., 2022). Respondents have to visualize an unknown situation in which information is imperfect (Whitehead et al., 2010). As a result, participants may "discount costs or income constraints or optimistically forecast avid recreation behavior" (Whitehead et al., 2010; p.92). According to Englin and Cameron (1996), it is also plausible that individuals are more capable to state their intended behavior in hypothetical circumstances rather than stating their WTP. Nevertheless, the CB method should still be subject to ample skepticism due to its hypothetical nature (Grijalva et al., 2002). In this regard, and even though the economic theoretical principles that sustain welfare analysis are well established (Freeman, 1993), no widely accepted theory explains the root causes of hypothetical bias or how to reduce its impact on results (Murphy et al., 2005). Nonetheless, there are various well-enunciated hypotheses on people's behavior in surveys involving hypothetical payments for changes in a public good (Loomis, 2011). Each of these hypotheses suggests a different approach to assuaging hypothetical bias (Loomis, 2011).

2.1.4.1 Paths to Minimize Hypothetical Bias in the CV method

Ex-Ante Approaches

According to Loomis (2011), three main ex-ante procedures have been proposed to alleviate hypothetical bias via survey design. Firstly, Carson and Groves (2007) argue that two conditions have to be met: (1) the respondents must view their answers as potentially influencing policy (policy consequential), and (2) they have to care about what the possible outcomes of their responses might be (payment consequential). That is, it must have some impact on their future utility (Loomis, 2011). A hypothetical designed market that meets these two criteria is deemed to be "strongly consequential" to the respondents, in the sense that the valuation scenario and the vehicle of payment are certain and binding (Loomis, 2011). Carson and Grooves (2007) and Vossler et al. (2012) empirically show that such surveys encourage truthful preference revelations. Secondly, another ex-ante approach to diminish hypothetical bias is the use of "cheap talk" (Loomis, 2011). This approach involves designing a survey that explicitly communicates to participants what hypothetical bias is and why it might occur (Cummings and Taylor, 1999). The authors found promising results from experiments that attempted to test the efficacy of cheap talk design. However, subsequent studies have demonstrated that cheap talks do not universally reduce hypothetical bias (Murphy and Stevens, 2005; Loomis, 2011; Aadland and Caplan, 2006). These studies suggest that the use of this ex-ante procedure may have an impact on consequentiality and that it only functions optimally under very specific conditions. Therefore, it is generally not recommended (Murphy and Stevens, 2005; Aadland and Caplan, 2006). Finally, other procedures have been suggested including "honesty oaths" (Jacquemet et al., 2017). Empirical evidence provided by Jacquemet et al. (2017) suggests that a person is more likely, to tell the truth after signing an "honesty oath". Nonetheless, according to Johnston et al. (2017), the efficacy of oaths is questionable because they may have unintended impacts on CV methods, and firmly disapprove their usage.

Lastly, the NOAA panel (Arrow et al., 1993) strongly recommended reminding participants that their stated WTP has an impact on their budgets and that expenditure in alternative goods should be reduced. Loomis et al. (1994) empirically studied the effectiveness of including "budget constraint reminders" with two identical surveys except for such a reminder. The results did not yield any significant statistical difference and obtained the same WTP in both surveys (Loomis et al.). On the contrary, Kotchen and Reiling (1999), following an identical

procedure, show that the inclusion of economic reminders substantially improves the statistical efficiency of the benefit estimates of environmental protection. Likewise, Johnston et al., (2017) argue in favor of the inclusion of budget reminders as a necessary element of CV research. Overall, the validity of these and other ex-ante procedures in abating hypothetical bias is moderate and in some cases counterproductive when two or more ex-ante approaches are used in combination (Johnston et al., 2017). Consequently, Johnston et al (2017) advocate for the use of a consequential design with a binding payment in combination with a limited number of ex-ante procedures carefully selected as the most effective approach.

Ex-Post Approaches

Among the ex-post approaches to alleviate hypothetical bias, Champ et al. (2009) test the validity of including follow-up questions about how certain the individual is that she would actually pay the amount if necessary. Respondents that are not confident about their answer are coded as "no" in the WTP statistical analysis of the CV (Champ et al., 2009). However, little evidence supports that allowing respondents to express their level of confidence with WTP questions reduces hypothetical bias (Loomis, 2011). Moreover, follow-up certainty questions substantially increase the length of surveys which could in turn lead to respondent fatigue and/ or the use of "simplifying choice heuristics" which decreases the validity of responses (Johnston et al., 2017). Additionally, the NOAA panel suggested the application of a "fudge factor" - that deflates the stated WTP by dividing it by two (Arrow et al.,1993). Yet, the 50% downward gauge of results in CV methods lacks objective and empirically defined criteria, so there is no guarantee that the calibration factor is precisely determined (Hausman, 2012; Johnston et al., 2017; Fox et al., 1998). Hence, generally, ex-post procedures are a better strategy for diminishing hypothetical bias (Johnston et al., 2017). In particular, consequentialism and incentive compatibility has proven to be very effective (Johnston et al., 2017).

General Theory and Conclusions

Despite all the research undertaken, a general theory of respondent behavior is yet to be enunciated. Each of the above-mentioned approaches is based on a different hypothesis about the source of hypothetical bias, even though some of these hypotheses are anchored on shared principles across the board (Loomis, 2011). The utmost objective is to elaborate a single all-encompassing theory of hypothetical bias that explains why respondents state a WTP that tends to exceed their actual WTP (Loomis, 2011). More empirical research must be carried out on testing the different hypotheses about the origin and magnitude of hypothetical bias and the validity of different survey designs to assuage the bias (Loomis, 2011). These tests will sharpen not only survey designs but also determine the optimal ex-post calibration of results depending on the survey setting (Loomis, 2011). Even though a general theory of hypothetical bias has not been articulated yet, skepticism, criticism, and lack of consensus have proven to be productive over the last three decades (Carson, 2012). A substantial amount of research has been devoted to the refinement of CV techniques (Carson, 2012) and empirical research has demonstrated that a diligently constructed survey scheme based on CV analysis is certainly more useful than no estimation in most situations for both CBAs and damage assessments (Kling et al., 2012). A CBA that omits the estimation of non-market values "will be at best incomplete and at worst completely misleading" (Carson et al., 2001; p.197) so the demand for SP value estimates is unquestionable (Carson, 2012).

2.1.4.2 Paths to Minimize Hypothetical Bias in the CB method

Combining Data

The most common approach for reducing hypothetical bias and maximizing convergent validity in a CB study is to compare and combine revealed and intended behavior (Whitehead et al., 2008). The aim is to capitalize on the "contrasting strengths of the various approaches while minimizing their weaknesses" (Whitehead et al., 2008; p.7) in an attempt to enrich the resulting data (Louviere et al., 2000). Whitehead (2005), and Grijalva et al. (2002) provide an excellent illustration of estimating behavioral change from an environmental quality improvement with revealed and stated behavior. However, revealed preference data can not be always obtained. Past behavior is inexistent for certain environmental changes in public goods (Whitehead et al., 2008). In the present research setting, there is no revealed preference data on past attendance to public clean-ups in a city where this solution has not been widely applied. As a result, the only possible approach is to design CB surveys that capture data on this hypothetical behavior and estimate intended behavioral responses to an environmental change (Whitehead et al., 2008). Another popular approach is to merge contingent behavior surveys with observed data in travel cost models (TCM). For instance, Alberini et al. (2007) combine past and hypothetical future fishing trips to the Lagoon of Venice in response to a hypothetical change in the price of the trip and/or in the catch rate. But yet again, this approach can not be utilized for the present research due to its characteristics.

Consequentiality

As previously stated, CB methods share a hypothetical quintessence with CV schemes, and consequently, similar concerns arise regarding their validity (Xie et al., 2022). As declared before, the main threat to validity is the problematic hypothetical bias, so CB methods demand a special focus on its incentive compatibility (Xie et al., 2022). As in CV, an incentive-compatible design compels respondents to veraciously reveal their intentions (Xie et al., 2022). It is important to note that, incentive-compatibility has to be constructed differently depending on the good's essence (public or private good) (Xie et al., 2022). On the one hand, experimental evidence has determined that CB elicitation for private goods is potentially affected by two hypothetical biases rooted in strategic behavior: price and provision bias (Xie et al., 2022). Price bias would lure participants that want to signal high price sensitivity into understating their actual intended trips to the site if they believe that the costs of improvements will fall upon them (Xie et al., 2022). Provision bias in CB would mean that participants overstate their intended recreational trips to a natural park to increase the probability of improvement provision (Xie et al., 2022). On the other hand, no research has attempted to explore the origin and impact of hypothetical bias in CB methods involving public goods. All the literature reviewed concerning intended behavior and willingness to participate in public clean-ups offers little explanation of the importance of incentive compatibility and how to minimize hypothetical bias. All the research consulted pursue an analogous strategy to CV methods.

General Theory and Conclusions

Overall, existing literature offers little evidence and counseling for conducting research on CB related to public goods. While great efforts have been devoted to the identification and mitigation of hypothetical bias in CV surveys, identical research is limited for the CB method. Existing literature shows evidence of considerable gains from combining revealed and intended behavior data, but its use is restricted to those studies that can collect historical behavioral data. Moreover, approaches for mitigation of hypothetical bias in CB studies have only been applied and tested in the context of private goods (generally recreation trips), but not for public goods. The vast prevailing gap of knowledge in this regard stands in the way of an optimal CB survey design. Further exploration should be conducted in this regard. Albeit, it is likely that hypothetical bias in CV originated similarly in CB studies and that has comparable effects on

validity. Consequently, a parallel strategy appears to be the optimal path to mitigate hypothetical bias in CB studies concerning public goods.

2.2 CV and CE Studies on Litter

As stated in the introduction, the closest comparative research available are articles that estimate the social costs of marine litter with the CV method. CE investigations are also studied due to the potential relevance of their approach and results for this study. With this intention, Shen et al. (2019) provide a list of three major studies devoted to the economic evaluation of the social costs of marine litter. Firstly, Smith et al. (1997) were the first to utilize a CV survey design to quantify the non-commercial values people placed on curbing marine pollution on recreational beaches in New Jersey and North Carolina (Smith et al., 1997). The authors employed different clean-up programs depicted in photographs as mechanisms of change that resulted in differing levels of beach litter and respondents were asked to value them. The estimated median WTP to cleanup marine litter ranged from US$ 21 to US$ 72 per person per year in annual income taxes (1992 price levels). Concerning socio-demographic variables, the authors were only concerned about the potential effect of selection bias in the results but did not exploit heterogeneity to study determinants of the WTP.

Secondly, Loomis and Santiago (2013) compared estimates from CV and CE survey methodologies in a split-sample design. The authors conducted in-person surveys at five beaches in Puerto Rico asking visitors to state their WTP for improving water clarity and removal of marine debris. The two different methodologies result in similar estimates of the mean WTP between US$ 98 and US$ 103 per visitor per day (2011 price levels). In this study, the authors controlled for various socio-economic variables (age, gender, education), perceived water clarity and cleanliness, time spent at the beach, perceived crowding level, and residency. Only the absence of trash and water clarity showed statistically significant effects on bid level. The authors also test for interaction terms between wave height and: (1) participation in water recreational activities such as snorkeling and (2) children in the group. They intended to test if wave height only mattered to those two groups of beach visitors because small waves benefit snorkeling recreation and the presence of children on beaches. Nonetheless, the results were not statistically significant to prove the hypothesis.

Lastly, Brouwer et al. (2017) estimate the social costs of marine debris across different European countries and coastal areas exploiting the same DCE survey design. They interviewed beach visitors at six different beach locations in Greece, Bulgaria, and The Netherlands asking them to declare their WTP for an entry fee or an increase in local taxes to mitigate marine litter. They utilized a total number of 200 valid interviews for Greece, 301 for Bulgaria, and 149 for the Netherlands. The mean WTP for removing plastic litter from beaches is US$ 0.67 per year in Greece, US$ 8.25 per year in Bulgaria, and US$ 2.05 per year in the Netherlands (2017 price levels). Concerning the removal of cigarette butts left by visitors, the average WTP is US$ 0.47 per year in Greece, US$ 7.98 in Bulgaria, and US$2.90 in the Netherlands (2017 price levels). The study results suggest that, overall, female beach visitors are more prone to pay than male visitors, and individuals that belong to an environmental organization or that have a higher income than the average, are more likely to pay the entrance fee/tax. As expected, the higher the dissatisfaction is with the actual state of beach cleanliness the more likely the individual will support the cleanup program. Surprisingly, local residents are less likely to agree on the payment compared to tourists. The authors point out that this may be because residents do not feel responsible for the litter present on the beach. Shen et al. (2019) in turn measured the social costs of beach debris in ten coastal scenic spots along the East China Sea in Zhejiang in the East China Sea, interviewing in person a total of 805 beach

visitors. Shen et al.'s (2019) article aims to estimate social costs using the two SP valuation methods (CV and DCE). The mean social costs of marine waste pollution at a city level are approximately US$ 1.08 to US$ 1.40 per visitor paying an entrance fee when utilizing the CV method, and US$ 1.00 to US$ 1.07 per visitor paying an entrance fee when applying the DCE method (2019 price levels). Results show that individuals with higher incomes and those concerned with marine litter positively influenced the WTP. Older people stated a higher WTP for the clean-up programs, while members of environmental organizations had a lower WTP. The authors do not explore possible explanations for this negative correlation between belonging to an environmental group and WTP to cleanup local beaches.

Other studies have been carried out around the world. Three of them are worth mentioning. Zambrano-Monserrate and Ruano (2020) aim to estimate the cost of environmental damage from plastic waste in the Galapagos by employing a CV survey design. The study utilized a face-to-face interview format and focused on three cities in Ecuador: Quito, Guayaquil, and Cuenca. Participants responded to a question about their maximum WTP for a hypothetical program that would reduce plastic pollution in the Galapagos. It is important to highlight the fact that the authors included a "cheap talk script" aimed at increasing the validity of the results. The average WTP estimated ranged between US$ 4.90 to US$ 14.51 per year, with a median of US$ 7.65 (2020 price levels). Predictably, individuals who stated that natural resources are important or very important, and that were worried about plastic pollution in the Galapagos Islands were more inclined to pay for the cleanup program. Moreover, in contrast with Shen et al. (2019), Zambrano-Monserrate and Ruano (2020) find that people who were part of an environmental organization during the last 12 months were more prone to be WTP. Other control variables such as age, gender, and education showed no evidence of significantly affecting WTP.

The Tyllianakis and Ferrini (2021) article aim to investigate the WTP and the perception and attitudes of Indonesian citizens toward policies that reduce macro-plastic contamination. The study utilizes a questionnaire designed to estimate WTP through the CV method. The survey was distributed via an online questionnaire platform and was completed by 822 respondents from the main islands that constitute the Indonesian Archipelago. The mechanism of change utilized was a hypothetical independent organization set up to support local governments in collecting and disposing of plastic waste from beaches and rivers, so respondents would contribute by providing financial support through an annual donation. The results indicate that the mean WTP is £13.5 to £15 which represents approximately 2$ of the average monthly salary (2021 price levels). In addition, the results indicate that age has a negative relationship with a participant's WTP, whereas men were more likely to state a higher WTP. Education was positively correlated to WTP.

Lastly, Abate et al. (2020) conduct a CV method study for measuring the WTP of Norwegian households for a hypothetical marine plastics pollution clean-up and prevention initiative in the archipelago of Svalbard. The authors utilized a CV method instead of a CE design because two pilot surveys identified that respondents regarded the proposed initiative to mitigate marine plastic pollution as a single item and not as a multidimensional one (Abate et al., 2020). The final sample of the article accounts for 552 valid responses at the national level. The authors find that the average WTP for the cleanup and prevention initiative around Svalbard is US$ 642 per household per year (2020 price levels), a very high value compared to the rest of the studies analyzed. Abate et al. (2020) suggest that the motive for these high WTP numbers might be a strong desire among the Norwegian population to preserve intact the majestic ecosystem of Svalbard. In the study, the majority of respondents were concerned with marine pollution and that had a positive impact on the WTP. Concerning the socio-demographic determinants of WTP, and in line with the above studies, results show that gender, age, and education are

positively correlated with WTP to reduce marine plastic in Svalbard. The authors further find that having visited the island of Svalbard had no impact on the WTP, which suggests that Norwegians strongly value the Arctic for its mere existence. This could also mean that plastic pollution is perceived as a huge concern regardless of the location (Abate et al., 2020).

Conclusions

Overall, all the previous studies estimate the mean WTP value which is considered to be the welfare loss experienced by beach visitors as a result of waste pollution on beaches. These social costs were estimated based on public perception of the damage caused by litter and not by actual measurements of litter in marine coastal areas. Only the article by Abate et al. (2020) used actual measurements of plastic pollution in Svalbard and how the initiative would decrease the amount of litter. This approach likely enhances a better understanding by participants of the baseline conditions and the proposed change as recommended by Johnston et al. (2017) which could improve the validity and reliability of responses. However, data on waste pollution is scarce and varies over time. More efforts should be dedicated to increasing the amount of information on litter pollution to construct clearer CV survey schemes.

All the articles have a strong consequential nature and are incentive compatible as a way of reducing hypothetical bias and strategic behavior. The majority of them follow the recommendations of the NOAA panel (Arrow et al., 1993) and Johnston et al. (2017). All of them agree on the use of ex-ante procedures for mitigating hypothetical bias, most importantly, the use of binding payments, and strong policy implications. Zambrano-Monserrate and Ruano (2020) were the only ones to control for strategic bias by utilizing a "cheap talk" approach. The WTP results are similar for most of the studies, except for Abate et al. (2020) where the WTP is strikingly high compared to the rest. Concerning heterogeneity in perceptions, attitudes, awareness, and demographics, the majority of studies point in the same direction: income, age, education, and being a member of an environmental organization are positively correlated to WTP for the removal of marine litter. Lastly, it should be noted that none of the articles reviewed treat protest bids adequately, but rather as true zero bids. Only Abate et al. (2020) mention the existence of protest responses but decides to exclude them from the analysis.

2.3 CB Studies on Litter

In parallel with the previous section, the concept of willingness to participate in public clean-up programs has only been studied in the context of marine litter. To my knowledge, four articles have conducted research on intended behavior to participate in cleanups but none of them is explicitly considered by the authors as a CB study. The article by Brouwer et al., (2017), in addition to the CV study, also asks beach visitors from the Netherlands, Greece, and Bulgaria for their willingness to volunteer in beach litter removal actions. First, participants were asked if they were willing to participate in a voluntary cleanup campaign, and second, they were asked about the number of hours they would contribute to the volunteer action. Follow-up questions were asked if the respondent was not willing to participate. The results show that dutch visitors were the least inclined to volunteer since only 26% of the participants would participate in a beach cleanup, followed by Greek beach visitors with a 55% acceptance rate and a 72% acceptance rate among Bulgarian beach visitors. With respect to the frequency of participation, the Dutch population would again contribute the least with an average of 3.4 hours per person per year whereas the Greek population would be willing to help an average of 6.8 hours per visitor per year. Lastly, Bulgarian visitors would contribute the most with an average of 14.8 hours per visitor per year. It can be observed that as the acceptance rate rises, the number of hours also increases between countries. As anticipated, the dirtier a visitor sees the beach, the

more probable she agrees to participate in cleanups. Younger male visitors were less prone to help in cleaning up (both in the number of days and hours) than younger female visitors. Those respondents that were members of an environmental organization and had a higher income were more likely to participate in cleanups than non-member and lower income groups.

Secondly, Shen et al. (2019) survey the willingness to volunteer in public cleanups in the Zhejiang Province and the number of days per month they would like to spend cleaning. Roughly 74% of the respondents would participate in a beach cleanup program (Shen et al., 2019), a very close acceptance rate to Bulgarian respondents in Brouwer et al. (2017). On average, interviewees were willing to devote 1.5 days per month to cleaning. The three most significant covariates affecting the willingness to participate and the number of days were: perceived beach litter, individual income, and age. Respondents who were annoyed by the presence of beach litter were more inclined to participate and contribute with more hours, whereas visitors with a higher income or older were less willing to participate in those beach cleanups. Generally, older visitors had higher incomes and as a result, were more willing to pay rather than to contribute with their time (Shen et al., 2019). Thirdly, Adam (2021) investigates the attitudes of international tourists toward beach litter and the characteristics of their willingness to participate in beach cleanups during their vacations in Ghana. This study only focuses on perceptions and the binary decision of attending a public cleanup or not. Data was collected from three beaches in Accra and Cape Cost and a total of 685 valid questionnaires were completed by tourists that had just arrived in the country. The findings of the study reveal that female visitors have a stronger pro-environmental behavior because they were six times more likely to undertake beach cleanups than males. Furthermore, older and richer tourists are 5 and 5.5 times respectively more likely to participate in a beach cleanup than younger visitors. Finally, results suggest that there is a positive correlation between education and willingness to help in beach cleanups.

Last but not least, Lucrezi and Digun-Aweto (2020) survey the willingness of beach visitors to participate in clean-up programs at Elegushi Royal Beach (ERB) in Nigeria. The investigation utilized a "quantitative, descriptive, and non-experimental research design, using a structured questionnaire survey as the measuring instrument targeting beach visitors at ERB" (Lucrezi and Digun-Aweto, 2020; p.5). The authors also exploit demographics, environmental attitudes, perceptions of marine and coastal debris, attitudes toward beach cleanups, and attitudes toward improved legislation against litter contamination. The surveys were carried out in person at the beaches and provided 512 valid responses. Surprisingly, around 35% of the participants declared to have participated before in a beach cleanup, but only 24% were willing to participate in a similar event at ERB. As expected, previous participation in cleanups was correlated with the disposition to participate in future cleanup events. Lastly, participants that stated strong environmental attitudes, awareness of plastic pollution, and desire to impose more stringent fines on beach littering, had a higher disposition to participate in future cleanups.

Conclusions

As stated above, these studies closely evoke CB literature works on intended behavior, especially in the article by Brouwer et al. (2017) which surveyed the frequencies of interested participants. Shockingly, none of them report on the procedures employed to mitigate hypothetical bias, strategic behavior, or other biases that could jeopardize the validity and reliability of results. No follow-up studies were carried out to test the difference between intended and actual behavior. Further research should be conducted to explore the magnitude and source of bias in CB studies, and public cleanup seems to be an optimal setting for it. Notwithstanding these issues, this literature is at the vanguard of research about intended

participation in clean-up schemes while exploiting heterogeneity in participants' characteristics, which is a very relevant topic of research that should be expanded in future academic research. Regarding the observed results, participation rates range from 24% in Nigeria (Lucrezi and Digun-Aweto, 2020) and 26% in the Netherlands (Brouwer et al., 2017) to 75% in Bulgaria (Brouwer et. al., 2017). As expected, in all studies, sociodemographic variables such as income, age, and environmental attitudes presented a positive correlation to the willingness to participate. Interestingly, Adam (2021) finds that female tourists are substantially more likely to participate in clean-ups than male tourists. Yet again, no mention is given to the existence of protest bids in CB surveys which remains to be a hidden component in studies that utilize this methodology.

2.4 On Time and Monetary Contributions

As previously mentioned, the present study researches two fundamental types of contributions conducted by citizens for the improvement of the environmental and public good. These two essential gifts are money and time. Though, it could be simplistically argued that one equals the other in the market (but also outside of it), a more in-depth analysis is required to understand the complexities, interdependencies, and determinants of these two fundamental charitable contributions. The following section thoroughly examines the body of literature related to time and monetary contributions to charity. It should be noted that all studies present below make use of revealed preferences' data, which opposes the present study that makes use of stated preferences.

2.4.1 Background

Ever since Adam Smith published 1776 his magnum opus "An Inquiry into the Nature and Causes of the Wealth of Nations", individual self-interested and rational behavior has been the essential theoretical substrate of the global market economic theory where only hyperrational unemotional agents, also dubbed as "homo economicus", could provide the necessary conditions for a Pareto efficient outcome in market exchanges (Kirchgässner, 2010). A Pareto efficient paradigm entails a market outcome where no agent can be better off without making another agent worse off, meaning that only self-interested behavior in the market can lead to welfare maximization. Even though Adam Smith did recognize in "The Theory of Moral Sentiments" of 1959 the existence of altruism in economic behavior, neoclassical economics embraced the more extremist economic and social narrative, which implied a departure from the classical conviction that at least some moral behavior is a fundamental requirement for better social outcomes (Kirchgässner, 2010). If the neoclassical economic viewpoint holds, a well-functioning exchange economy relies solely on selfish market practices which do not require any moral standards (Kirchgässner, 2010). On the other hand, real-world evidence appears to suggest that the neoclassical economic rationale does not hold, and that, there is indeed a minimal moral basis underlying every economic and social exchange. A straightforward exemplification is the existence of donations to non-profit organizations that rely on voluntary time and monetary contributions, but also on some degree of morality in the world (Bauer et al., 2013).

2.4.2 Determinants of Charity Contributions

The occurrence of charities, privately funded by volunteer donations of individuals, provides a clear-cut example of world societies grounded on elementary moral principles. Such charitable organizations will then utilize a share of the money funds to augment their efforts by enlisting a

larger number of volunteers, that are fairly compensated for any sustained expenses (Ansink et al., 2019). Therefore, individuals willing to fund the provision of a public good can select among different types of voluntary donations (Ansink et al., 2019). These voluntary contributions made by philanthropic individuals are fundamentally driven by preferences and attitudes, that shape the decision-making process of an individual (Cappellari et al., 2011). But such moral variables are intricate to identify and arduous to integrate into the utility-maximizing framework (Cappellari et al., 2011). According to Andreoni (2006), philanthropy and charity actions are "one of the greatest puzzles of economics" and defines economics as a "science based on precepts of self-interested behavior which does not easily accommodate behavior that is so clearly unselfish" (Andreoni, 2006; p.1204). There appears to be a very wide heterogeneous range of (un)observed factors that differ across individuals and space, that determine the involvement in charitable actions, which cannot be easily harmonized with selfish economic rationality.

Andreoni (2006) proposes five solutions to this conundrum: (1) Certain charitable donations may stem from selfish behavior because giving is aimed at purchasing a certain future service. A person that gifts money to a medical research institute may hope to benefit from its findings one day; (2) The so-called 'enlightened self-interest' according to which individuals contribute with the expectation of receiving help from others in the future. An individual may contribute financially to a poverty-related charity to keep the organization running in the rare case any relative may be impoverished one day. This justification is considered to be a subclass of pure altruism; (3) People may deeply care about the well-being of others in their communities or the society at large, as well as for the environment, and collaborate to fund public or environmental goods (known as 'pure-altruism'). An environmentalist may donate money to save a rare animal species from faraway lands even if she never expects to see it; (4) another reason may arise from an 'impure-altruism' or 'warm-glow' motivation which implies that people get utility from the act of donation itself. A person may have a fixed budget for charity donations that allocates randomly among several organizations to attain a feeling of moral satisfaction; (5) Finally, the last solution would entail the enunciation of a new economic paradigm that integrates moral codes of conduct into economic modeling (Andreoni, 2006). All of these enunciations, except for the fifth solution, can be reconciled with the utility framework by including additional terms in the utility function (Cappellari et al., 2011). Albeit, the (5) solution would entail constructing a more comprehensive human behavior model that accounts for intrinsic moral motivations by integrating psychology and economics (see (Bénabou and Tirole, 2006).

In addition to preferences and attitudes, two other factors determine, to some extent, the charitable contributions of individuals. These two fundamental factors are the fundraising budget of the charity, and the run of informational campaigns about the volume of past donations to a certain charity cause (Cappellari et al., 2011). The first factor appears to give rise to two opposing effects: an increase in the number of donations because of fundraising campaigns, but a subsequent decrease if these campaigns are too costly (Cappellari et al., 2011). Costly campaigns are perceived by contributors as a loss in effectiveness at achieving the true objectives of the charity (Khanna and Sandler, 2000). The former effect seems to prevail over the latter (Khanna and Sandler, 2000). As for the second factor (increasing information about past donations), (Bénabou and Tirole, 2006) suggest that greater publicity induces donations to be perceived as being motivated just by social esteem. Lastly, government behavior has been empirically shown to have a substantial impact on donations. On the one hand, when a public good is provided through voluntary contributions, the allocation of government funds will crowd out voluntary contributions (Duncan, 1999). This crowding effect stems from the fact that according to literature, public and private charity contributions are close substitutes (Cappellari et al., 2011). On the other hand, governments have historically incentivized charitable donations with tax-deduction policies, which have a profound effect on charitable

contributions (Andreoni, 2006). To calculate such effects, tax-price elasticities are utilized, which render the impact of a reduction in tax compensations from charitable donations on the amount of charitable giving by individuals. For instance, Andreoni et al (1996), utilizing a static elasticity approach and a cross-sectional data set, determined that eliminating tax deductibility in the US would result in a 5.7% fall in donations and a 0.7% aggregate reduction in volunteer hours. These figures are significant, but it is important to note that tax-price elasticities greatly vary across studies, depending on the data sets and models utilized. For example, Auten et al. (2002), contrary to Andreoni et al. (1996), advocate for the use of panel data that enables the study of dynamic effects of tax deductions on charity contributions via prices and income effects. The authors find greater effects. For a taxpayer facing a marginal tax rate of 30%, a policy that would eliminate the charitable deduction would induce him/her to decrease charitable contributions by 25 to 36% had the tax code not been changed (Auten et al. 2002). These findings highly exceed the results suggested by Andreoni et al. (1996).

2.4.3 Modelling Charitable Contributions

When studying the nature and dynamics of charitable contributions of time and money, the interpretation of data hugely depends on the model utilized (Duncan, 1999). The two most widely used benchmark models are the "public goods model' and the "private consumption model". According to Duncan (1997), the main difference between these two interpretations of reality lies in the underlying motivations of contributors to give. In the public goods model, a charitable gift is only desirable if it increases the supply of the public good that is provided by the charity (Duncan, 1999). In the private consumption model, the 'warm-glow' effect of giving is sufficient for contributors to give, and as a result, charity is always desirable regardless of the outcome (Duncan, 1999). Often, researchers combine both models so contributors are motivated by what their donations produce as well as by how donating makes them feel (Duncan, 1999).

2.4.4 Empirical Findings

A wide array of studies have applied these different theoretical frameworks, providing a relevant and robust body of literature that shines a bright light on the intricacies of voluntary time and monetary contributions. The central article by Duncan (1999) presents a public goods model that studies the dynamics of an illustrative charity that maximizes the production of the public good given its financial resources, which are provided by contributors of time and money gifts. The author explores several dimensions of a charitable organization that had been studied before, including both time and money contributions, which had been historically considered to be mutually exclusive in charity modeling. Duncan (1999) provides an extension of previous empirical models and creates a creative framework for future empirical endeavors. Albeit, although the empirical framework utilized in the model closely represents economic reality, the data set employed presents several flaws. Firstly, cross-sectional data is utilized, meaning that the evolution of the variables included in the model is not studied, which is especially troublesome for prices and income. Secondly, wages had to be estimated with a Heckman model because they are unobserved for the author. Thirdly, charitable contributions are not observed at the individual level, but rather at the aggregate level, which constrains the precision and depth of the study. Nevertheless, according to the author, the data set obtained from the National Study of Philanthropy in 1976 was one of the few existing sources with detailed descriptions of time and money charity contributions on a national scale in the U.S. The sample data includes 2917 responses, of which 1892 were collected by the Survey Research Center and 1025 by the U.S. Census Bureau using identical questionnaires. The initial descriptive results show that, in the sample, 88% of households gave money, 45% gave time,

and 90% gave either time or money. The author employs the Tobit model to compute the total value of a household's charitable contribution on an ample set of covariates, such as the tax price of contributions (one minus the household's marginal tax rate), income, race, or education among others. The computed value of the dependent variable is calculated by the number of hours volunteered plus contributions of money. Time and money are also employed separately as dependent variables in some specifications.

Among all the results reported by the author, perhaps the most relevant empirical findings are price and income effects for the value, time, and, money specifications. Following previous methodologies, the author calculates price and cross-price elasticities using the net wage of an individual as the volunteering price and one minute an individual's marginal tax rate as the price of money gifts. Before Duncan (1999), several authors had calculated price and cross-price elasticities. For instance, Menchik and Weisbrod (1987) estimate a price elasticity of volunteering of 0.40 (all elasticities are expressed in absolute terms), which implies that an increase in the real wage rate of an individual will translate into a more than proportional decrease in labor contributions to charity. Additionally, Menchik and Weisbrod (1987) estimate a cross-price elasticity between the price of money donations and time contributions of 1.25, which in accordance with other studies, induces them to conclude that time and money contributions are gross complements. Simply put, increasing the price of money contributions (decreasing the tax-deductibility of money donations) would trigger a decrease in volunteering (Menchik and Weisbrod, 1987). Conversely, Duncan (1999) reports a low price elasticity of household contributions of 1.6 and an income elasticity of 0.02. Consequently, results in Duncan (1999) suggest that in equilibrium, contributors perceive charitable gifts of money and time as perfectly substitutable, meaning that an individual will perfectly exchange money gifts for labor supply in charity if the price of money contributions increases. This existing divergence stems from the use of the public goods model, as opposed to the use of the private goods model utilized by Menchik and Weisbrod (1987) and others. Overall, Duncan (1999) proposes an innovative theoretical framework that had a significant impact on future research in the field of time and money contributions and contradicted previous literature that concluded a complementary relationship between time and money gifts to charity.

Smith and Chang (2002) provide new theoretical foundations on the subject of time and money contributions to charity by expanding the canonical model of warm-glow contributions of time and money (private goods model) of Andreoni, Gale, and Scholz (1996). This study, as stated before, estimates comparative static predictions of monetary and volunteer contributions. Smith and Change (2002) base their empirical framework on Andreoni, Gale, and Scholz (1996), but include more than just two goods in the utility function (leisure time, and a composite commodity of goods and services purchased on the market) which allow them to study the indirect and direct interactions across multiple goods. This provides a more realistic view of substitution and complementarity between donations and other elements. Even though this approach was not empirically tested, the authors establish the theoretical possibility that the Hicksian supply of charitable labor may be an increasing function of an individual's wage, given that time and money are complementary in equilibrium. Prior studies (see Andreoni, Gale, and Scholz, 1996; Menchik and Weisbrod, 1987) also conclude that volunteering and money gifts are complementary, but suggest that volunteer labor supply is a decreasing function of wage. The study by Smith and Chang (2002) even though it lacks empirical support, shows that there exists an intricate web of interactions underlying the public and private goods consumption models that are not generally accounted for.

A third article that should be commented on was composed by Feldman (2010). The author revisits how the tax treatment of charitable giving affects individuals who divide his/her resources between time and money gifts. Feldman (2010) utilizes a forthright theoretical

framework that includes the joint consumption of donations of money and time in line with Duncan (1999) and introduces the consumption of alternative marketable goods following Smith and Chang (2002). This allows Feldman (2010) to account for interdependencies in the consumption of not only charitable goods but also alternative marketable goods. Importantly, Feldman (2010), contrary to Duncan (1999) opts for the use of the private consumption model. The data set utilized in the article is from the Independent Sector's biennial U.S., "Giving and Volunteering) from 1996 and 1999, which evaluates monetary contributions to charity in both years, and volunteering activity for the previous 12 months of those respective years. The 1996 and 1999 surveys incorporate 2,719 and 2,553 individuals, all of whom were over 18 years old. Respondents present in the data set were asked about preferences and motivations for charity giving, donating quantities both of time and money, to which organizations they allocated those resources, and several demographic inquiries such as tax-price levels, income, age number of children, or level of religiosity among others. Descriptive statistics show that less than 40% of the sample donated both money and time and that religious organizations were the ones receiving the lion's share of donations (47% of the sample donated money and 23% volunteered) (Feldman, 2010). The author utilizes a bivariate probit model in which the dependent variables are binary indicators of whether or not a household gifted money or time to at least one of the 12 charitable organization categories contemplated in the data set. Results of the empirical application of the model generate conclusions of interest. The study shows that an increase of $10,000 in income from the sample mean is associated with an increase in the probability of donating money of 0.023 at the 1% confidence level, and an increase of 0.019 in the probability of supplying charity labor, at the 5% confidence level. These results contradict previous studies that presume time charity contributions should be negatively correlated to increases in income. Conversely, if we consider volunteering as a normal good, according to economic rationale, consumption of it should increase as income rises (Feldman, 2010).

Additionally, there is empirical evidence that membership in charitable organizations augments with increases in social status, which is closely correlated to income levels (Feldman, 2010). Among the independent variables considered in the article, only employment status and the number of children present a positive significant correlation to volunteering. These two variables and education do not have a significant impact on monetary contributions. At last, owning a house, regularly attending religious events, past membership in youth groups, and having grown up with parents that volunteered presents a positive significant correlation to both types of charity contributions. The paper also sheds a light on the long-standing debate of the intrinsic relation between time and money contributions in equilibrium conditions. The author utilizes the joint consumption of time and money donations and estimates that donations of time and money are substitutes, contrary to the previous literature suggesting that time and money gifts are gross complements. Consequently, a reduction in the tax-deductibility of money donations to charity catalyzes an increase in the probability of money contributions but depresses the probability that the average household volunteers (Feldman, 2010). Even so, the author recognizes the existence of positive effects on time contributions that arise from intrinsic motivation to volunteer, increases the information on a charity's objectives, households' alignment with those objectives, and/or decreases volunteer costs (Feldman, 2010). These effects may offset the impact of a reduction in the tax deduction on time contributions. Nevertheless, Feldman (2010) presents compelling evidence of the substitutable nature of time and money contributions to charity.

At last, an important paper that shall be discussed is the one by Cappellari et al. (2011) on time and money contributions in Europe. Previous literature on this topic has been largely devoted to the computation of the tax-price elasticity of charitable labor supply and money gifts in the U.S., mainly because no general tax benefits for donors existed in Europe before 2010

(Cappellari et al., 2011). Wright (2001) explains the existence of the institutional divergence with cultural differences in the role of the State in the economy and the attitudinal perceptions towards money and wealth. As a result, Wright (2001) concludes that institutional frameworks and cultural differences shaped the reasons behind charity giving on both sides of the Atlantic, with Europeans not giving to charity to reduce their tax rate but for alternative intrinsic motives (Wright, 2001). Seems evident that the unavailability of European survey data on charity labor supply and money gifts to charity can partially cast a light on the lack of scrutiny devoted to the development of an exhaustive behavioral model which accounts for individual preferences and attitudes as opposed to monetary/tax incentives (Cappellari et al., 2011). Thus, the article by Cappelari et al. (2011) discusses the role of attitudes for giving on individual behavior. The authors utilize the standard static labor supply framework of the private consumption goods model, accounting for both time and money contributions, and for domestic work. Cappelari et al. (2011) argue that there are two main drivers of charity giving, the warm-glow effect, and the social signaling motive. The former implies an individual that derives utility from the very act of giving, while the latter entails that giving is motivated by the hope to signal generosity and to receive social esteem (Cappelari et al., 2011). As discussed by Smith and Chang (2002), any model related to time donations should account for the fact that individuals generally have other alternatives to non-market labor so Cappelari et al. (2011) take into consideration not only hours of volunteer labor but also domestic work hours and leisure time in a joint utility function. For simplicity, the authors terminate the long-standing debate and, assume that individuals do not have strict preferences for time versus money donations and consider the two forms of donation as perfect substitutes.

In the article, the authors make use of a cross-sectional survey yearly administered by the Italian National Statistical Office to a representative sample of the entire Italian population. The sampling unit of analysis is at the household level and the survey scheme provides a very rich source of information at the micro-level of Italian households (Cappellari et al., 2011). After clearing, the authors report a resulting sample of 11,331 men and 11,038 women, which for instance, report employment rates of 85% and 54% respectively, or 97% of the females reporting domestic work, while only 57% of the males do. On the money and time charity gifts, survey participants are asked about their contributions in the last 12 months before the interview (Cappellari et al., 2011). Sample data provides insightful information on the probability of charity donations by males and females. Firstly, money donations are predominantly preferred among sampled individuals, and males donate more, both money and time. Secondly, the overall majority of the sample does not donate, Only around 20% of the sample donated money in the last 12 months, while just around 10% of the sample reveals having volunteered. Thirdly, when looking at conditional probabilities, being a volunteer increases the probability of donating money approximately by four times if compared to an individual who does not volunteer. Additionally, volunteerism is augmented six-fold if money donors are weighed against non-donors of money. Nonetheless, the authors advise that sample probabilities do not reflect true correlations between time and money gifts due to the "compositional effects" that afflict sample descriptives (Cappellari et al. 2011). A more in-depth analysis demands multivariate regression analysis (Cappellari et al. 2011).

The authors utilize a simultaneous equations model that studies the relationship between the four fundamental dependent variables (money and time donations, hours of domestic work, and market work. As mentioned before, the paper aims to investigate the role that impure altruistic attitudes play in explaining charity donations. More specifically, the article explores the effect of warm glow and social esteem aspirations. Since these characteristics are unobserved, the authors employ two proxy explanatory dummy variables that represent the absence of impure altruism. Firstly, those individuals that report seeing friends less than one time per month are considered to place a lesser weight on the importance of social signaling via

donations. An individual with an extensive social network is deemed more likely to be impurely altruistic. Secondly, based on previous studies, the authors utilize religiosity as a proxy to measure the degree of warm-glow donations. An individual who adheres to religious norms is more likely to engage in donations based on warm-glow reasoning. Nevertheless, the authors recognize that this dummy variable may be troubling because religious people that are very involved in religious activities are very concerned with social reputation inside the community, and consequently, this proxy dummy variable may capture both warm-glow and social aspirations. Moreover, a wide set of covariates is included such as the number of children, education, estimations of income, health, and others. Finally, the authors have discrete information on time and money donations (whether or not to donate) and continuous data on domestic and market work hours (specified in the logarithmic form). The authors utilize two Probit and two Tobit models, allowing for unobserved correlations between specifications in the error term.

The results obtained by the regressions are consistent with the hypothesis formulated by the authors and offer a very rich understanding of existing sample processes. Both proxy dummy variables (religiosity and social network) significantly affect the provision of charitable donations. Rarely meeting friends (small social network) diminishes the probability of volunteering by 4.5% in the case of females, and by 6.2% in the case of males; while the probability of donating money decreases by 2.9% for women and by 5.3% for men (Cappelari et al. 2011). Similar results are found for non-religious individuals. For them, the probability of volunteering decreases by 2.7% for women, and by 3.1% for men (Cappelari et al. 2011). All coefficients are statistically significant at the 1% level. This evidence suggests that there exists substantial variability across genders in the decision to donate money and/or time. Finally, looking at the impact of other control variables the probability of donating both time and money is generally increasing in education level and age for both men and women. Moreover, residents from Northern (wealthier and more developed) regions are more prone to donate than Southerns, which is in accordance with previous Italian literature that points to the fact that higher income levels and the existence of non-for-profit organizations are more prevalent in the North (Cappelari et al. 2010). Unsurprisingly, domestic and market work hours present a negative relationship with volunteering (Cappelari et al. 2011). Concerning inter-dependencies between time and money gifts, results from unobserved attitudes, show a slight positive correlation between them, which may not signal complementarity, but rather a preference for combined undertaking or conditionality (Capellari et al 2011). Accordingly, and aligned with Feldman (2010), substitution effects are present when evaluating price and income effects, which suggest that, on the one hand, substitution effects between time and money donations are in line with standard microeconomic theory, but show on the other hand that there are strong direct complementary processes that may reverse the net effect (Cappelari et al. 2011). Overall, this paper proposes a comprehensive behavioral model of time and money gifts that stresses the importance of demographic factors, attitudes, perceptions, and impure altruistic motives in the decision process of donating, while exploiting high-quality revealed sample data.

2.4.5 Conclusions

Contrary to the neoclassical economic standpoint, both at the micro and macro level, world economic systems operate based on a moral foundation that underlies every market and non-market transaction. One clear example is the existence of charities that privately fund the provision of a public good. These not-for-profit organizations generally carry out their activities thanks to volunteer contributions of money and time. The occurrence of these organizations gave birth to a vast and muscular body of literature that studies all the complexities of charitable organizations and donations and proposes sound models to optimize

their functioning. But, what induces an agent to willfully donate their money and/or time to a cause that may not generate any benefit for him? According to research, two fundamental processes motivate a person to donate: pure and impure altruism (Andreoni, 2006). For instance, Cappelari et al. (2010) show that warm-glow and social aspirations substantially affect the decision to donate, proving that volunteerism and donations may be intrinsically stimulated by many different factors. Additionally, many empirical articles have been published aimed at discovering the relationship between time and money contributions to charity, the elasticities of price and income derived from charity contributions, and the existing heterogeneity in donators in terms of gender, age, or religiosity. Many decades of empirical research established the assumption that money and time are complementary goods, which move jointly in equilibrium conditions (see for instance Menchik and Weisbrod, 1987). To empirically test this assumption, donations in the U.S. subject to tax deductibility are utilized. In other words, they observe how time and money contributions react to an increase or decrease in the tax rate of contributors which resulted in the conclusion of complementarity. However, with time, and the refinement of models and data sets, alternative empirical results suggested that, if time and money contributions are considered to be normal goods, real-world data show that they are perfect substitutes (see Duncan, 1999). Agents perfectly exchange money gifts for labor supply to charity and the other way around. This assumption appears to be solid and has been generally accepted, even though some effects may impact this relationship and the outcome of a complementarity relationship (Feldman, 2010; Cappelari, 2010).

3 Materials and Methods

In the present investigation, the selected methodologies are the CV and CB survey schemes and payments and frequencies are elicited with the use of the payment card format. Time is measured with the concept branded as WTC, and monetary contributions are gauged in terms of WTP. Essentially, participants are questioned "how many hours per year are you WTC" in the case of time contributions, and "how many euros are you WTP" in the case of monetary gifts. Both instruments would outcome the same hypothetical scenario, which is a "clean city" free of litter. The "dirty city" represents the status quo in Zaragoza and the environmental change is measured in time and monetary terms. In this chapter, the intricacies of the econometric treatment of results will be discussed, justifying every decision made to reduce biasedness and inconsistency of results. There are 3 sub-sections in this chapter: the first two explain the statistical treatment of results for both the time and monetary contributions retrieved from the sample data; the last sub-section, briefly explains the sample mean model utilized to measure the aggregate welfare effects of the proposed policy.

3.1 Time Dimension

To correctly analyze the data obtained in the study's time dimension, several aspects must be emphasized. Firstly, zero frequency bids stated in the survey are considered to be true zero bids, which means that the statistical treatment of this sample data is conducted differently from the money dimension of the study. All respondents that are not willing to cleanup, would not do so, even if the contextual elements of the survey change so their negative to participate in urban clean-ups would be constant across different scenarios. Consequently, zero bids can be taken at face value for the statistical examination. Following Brouwer et al. (2017), two models are utilized to analyze the data. Firstly, a binary Probit participation model to explore the dominant factors driving the decision to be willing to participate in urban clean-ups or not. Thereby, the dependent variable, is dichotomous - whether you want to participate or not - that will be determined by the set of sample characteristics. Nonlinear regression models like the Probit are precisely designed for binary dependent variables (Stock and Watson, 2015). Since the outcome variable is dichotomous, the regression models the probability that Y=1, which means that the respondents are willing to participate in cleanups. Thereby, adopting a Linear Probability Model (LPM) is not appropriate. According to Koster (2019), there are three primary flaws related to the use of the LPM in binary choice regressions: First, the LPM "can not estimate the structural parameters of a non-linear model" (Koster, 2019, p.11). As a result, the LPM may result in wrong predictions above 1 or below 0; secondly, the LPM will produce inconsistent marginal effects; and lastly, if a measurement error exists in the dependent variable, the LPM can not correct it. Consequently, a nonlinear regression model, using a cumulative normal distribution function is required to force the predicted probability values to be between 0 and 1 (Stock and Watson, 2015).

In the present study, the specification of the multivariate Probit regression model with several regressors is:

$$Pr(Y = 1 \mid x_k) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_k)$$

where $\Phi(\cdot)$ is the standard cumulative normal distribution function, β_0 and $\beta_1, \ldots, \beta_m$, respectively represents the intercept and the coefficients of the independent variables included in the model. To compute the probability of Y=1 (willingness to participate) given the values

of the exploratory variables, the z-score is computed, which is the number of standard deviations from the mean (Stock and Watson, 2015). The marginal effects will be given by:

$$\frac{\Delta Pr(Y = 1 \mid x_k)}{\Delta x_k} = \beta\phi(\beta' x_k)$$

where ϕ is the density of the normal distribution. The model will be estimated using the maximum likelihood estimator (MLE) which utilizes the unknown values of the coefficients that maximize the likelihood function (Stock and Watson, 2015). Thus, the MLEs are "the parameter values most likely to have produced the data" (Stock and Watson, 2015; p.405).

Secondly, after the determinants of participation probability are modeled, participants were also questioned about the preferred frequency of participation. As previously mentioned, the desired frequency is elicited with the use of a payment card ranging from 0 to 36 hours per year. To explain the variation in the number of hours that citizens are willing to participate in urban clean-ups, a multivariate Tobit regression model is used. According to Halstead et al. (1991), when studying the determinants of variation in a CVM context, Tobit analysis provides better estimation results than using OLS. Furthermore, the authors prove that the Tobit model is a better-suited regression technique for data sets with large numbers of zero bids, like the present one. The Tobit model also enables the decomposition of the data set to study more in detail the effects of the exploratory variables on existing non-zero respondents and estimates the chances of zero bidders changing to positive frequencies based on variation in their characteristics (Halstead et al., 1991). The statistical application of the latter feature of the model provides a safety net for the true-zero frequency bidders' assumption of the study. If the model identifies a respondent that could participate in the clean-ups based on the independent variables, it will estimate the positive frequency. Lastly, the Tobit regression technique models the censored nature of the outcome variable, which only has zero or positive values (Brouwer et al., 2017). Once again, the MLE technique will be applied to estimate the unknown coefficients, which will be unbiased and consistent parameter estimates, whilst more information than in the binary Probit model is incorporated (Halstead et al., 1991). Essentially, the Tobit model assumes that:

$$E(Y_j) = X_k'\beta$$

where the expected value of the frequency bid selected by the respondent is given by the set of independent variables used in the Probit model, X_k', and its respective coefficients β. Then, the estimates of the model are derived from:

$$E(Y_j) = X_k'\beta F(z) + \sigma f(z)$$

where $F(z)$ is the cumulative standard normal distribution function; $f(z)$ is the standard normal density function of a normal, random variable with mean zero and variance σ^2; z is the normalized Tobit index: $\frac{X\beta}{\sigma}$; and σ is the standard error of the regression. The Tobit model can also be expressed as:

$$Y_j = \begin{cases} X_k'\beta + e & \text{if } Y_j > 0 \\ 0 & \text{if } otherwise \end{cases}$$

Both models were computed using the statistical software Stata 17.0 by utilizing the commands *probit* and *tobit*, and correcting for the robustness in the standard errors.

3.2 Monetary Dimension

The econometric strategy followed for studying money contribution has been cemented on the existence of protestors, that comprise the major share of zero bids. As discussed prior, CV studies tend to suffer from a high percentage of protestors as if it was a physiological component of the methodology. Protestors in the literature are assumed to have a positive WTP (Collins and Rosenberger, 2007; Strazzera et al., 2001; Strazzera et al., 2003; and Brouwer and Martín-Ortega, 2012) but those respondents decide not to reveal their true valuation because they react to some component of the questionnaire. On the contrary, true zero bids are considered to be genuine and legitimate zero valuation bids. Generally, protest responses are excluded from the analysis but this approach is very likely to create a selectivity bias (Calia and Strazzera, 2001) and a loss of potentially useful information (Collins and Rosenberger, 2007). A selectivity bias will arise when the decision to protest is not random, and the two subsamples (protestors and non-protestors) exhibit distinct characteristics in terms of socio-demographics or attitudes and perceptions. In the present study, notable differences between subsamples were discovered in the discriminant analysis, which will be covered in the coming sections. As a result, a statistical model was required to correctly treat each type of zero bid.

The econometric strategy is divided into two phases and was conducted following Strazzera et al. (2001), Strazzera et al. (2003), and Collins and Rosenberg (2007): firstly, a Tobit model, excluding zero bids is estimated, deliberately incurring in a sample selection bias to expose its existence. Afterward, a Full Information Maximum Likelihood (FIML) model is computed, accounting for all the observations in the sample, and correcting for the selectivity bias. The FIML sample selection model integrates protest responses via two equations that model the joint decision process of a participant, which involves the choice of a monetary value and the decision of whether to reveal the bid or not. Using a FIML sample selection model enables the accurate estimation of positive bids from protest zero bids based on respondents' observed socio-demographics, attitudes, and perceptions (Collins and Rosenberger, 2007). In other words, if a protestor's characteristics closely resemble those of a positive bidder, the FIML model will mimic the bidding process of the non-protestor and compute a positive value for the protestor (Collins and Rosenberger, 2007).

Let Y_1 represent the monetary value that a respondent places on the environmental good, let Y_2 be a binary variable that takes the value 0 if the individual decides to not reveal the amount and 1 otherwise, and let x and z be vectors of independent variables for the two-equation system. The first WTP equation can be written as:

$$lnY_{1i} = x_i'\beta + \sigma u_i$$

where σ is a scale factor. Then, Y_{1i} is only observed when $Y_{2i} = 1$, and the binary model is expressed as:

$$Y_{2i} = \begin{cases} 1 & \text{if } z_i'\gamma + \epsilon_i \geq 0 \\ 0 & \text{if } z_i'\gamma + \epsilon_i < 0 \end{cases}$$

representing the decision equations. The joint distribution (u_i, ϵ_i) is assumed to be bivariate normal with zero means, variance equal to 1, and correlation ρ. When $\rho = 0$, the two decisions are independent and the parameters of the two equations can be estimated individually.

The FIML model is then given by:

$$l = \sum_{Y_{2i}=0} (1 - I_i) ln\Phi(-z_i'\gamma) + \sum_{Y_{2i}=0} I_i ln\Phi\left(\frac{lnY_{i1} - x_i'\beta}{\sigma}\right) + I_i ln\Phi\left(\frac{z_i\gamma + \rho\left(\frac{(lnY_{1i} - x_i'\beta)}{\sigma}\right)}{\sqrt{1-\rho^2}}\right) - I_i ln\sigma$$

Maximization of this function generates a simultaneous estimation of the parameters of both the participation and WTP equations. The maximization of the FIML model function was computed in Stata 17.0 utilizing the command *sem*.

3.3 Sample Mean Model: Welfare Estimates

The sample mean of all monetary and time bids of the sample is estimated by the following equation:

$$\overline{b} = \frac{\sum_{i=1}^{b}}{n}$$

where b is the bid selected by the participant. The sample means are calculated separately for time and money contributions and then compared by estimating the average income level in the sample, to assess which one results in higher economic benefits.

4 Data

This chapter discusses the data set utilized in the present work. The first section contains a description of the questionnaire and explanations for the decisions undertaken in the design process. The second section expounds on the survey pretesting process, the sampling technique employed, the data collection procedure, and the survey mode selected. The third section presents a description of the variables employed in the multivariate regression analysis and explicates the rationale behind the selection and measurement unit of variables. In the fourth section, the sample is studied by conducting a general descriptive analysis of the intricacies in the observed data. Finally, the fifth section includes a discriminant analysis of protest and non-protest bids aimed at studying sample differences between these two subgroups, to determine the existence of a potential selectivity bias if protest bids are censored from the study.

4.1 Questionnaire

The questionnaire was majorly designed according to the guidelines of Johnston et al. (2017) and Arrow et al. (1993) which are the backbone of the study. Both manuals enable the study to be in accordance with existing peer-reviewed practices in the field of contingent valuation and contingent behavior. The aim was to create a survey scheme that minimizes potential biases in value estimates to maximize the content and construct validity of the study. All ex-ante procedures previously discussed in the literature were taken into account to minimize the potential biases that could jeopardize the validity of the results, in addition to elements from relevant articles in the field of CV and CB methods. The following subsections describe the complexities of each of the blocks that form the questionnaire.

Introduction and Attitudes and Preferences' Enquiries

The first block of the questionnaire is formed by the introduction to the questionnaire and four inquiries regarding attitudes and perceptions towards existing urban litter in Zaragoza, the protection of the environment, and littering. The introduction section is composed of a welcoming slide that states the conditions and assurances of participation. The participant is ensured of the absolute anonymity of participation in the survey. No personal data was retrieved unless explicitly specified by the respondent, and all responses were permanently secured on a password-protected platform with encryption. The participant is informed that she would not be able to change any stated response and petitioned to ponder thoughtfully before answering any of the inquiries. After acceptance of the participation conditions and assurances, the participant faces questions about attitudes and perceptions. This question set is purposely placed at the beginning of the survey to obtain pure and unbiased responses on environmental attitudes and perceptions to be able to study the statistical association with stated WTP and WTC in the next sections of the survey. Nonetheless, this placement decision lacks an empirical basis because no previous studies exist on the matter. Further exploration is needed to assess the potential biases created by the placement of questions about attitudes and perceptions at the beginning of a survey, on posterior stated responses. Additionally, the placement of this set of questions at the end of the survey seems to be problematic as well. Stated bids may anchor participants and condition their responses to the question set about attitudes, perceptions, and preferences. It seems necessary that experimental studies should be conducted to determine the magnitude of bias and the optimal place in the survey for this set of questions.

The first question is a multiple choice format inquiry about the participant's level of satisfaction with perceived cleanliness in Zaragoza's urban spaces. The participant can choose between (a) very dissatisfied, (b) dissatisfied, (c) satisfied, and (d) very satisfied. Next, the participant faces three five-point Likert scale questions clustered together in the same slide about whether the responders find themselves described by the statements provided. The Likert scale format was selected to increase the information available about each respondent's opinion. The first statement "minimizing environmental degradation has to be a political priority" is aimed at measuring how important the environment is for participants and whether it should be one of the top priorities in the political agenda. For the analysis, this variable is named"priority". Subsequently, respondents are asked to state how bothered they are when they see somebody littering a cigarette butt. Littering cigarette butts is a cultural behavior that is not generally punished socially nor institutionally so stating high punctuation on the Likert scale, signals a strong pro-environmental attitude and a high level of awareness concerning the environmental damage caused by cigarette butts (or any type of non-biodegradable residue). This variable is coded as "butt". Finally, to measure revealed behavior on daily cleanup efforts, the following statement is enunciated: "If I happen to see a littered can in Zaragoza, I always pick it up for correct disposal in the closest bin". It is coded as "cleanup". A respondent that is perfectly defined by this statement prompts a very high level of commitment to tackling waste pollution at the individual level and also an altruistic stance against environmental and urban degradation.

Scenario Descriptions

SP surveys capture value estimates of changes in economic welfare caused by an alteration of the status quo (Johnston et al., 2017) and it is fundamental to precisely describe the proposed change undervaluation, relative to the baseline conditions. A precise description shall include both factual and subjective information potentially perceived by the respondent, the degree of uncertainty concerning the achievement of the proposed change, the spatial location of the proposed change, and the scale of the proposed change (Johnston et al., 2017). This information block must be "understood, accepted, and viewed as credible by respondents" to capture the true economic welfare change (Johnston et al., 2017; p.326). In the present study, the status quo conditions are related to the actual state of Zaragoza in terms of cleanliness and waste contamination and the scenario is called for simplicity "dirty city". It should be noted that, even though the construction of the baseline scenario is rooted in real-world conditions, it is still contingent on the attitudes, perceptions, and information level of the participant. An individual may not perceive the city as dirty if she is not informed about the adverse consequences of litter and as a result, does not consider waste pollution to be a problem. The participant may also have low standards for urban cleanliness. According to Johnston et al. (2017), state-of-the-art contingent surveys shall include factual data on the baseline conditions and the conditions proposed to maximize objectivity in the hypothetical scenario. Since data on urban waste pollution in Zaragoza does not exist, explanations were conducted only with the support of images of littered places in the city.

The proposed change in the baseline conditions of the "dirty city" is an environmental and aesthetic improvement, which produces a "clean city". The second scenario is presented as an alternative and opposed setting where all the problems associated with the status quo are avoided. The aim is to induce the participant to visualize the impact of the proposed change in the level of cleanliness in the city, and what would imply for his/her utility in measurable terms. Thereafter, the survey explains that the proposed change is to be attained by the implementation of the "New Cleaning Plan" and it is presented as a comprehensive policy promoted by the city council to tackle the environmental problem effectively and efficiently. The primary objective of the hypothetical policy is to remove all existing and future residues

improperly discarded that pollute streets, rivers, parks, and squares. Note that the plan proposed, represents the mechanism of change that delivers the improved scenario with certainty. Furthermore, it is clearly stressed that the policy only targets the cleaning of the city as the first step towards a paradigmatic change in the city, that would solve the problem in the short and medium term, but not in the long term since the root cause of littering behavior can only be solved via educational campaigns. This additional information was given to avoid protest demeanor by some better-informed participants. It should be highlighted the fact that the plan is explicitly presented as binding and mandatory to secure incentive compatibility, and consequentialism, and to prevent free riding. The "New Cleaning Plan" is presented to be a certain policy intervention and the overarching mechanism of change, and the referendum survey is aimed at determining the optimal implementation based on citizens' will.

Payment and Frequency Vehicles

In the fourth block of the survey, once the participant has comprehended the baseline conditions, the proposed change, and the mechanism of change, both policy instruments (contained in the cleaning plan) are unveiled. It is important, that the participant is not informed about the two instruments included in the policy until she faces the block dedicated to that specific policy measure. The unveiling tempo of the information is very relevant to avoid potential biases before the valuation inquiries. In this regard, and to control for unobserved factors that lead to conditioned responses, the participants face both policies, in a randomized order, e.g. a participant may respond to the payment question first and next to the frequency question, while another one answers them in the opposite order. The software platform utilized allows control for the randomization process so the two combinations are equally displayed across the entire sample. Albeit, the software does not allow to construct a binary variable that explores the potential "order effect" of the two policies.

The two policy measures contained in the new clean program are, in the first place, the creation of a green tax levied on all households residing in Zaragoza. The green tax, collected every four months, is to be utilized to fund the creation of a new cleaning brigade that would patrol Zaragoza daily, cleaning all the existing residues. It is also communicated that those families which are at risk of economic exclusion are to be exempted from the payment. This last piece of information is added to increase the credibility of the hypothetical policy. It is remarked that the "clean city" scenario would be achieved with the implementation of this policy measure with total certainty. Thereafter, the participant is informed that the city council would like to know his/her WTP for the green tax to set the optimal tax level according to citizenship preference. Lastly, the respondent is reminded of the existing opportunity cost in the decision of paying the green tax because the money individually allocated to the payment of the green tax inevitably leads to a decrease in income that should be compensated elsewhere (e.g. leisure activities such as going to the cinema). It is an attempt to increase truism in responses and reduce inflated responses. The valuation question response format selected is a payment card, as it is the most effective way to eliminate starting point bias and anchoring effect on initial bid amounts. The bids presented are: 0€, 1€, 2€, 5€, 10€, 15€, 20€, 25€, and 30€. The selection of bids was delimited by the responses obtained in the focus groups and the bid range is limited to a maximum of 30€ to avoid outrageously high responses which could be subject to bias. The 0€ bid is included to allow participants to reject the green tax scenario and are asked a follow-up question to understand this decision. Recall that the zero bid may imply a protest against some element of the survey or a legitimate zero bid that places a zero value on urban cleanliness. The participants are given several options: (a) urban waste pollution is not a priority for me; (b) I pay enough taxes; (c) I can not afford to pay the tax; (d) I believe the city council should utilize public money that is already collected more effectively and efficiently; (e) other reasons.

Respondents that claim to have other reasons to reject the green tax are directed to an open-ended question on their specific reasoning.

The second policy strategy contained in the new cleaning program is the creation of a cleaning brigade composed of citizens that would conduct the same activities as the professional group. The questionnaire unmistakably explains that this policy entails the contribution of labor. The city council would then create a cleaning brigade for each neighborhood and residents would be in charge of cleaning excess residues every week, based on their availability. The participants are then informed that the city council is interested in surveying their disposition to help and the frequency of intended participation. To increase realism, the questionnaire states that the equipment and additional expenses required would be financed by local businesses. Following the same procedure as in the WTP scenario, a slide was included before the frequency card reminding participants of their time constraints. Becoming a member of the cleanup brigade would have to be a priority for them, and would demand a share of their leisure time. Again, the intention is to induce participants to respond more realistically. The frequencies are presented in a payment card format with the following options: "I am not willing to help", "once every six months", "once every four months", "once every two months", "once every month", and "once every two weeks". These frequencies were also set in accordance with the results obtained in the focus group sessions and once again, there is a limit on the frequency that the participant can select to avoid inflated responses. Those individuals that select the option "I am not willing to participate are petitioned to specify the reason why. The participant is given several options: (a) I have physical problems"; (b) urban waste pollution is not a priority for me; (c) I am not willing to collect litter generated by others"; (d) I do not have time; (c) other reasons. Those respondents who select option (d) are asked to state their specific justification in an open-ended question. As can be observed, the presentation and questioning of the WTP and WTC scenarios are conducted following a parallel and identical structure and wording. According to the little empirical evidence found, it seems likely that CB and CV surveys suffer from the same biases and therefore, the best approach to combat them is using a parallel structure. However, this lacks a solid empirical foundation, which should be explored in future research.

Demographics and Lottery Participation

In the last part of the questionnaire, the participants are requested to provide personal information on several dimensions that were found to be relevant to this type of research. The socio-economic variables controlled in the questionnaire are gender, age, level of education, district of residence, and reported household gross monthly income. In the last demographic question, respondents are allowed to not reveal their income. Lastly, as promised in the message included in the distribution of the survey link, participants that are interested in participating in the lottery of 30€ are permitted to give an e-mail address. The participants are reassured that the e-mail address will only be used for conducting the lottery and contacting the winner for the transfer of money. These participants give also consent to receive the results of the study.

4.2 Data Collection

Survey Pretesting

The questionnaire's quantitative and qualitative quality was pretested in two different focus groups composed of 8 and 13 participants respectively. In accordance with Johnston et al. (2017) all participants were selected carefully, to obtain a representative pretesting sample of the target population, Zaragoza. The focus groups were initiated by a brief introduction to the

study, the objective of the session, an enumeration of the different phases of the session, and an expression of gratitude for participation. It was explicitly stated that the objective was to refine the questionnaire and the participants were encouraged to write down any doubt or suggestion they had while completing the questionnaire. In the second phase, the draft questionnaire was administered. An open-ended elicitation format was utilized, so participants were permitted to evaluate the environmental change without any bid range constraint. The objective was to determine the design of the payment card in accordance with the pretested sample. As soon as all participants had finished the completion of the survey, the third phase was initiated. By rising their hands, all participants had the opportunity to express their views about the survey. Once everybody had given feedback, another questionnaire was provided so participants would supply desired comments on subjects such as the level of fatigue experienced during the survey, the credibility of the hypothetical scenarios, suitability of demographic questions, etc. After the two focus groups were carried out, a final version of the questionnaire was created and tested again on a concluding focus group formed of 6 persons. The participants were not the same as in the previous two focus groups to test the validity of the feedback given by the first two focus groups. The session was identically conducted to the two other focus groups and participants only highlighted minor issues that were easily resolved afterward. Overall, the focus groups proved to be remarkably beneficial for the development of the final survey version. After the pretesting phase of the first two groups, the survey was clearly understandable and plausible to respondents in the third focus group, all unnecessary information was eliminated and the bid and frequency vector was identified.

Sampling Procedure, Collection of Data, and Survey Mode

For the collection of data, the snowball and convenience sampling procedure was utilized. Snowball sampling entails the selection of a small number of initial contacts who fit the sampling requirements and are invited to complete the questionnaire (Parker et al., 2019). Then, those participants are trained and petitioned to further disseminate the survey among potential respondents from their social networks (Parker et al., 2019), creating an entire distribution web out of a smaller one. Although it is a very effective and practical sampling procedure, the snowball technique is subject to a substantial degree of criticism summarized in two main arguments: firstly, it deviates from probability-based sampling techniques, which hinders the capacity of the sample to meet random sampling criteria; secondly, the snowballing technique greatly hampers the ability to obtain a representative sample of the population of interest, impeding the extrapolation of results (Parker et al., 2019). Interestingly, Noy (2008) suggests that snowballing samples are likely to be overrepresented by women due to their propensity to be more cooperative and inclined to help.

Specifically, the network-based sampling procedure for data collection in the present study was conducted in two simultaneous phases grounded on the two main distribution channels at disposal. Firstly, the survey was distributed via WhatsApp to relatives and friends who were fixed as the initial nodes. Along with the survey link, a message was included briefly explaining the study (emphasizing that the targeted participants were residents of Zaragoza) and the existence of a participation lottery, which set the economic incentive to respond. The lottery was conducted among all participants that wittingly provided their email addresses and the winner got 30€. More than half of the sample gave the email address (a total of 256 respondents), but it is unclear the extent to which the lottery increased participation rates. For future research, it would be interesting to empirically test the impact of economic incentives on response rates in CV and CB surveys. Secondly, the message including the link was posted on several private Facebook groups of different districts in Zaragoza and groups related to news of the city of Zaragoza. The survey remained open for ten days until the target sample size (more than 400 valid responses) was reached. The number of acceptable responses obtained

was 447. Overall, the snowball sampling technique proved to be very effective as the initial links created a solid sampling momentum that peaked after three days. After that, the response rate was lower but still contributed substantially to the final sampling outcome.

Concerning the survey mode employed, Johnston et al. (2017) argue that internet-based methods for surveys present several disadvantages that should be taken into consideration: firstly, there may be differing levels of computer literacy between age groups that would certainly impede certain individuals to participate in the survey; secondly, certain sociodemographic or economic groups may not have access to the internet and consequently not accounted for in the study; lastly, participants are more likely to respond vehemently in internet-based surveys. In the article by Lindhjem and Navrud, (2011), the authors explore how internet surveying may influence stated preferences in the measurement of environmental goods. The authors set up an experiment as part of a national CV survey aimed at estimating WTP for a policy related to environmental protection by creating two groups drawn from the same population. The first group completed the survey in a face-to-face interview and the second responded to the same questionnaire online. The researchers did not find any evidence of social desirability bias in the interviewed group or reduced-response time in the internet-based group. Moreover, the non-response rates, the share of "don't know", zeros, and protest responses in the WTP question were practically identical between both groups. Hence, these results support the use of the internet survey modes for the valuation of environmental goods as using internet-based survey schemes does not appear to bias the results (Lindhjem and Navrud, 2011).

4.3 Description of Variables

The independent variables presented below, in table 1, constitute the variables of interest in the study and are utilized in all econometric specifications across time and monetary contributions. In the survey, as mentioned before, a larger number of variables was collected from participants. However, some of them, such as "gender" or "butt" are only useful for the descriptive analysis of the sample and are eliminated from the econometric analysis of results. This decision was taken because those variables do not have any statistical significance and are not reliable controls. The only control variable that has been unconditionally excluded from the entire analysis is the "district of residence". It was decided because it does not provide any additional information to the average reader nor statistical significance as a spatial control. The set of variables of interest presented below had to be rearranged to maximize its statistical significance. Two-sided variables were created for 'satisfaction', 'education', and 'income' and are displayed in table 1. The reason for doing so is that some categories of those variables are overrepresented and problematic when conducting statistical estimations. In spite of this, if those variables are rationally transformed into dichotomous variables, some degree of statistical significance can be achieved, or in the case of "middle-high income", can be used as a valuable control variable. It is important to note that the null side of satisfaction also includes indifferent participants that believe the city is not clean or dirty. After a thorough analysis of observations, it was clear that the tendency of indifferent participants was to, later on, show WTP and WTC. Anyhow, "satisfaction" is not further utilized for econometric inference because it lacked statistical significance. Regarding "age group"', some minor readjustments are made to achieve more representativeness in some categories. The rationale will be further explained in the next section. Finally "priority" and "cleanup" are continuous variables ranging from 0 to 5 that measure pro-environmental attitudes and behaviors.

Table 1*: Description of exploratory variables utilized in the econometric analysis

Variable	Description	Unit of measurement
Satisfaction	If respondent is satisfied or very satisfied (categories 4 and 5) with overall urban cleanliness in the city.	Dichotomous variable: 0 = no; 1 = yes
Priority	How well are you described by: "Minimizing environmental deterioration should be a political priority".	Likert-scale: 0 (not described at all) to 5 (perfectly described)
Cleanup	How well are you described by: ""If I happen to see litter, I always collect it for proper disposal".	Likert-scale: 0 (not described at all) to 5 (perfectly described)
Highly educated	If the respondent holds a university degree or postdoctoral studies.	Dichotomous variable: 0 = no; 1 = yes
Age group	Age interval categories.	Ordinal variable: 1 to 5
Middle-high income	If respondent states a household gross monthly income higher than 5,001 euros.	Dichotomous variable: 0 = no; 1 = yes

4.4 General Descriptive Statistics

4.4.1 Socio-Demographic Characteristics

A total of 447 valid responses were recorded and utilized, while 220 incomplete questionnaires were excluded from the initial sample data. The main characteristics of the participants are displayed in table 2. In terms of gender, approximately one-third of the sample is constituted by men, while around 65% is represented by women. This considerable disparity is in line with Noy (2008), which suggests that samples obtained with the snowball procedure tend to be overrepresented by women, with respect to men. This sample feature is problematic and has impeded further econometric analysis of this variable. Only 3 participants stated "other" as their gender. Concerning age, the variable is measured in intervals, which had to be rearranged after retrieving the sample data to increase the number of participants in each of them. Changes were made carefully so the logical interpretation of results was not affected. Thus, the range of the second and fifth groups is larger compared to the rest of the groups. It can be observed that shares in the lower (14,32%) and upper (9,17%) age groups are the smallest, whilst participants between 26 and 45 years represent the largest share in the sample (27.52%). In terms of education, roughly 60% of the sample is formed by highly educated individuals holding a university degree or postdoctoral studies, while approximately 40% is formed by citizens holding secondary education and a professional degree. In terms of income levels, the sample presents an unbalanced composition due to the overrepresentation of individuals (more than 50%) reporting a gross monthly household income smaller than 2000€, while 14.54% earn more than 8000€. Around 28% of respondents reported a household income level ranging from 2000€ to 8000€. Albeit, this socio-demographic variable is difficult to interpret because the household composition is unknown, which hinders the capacity to generate more precise econometric estimates of income among participants. Perhaps, it is more effective to request for participants to state their individual net/gross earnings per month based on income categories.

Table 2*: Sample Socio-Demographics

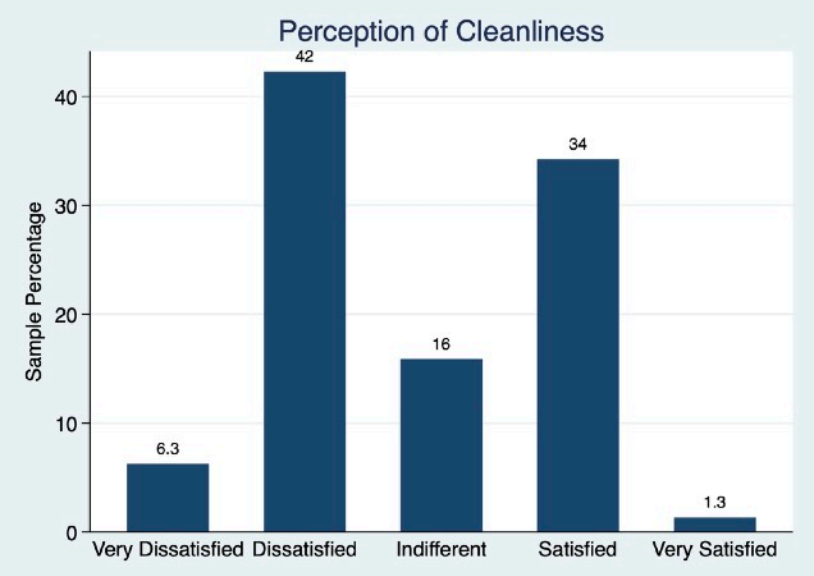
		Frequency	Percent	Cumulative
Gender	Female	292	65,32 %	65,32 %
	Male	152	34,00 %	99,33 %
	Other	3	0,67 %	100,00 %
	Total	447	100,00 %	
Age groups	18 - 25 years	64	14,32 %	14,32 %
	26 - 45 years	123	27,52 %	41,83 %
	46 - 55 years	110	24,61 %	66,44 %
	56 - 65 years	109	24,38 %	90,83 %
	66 - 85 years	41	9,17 %	100,00 %
	Total	447	100,00 %	
Education level	Secondary Education	60	13,42 %	13,42 %
	Professional Degree	117	26,17 %	39,60 %
	University	178	39,82 %	79,42 %
	Postgraduate	92	20,58 %	100,00 %
	Total	447	100,00 %	
Gross monthly income income per household	I prefer not to disclose	25	5,59 %	5,59 %
	Less than 1000€	118	26,40 %	31,99 %
	1001€ - 2000€	114	25,50 %	57,49 %
	2001€ - 3500€	67	14,99 %	72,48 %
	3501€ - 5000€	26	5,82 %	78,30 %
	5001€ - 6500€	12	2,68 %	80,98 %
	6501€ - 8000€	20	4,47 %	85,46 %
	More than 8000€	65	14,54 %	100,00 %
	Total	447	100,00 %	

4.4.2 Perceptions, Attitudes, and Preferences

Graph 1 presents the results of the initial inquiry related to satisfaction levels concerning cleanliness in the city of Zaragoza. As can be observed, a majority of the sampled citizens

(48.3%) are dissatisfied with the overall state of cleanliness in Zaragoza. On the contrary, 35.3% of the sample perceive the city as clean and state to be satisfied with status quo conditions. 16% of the sample believes the city is not clean or dirty. Observed discordance in opinions signals how differently citizens evaluate existing and perceptible urban litter pollution in Zaragoza. For one dissatisfied person, actual litter levels in the city are perceived as detestable, whilst a satisfied individual perceives that same amount of waste as tolerable. Notwithstanding the clear perception divergence in the sample, results manifest a larger share of the sample being disenchanted with the baseline litter conditions in the city, which strengthens the necessity of further research on the topic.

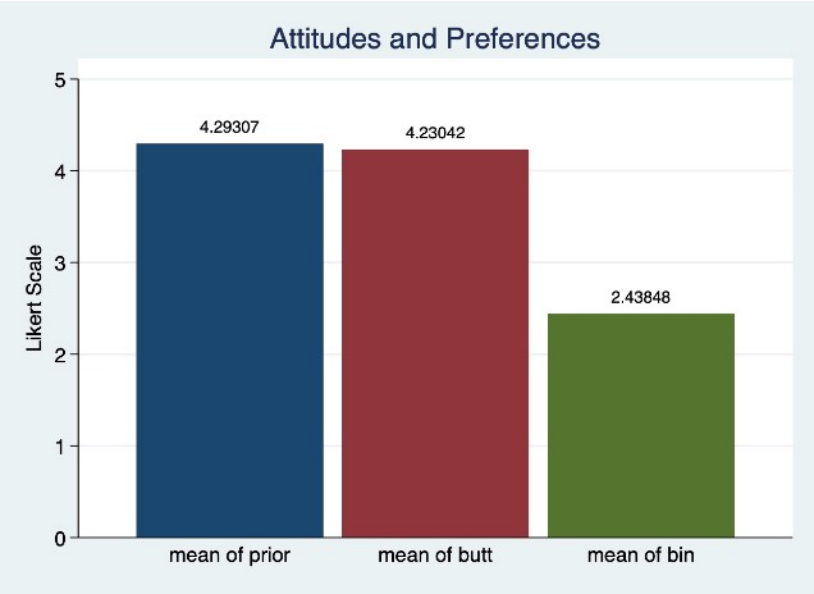
Graph 1: Perceptions



Graph 2 shows the mean results of three inquiries related to attitudes and preferences towards the environment and littering. The first variable "prior" represents the question: "I believe that minimizing environmental degradation has to be a priority". The mean response was 4.3 out of 5, meaning that the average respondent demands the protection of the environment to be a prime political concern. A more comprehensive cognition of attitudinal viewpoints would require extra questions concerning other priorities such as economic conditions to obtain a wider angle of the topic, but no further questions were included to avoid respondent fatigue. Nonetheless, in perspective, according to the Eurobarometer's last report on public opinion in the European Union, "environment and climate change" stands as the fourth main priority for Europeans (Eurobarometer, 2022). Combatting sharply rising living costs, securing the energy supply, and resolving international tensions comprise the three top concerns for European citizens (Eurobarometer, 2022). Secondly, the variable "butt" stands for "I am bothered when somebody litters a cigarette butt" and shows a very similar mean relative to the first inquiry. This convergence indicates that the average participant appears to associate the protection of the environment with litter behavior condemnation, thus establishing a robust link between environmental degradation and cigarette butts. Thirdly, "bin" represents the statement "If I happen to see litter, I always collect it for proper disposal". Results show that, to some extent, participants take individual action against urban litter, but not consistently. Interestingly, there is a noteworthy contrast with the mean of "butt" and "prior". This existing gap symbolizes the gulf between attitudes and deeds. A person may mentally reject litter behavior, and consider the environment a political priority, but does not take frequent personal action in tackling the

problem at the individual level. Despite that, overall descriptive results gravitate towards a general pro-environmental stance in the sample.

Graph 2: Attitudes and Preferences



4.4.3 Citizens' General Disposition to Support Environmental Change

According to sample results presented in table 3, close to 67% of participants decided to participate in the contingent market proposed and consequently, would financially support the creation of a "green tax" dedicated to maintaining Zaragoza constantly free of litter. However, to fully comprehend this result, the remaining 33% of the sample has to be further analyzed according to Table 4. This table contains the follow-up question offered to those participants that stated a zero bid. As has been mentioned previously, state-of-the-art CV methods should make a clear distinction between protest and true zero bids, to not jeopardize the aggregation of economic benefits. The task of identification and differentiation of protestors and true zero bidders was conducted in accordance with existing literature (see Brouwer and Martín-Ortega, 2012; Meyerhoff and Liebe, 2006, Lo and Jim, 2015; Strazzera et al., 2003; Chen and Qi, 2018; Jorgensen and Syme, 2000).

In the case of money contributions, protest bids predominate over true zero bids. Starting with true zero bids, only 14 respondents affirm not being able to pay the green tax. Although some authors argue that reasons related to the inability to pay should be considered as protests (Milon, 1989), it is generally conceived that insufficient income should be expected to be a true zero bid (Brouwer and Martín-Ortega, 2012). Nonetheless, this matter is still subject to debate because respondents have intricate and unobserved belief systems and situations that influence their decision to not pay, which can not be fully captured in conventional CV surveys (Brouwer and Martín-Ortega, 2012). Anyhow, in this study, these participants are regarded as true zero bidders. Concerning protest responses, nearly 57% state that local authorities already collect enough money and that if utilized efficiently, litter accumulation could be easily eradicated in Zaragoza. 24.32% directly reject the imposition of another tax on top of existing ones. Among the 13% that selected "other reasons", a majority states that education programs are more effective in tackling litter, or further developing the previous option of "I already pay enough taxes". Therefore, responses in the open-ended section of the study are also taxonomized as

protests. As stated by Halstead et al. (1992, p. 161), participants that "unconditionally reject some contextual component of the contingent market, should also be unconditionally treated as protestors". Kriström (1997) argues that an alternative payment vehicle should be proposed to induce "protesters" to disclose their true WTP. In the present research no further schemes were included to not introduce response fatigue bias in the survey. Concluding, the monetary dimension of the study faces two observed opposing forces: positive bidders and protestors. True zero bidders in the CV study solely represent approximately 3% of the total sample.

Concerning time, a very similar positive result is obtained. Almost 65% of the sample participated in the proposed time contingent market and stated a positive frequency of participation in urban public cleanups. Once again, to comprehend this result, the stated reasoning behind zero bids must be analyzed with the help of follow-ups. As can be seen in Table 4, 22.29% of those participants were not willing to participate, stating that they do not want to clean others' litter and nearly 25% declares not having enough time to participate in clean-ups. 12.10% allege having physical problems that would impede him/her to participate. These three statements can be directly regarded as true zero bids. Lastly, the lion's share of this sub-sample is found in the open-ended alternative "other reasons". Overall, open-ended statements revolve again around education being the only solution to the problem and the inefficient and ineffective use of public funds by the city council. Thereby, one could argue, based on the CV literature, that this share of respondents are protestors. Consequently, only those alleging 'physical problems', 'lack of time', and 'not willing to clean others' litter' could be considered true zero bidders. On the contrary, one could also argue that respondents in the open-ended questions are in reality, true zero bidders. No declaration was reported stating that if, for example, the cleanup organization was not managed by the city council but self-governed by citizens, she would participate. This would entail an alternative vehicle and a protest bid rooted in mistrust towards the public authorities. Seems obvious that a protestor in the CV dimension would state a positive bid if she believed that the city council manages public funds efficiently. But that is not the case in the WTC scenario. Participants that would not participate in cleanups appear to simply reject the idea of contributing with their time to remove existing urban litter and would not do it under any circumstance. In whichever way, no further information can be extracted from the sample data utilized in this study, and the optimal treatment of zero frequency bids in CB litter methods is empirically unknown so intuitive reasoning must be utilized. As a result, in this work, all zero frequency bids in this study are regarded as true zero bids for the reasons explained above.

Table 3*: Sample's willingness to pay and cleanup

		Frequency	Percent	Cumulative
Are you WTP?	No	148	33,11 %	33,11 %
	Yes	299	66,89 %	100,00 %
	Total	447	100,00 %	
Are you WTC?	No	157	35,12 %	35,12 %
	Yes	290	64,88 %	100,00 %
	Total	447	100,00 %	

Table 4*: Identification of zero bids

	Identification		Frequency	Percent	Cumulative
Not WTP: Reasons	Protest bid	"I pay enough taxes already"	36	24,32 %	24,32 %
	True zero bid	"I can not afford to pay the green tax"	14	9,46 %	33,78 %
	Protest bid	"I believe public money that is already collected should be better utilized"	85	57,43 %	91,22 %
	Protest bids	"Other reasons"	13	8,78 %	100,00 %
		Total	148	100,00 %	
Not WTC: Reasons	True zero bid	"I am not willing to collect litter that I do not generate"	35	22,29 %	22,29 %
	True zero bid	"I do not have time"	39	24,84 %	47,13 %
	True zero bid	"I have mobility problems"	19	12,10 %	59,24 %
	True zero bids	"Other reasons"	64	40,76 %	100,00 %
		Total	157	100,00 %	

Lastly, table 5 displays a matrix of conditional frequencies of WTP and/or WTC in the sample. According to the results, an individual that refuses to cleanup would state a zero bid as WTP with a probability of 59.24%. In contrast, given that she rejects participating in clean-ups, she will state a positive bid in the green tax scenario with a 40.76% probability. Conversely, an individual who is willing to contribute will also be willing to pay with a probability of 81.03%. Conversely, given that a participant is willing to contribute with time, she will not be willing to pay with only a chance of 18.97%. These conditional probabilities superficially suggest that time and money contributions are strong complementaries. In other words, there exists a positive correlation between WTP and WTC. Hence, participants are willing to contribute with both time and money, and would not substitute one for the other. However, as stated by Cappellari et al. (2011), sample probabilities can paint a misleading picture and should not be interpreted as correlations due to existing "compositional effects". Additionally, protest bids completely disrupt the use of conditional probabilities to establish statistical correlations. Therefore, an accurate understanding of the true underlying relationships requires multivariate econometrical analysis (Cappellari et al., 2011).

Table 5: Conditional probabilities in the sample

	Are you WTC?		
Are you WTP?	No = 0	Yes = 1	Total
No = 0	93	55	148
	59,24 %	18,97 %	33,11 %
Yes = 1	64	235	299

	Are you WTC?		
	40,76 %	81,03 %	66,89 %
Total	157	290	447
	100,00 %	100,00 %	100,00 %

4.5 Discriminant Descriptive Analysis of Monetary Contributions

Once the three different types of respondents in the dimension of money contributions have been accounted for, a more in-depth analysis is required to assess whether there exist systematic differences between protestors, true zero bidders, and positive bidders. Almost identical characteristics across subsamples would indicate that the decision to protest is randomly distributed in the sample divide and, eliminating them, would not prompt sample selection bias (Brouwer and Martín-Ortega, 2012). Albeit, this is unlikely to hold. Following the footsteps of the 'CV protest literature' (see Strazzera et al., 2003; Brouwer and Martín-Ortega, 2012; Collins and Rosenberger, 2007; and Chen and Qi, 2018), variable means of characteristics of interest are compared between subsample populations by conducting a 'discriminant analysis'.

Overall, significant differences are found across subsamples and variables. Concerning perceptions, attitudes, and behavior, true zero bidders are notably less satisfied than protestors and positive bidders, while the latter, are the most satisfied. Consequently, dissatisfaction and disenchantment with the actual state of the city can be regarded as a driving factor for protesting in the survey against public authorities, polluting companies, etc. Moreover, the fact that protestors tend to be more dissatisfied with status quo conditions means that they hold a positive WTP for an environmental improvement but choose not to state it. Secondly, positive bidders are found to have a stronger pro-environmental attitude ("priority"), closely followed by positive bidders. In turn, protestors present on average a weaker stance in this regard. It appears that protestors in the sample, are somewhat dissatisfied with the overall management of the city in many different dimensions (not only cleanliness) and believe that other problems should be tackled before the environmental and aesthetic degradation of urban spaces. Nonetheless, the mean response in "priority" of protestors is high (above 4) which indicates that they have a positive WTP for environmental public goods such as urban cleanliness but decide not to reveal the amount because of rejection of some contextual element of the contingent market proposed. Thirdly, true zero and positive bidders are notably more committed to litter removal practices than protestors, which present the lowest score. Seems well-founded that protestors would also reject personal citizen action against litter because, according to the results, in their eyes, it is the fault of uncivilized citizens and it should be the city council's task to deal with the problem via existing taxes and educational campaigns.

Table 6**: Discriminant analysis of perceptions, attitudes, and behavior

Variables	Type of bidder	N	Mean	SD
Satisfied (yes = 1)	True zero bidder	14	.143	.363
	Positive bid	299	.378	.486
	Protestor	134	.328	.471
	Total	447		

Priority (0 - 5)	True zero bidder	14	4.429	.852
	Positive bid	299	4.398	.908
	Protestor	134	4.045	1.268
	Total	447		
Cleanup (0 - 5)	True zero bidder	14	2.5	1.225
	Positive bid	299	2.562	1.39
	Protestor	134	2.157	1.486
	Total	447		

The discriminant analysis results of socio-demographic variables of interest are presented in table 7. According to descriptive results, positive bidders are considerably more academically educated than true zero bidders. This result could emerge from the fact that education is positively correlated with income and positive bidders are in a better financial position to state a positive contribution and are less likely to protest against the payment vehicle. In contrast, protestors have a positive WTP but do not want to express it in terms of money collected through a tax. It should be noted that, up until this point, all descriptive results from the discriminant analysis are in line with those suggested by Meyerhoff and Liebe (2006), which conduct a multivariate regression analysis on the determinants of protest behavior in CV surveys. Albeit, with respect to age, older respondents in the sample are more prone to protest than younger participants, something that is contrary to the study by Meyerhoff and Liebe (2006). The authors also find that income presents a positive correlation to non-protest behavior, which is not the case in the sample of the present study. Anyhow, as stated before, income is a problematic variable in this study due to its unbalanced composition in terms of observations. Concluding, there appears to be ample evidence of considerable differences between subsamples, which suggests that the decision to protest is not random but subject to observed characteristics. Thus, a specific econometric approach is required to not bias subsequent results.

Table 7**: Discriminant analysis of socio-demographics

Variables	**Type of bidder**	**N**	**Mean**	**SD**
Highly educated (yes = 1)	True zero bidder	14	.357	.497
	Positive bid	299	.666	.473
	Protestor	134	.493	.502
	Total	447		
Age group (1 - 5)	True zero bidder	14	2.5	.855
	Positive bid	299	2.779	1.274

Variables	Type of bidder	N	Mean	SD
	Protestor	134	3.097	1.025
	Total	447		
Middle-high income (1 = yes)	True zero bidder	14	.357	.497
	Positive bid	299	.204	.404
	Protestor	134	.231	.423
	Total	447		

5 Discussion of Results

In this chapter, the results obtained from the study are analyzed. The computation of results was conducted with the use of the maximum likelihood method and the models explained in Chapter 3. The results are presented and discussed, presenting factual data and possible explanations of the underlying dynamics of results. In order to test the construct validity of the findings, similar studies are utilized to compare the results. It should be noted that only statistically or economically relevant results will be analyzed.

5.1 Willingness To Cleanup

As discussed previously, two regression models are computed in the time dimension of the study. In the survey, respondents were asked if they were willing to participate in public cleanups in the city of Zaragoza via a payment card ranging from 0 hours to 36 hours per year. Table 8 presents the estimated results which essentially explore the driving factors of a positive WTC. In the first column of the table, the results of the binary Probit model are displayed. In this specification, the dependent variable consists of a two-sided variable based on whether a respondent is willing to cleanup or not. As a result, the sample is divided into two groups: positive frequency bidders and zero frequency bidders. As discussed previously, all zero bids are considered as genuine and legitimate true zero responses. In the second column, the results of the Tobit model can be observed. This regression was conducted based on the actual number of hours respondents would be willing to participate in cleanups per year. Note that, all observations were utilized in both models, avoiding sample selection bias, which has been a troublesome task. All coefficients represent marginal effects on the dependent variables and are studied in isolation, holding all other variables constant. In the case of the Probit model, marginal effects had to be calculated after computing the model, but in the case of the Tobit model, since the dependent variable is expressed in logs, marginal effects can be readily interpreted. In terms of significance, several consistent and promising results are found across the two models and those will be compared to statistically relevant results in the literature. Therefore, an in-depth analysis of the determinants of WTC is conducted, firstly, studying the results of the binary Probit model, and secondly, reviewing estimations in the Tobit regression.

Starting with the binary Probit regression, it can be observed that a 1 unit increase in the variable 'priority' (stronger pro-environment attitude), increases the probability of positive WTC by 4.8%, at the 5% significance level. As shown by relevant literature (Bouma and Koetse, 2019; Spash, 2006), attitudinal beliefs play an important role in participation in contingent environmental markets. Thereby, a positive correlation appears to be correctly established. Secondly, a 1 unit increase in 'cleanup' (pro-environmental behavior) will be translated into an increase of 5.8% in the probability of participating in cleanups, at the 1% confidence level. Intuitively, a positive correlation could be expected. A person that is already collecting urban litter daily, will on average be more inclined to collaborate in a collective cleanup. With respect to the literature, Brouwer et al. (2017) controlled for pro-environmental deeds by asking whether respondents were members of any environmental organization and find a positive correlation with willingness to participate in beach cleanups. Albeit, as discovered in the "time vs money" literature (see for instance Cappellari et al., 2011), individuals tend to contribute to charity based on impure altruistic purposes, such as warm glow. These contribution motives can be regarded as weaker and may not be so connected to participation in cleanups, which entails contributing with time, instead of money (financially contributing to an environmental organization). In contrast, collecting litter could be considered to be a pure altruistic deed and a contribution of time, that may establish a more

muscular link to WTC. Nevertheless, both variables, 'cleanup' and 'membership' present a positive correlation which proves that cleanup behavior is tethered to pro-environmental attitudes and actions. Concerning socio-demographic variables, 'highly educated' reports the largest effect on the probability of joining a cleanup. For those respondents that have at least a university degree, the probability of participating in a public cleanup is 8.7% higher than those respondents that have a lower academic qualification. This coefficient is statistically significant at the 5% level. In Brouwer et al. (2017) and Shen et al, (2019) identical results are found, which indicate that, on average, citizens that are more academically educated will report higher levels of ecological values and awareness. As a result, they are more likely to join an urban cleanup. Secondly, estimated results show a negative correlation between age and WTC, meaning that the older a respondent is, the less likely it is that she will accept joining a public cleanup. Specifically, if the age group increases by 1 category, the probability of joining a cleanup decreases by 7.3%, at the 1% significance level. Generally speaking, older citizens are often richer and less physically apt than younger citizens and therefore, should be more likely to report a lower WTC, but a higher WTP. This result is aligned with results estimated by Brouwer et al. (2017) which also find a negative relationship.

Finally, the coefficient analysis of the dichotomous variable 'middle-high income' is ruled out because it is not found to be statistically relevant. Even so, it is feasible to study the sign of the coefficient as a statistical association with WTC. The income variable then, indicates that individuals with gross income levels per household higher than 5001 euros per month will be less inclined to participate in cleanups, compared to those respondents with less income. Generally speaking, a negative sign seems to be in accordance with reality since richer individuals have a higher opportunity cost of time in terms of wage but also in terms of alternative leisure activities that would substitute cleanups. Richer individuals interested in tackling litter pollution should on average prefer to contribute with money. Furthermore, income is, in theory, tightly tethered to age and education and as a consequence, consistent results should be found across variables. However, that is not the case. On the one hand, the negative sign of the coefficient of income appears to be aligned with age. Older individuals will report higher income levels. If, as stated before, older individuals are less likely to participate, then a negative correlation between WTC and income seems appropriate. Conversely, the negative sign of income is not consistent with education. Richer individuals tend to be more academically educated and more informed about waste pollution. Richer and highly educated respondents should be more willing to participate in cleanups than less rich/educated citizens. Thus, a positive correlation should be revealed. Interestingly, results denote the existence of these two opposing effects, which result in age having a greater impact on income, due to the negative sign. As a result, the probability of a positive WTC should be expected to be decreasing as a function of income. Strikingly, existing literature reveals contradictory outcomes concerning income and WTC. On the one hand, Brouwer et al. (2017) report a positive correlation between income and WTC, whilst Shen et al. (2019) report a negative sign on the income coefficient. Anyhow, it should be remarked that the income variable in this study is highly problematic, which hinders the capacity to confidently extract empirical conclusions.

In the second column, estimations from the Tobit regression are showcased. A superficial analysis rapidly reveals an expected doubling (relative to the binary Probit model) in the size of the coefficients for every variable, except for education, which decreases in size and also loses statistical significance. For instance, now a unit increase in 'priority' translates into an increase of 10.9% in the number of hours that an individual is willing to cleanup in one year, with a 10% level of significance. In terms of pro-environmental behavior, a unit increase in 'cleanup' correlates with an 18.5% increase in the number of participation hours in a year with a significance level of 1%. In comparison with Brouwer et al. (2017), the effect of 'member environmental organization' loses all statistical significance when modeling the maximum

number of cleanup hours with the Tobit regression model. Therefore, 'cleanup' much better explains variation in the number of hours than being a member of an environmental organization. As discussed, individuals that collect litter regularly signal a higher level of commitment to environmental issues than individuals that financially contribute to an environmental organization. In terms of socio-demographics, 'highly educated' is not significant anymore in the Tobit model, but still has a positive sign. Perhaps, the fact that education loses significance when modeling the frequency of participation, indicates why the effect of age on income is stronger when studying the probability of participation. As can be observed, age is still highly significant and increases in the magnitude of impact. A unit jump in the age category decreases the number of hours by 16.1%, at the 1% significance level. Additionally, it is not surprising that age has a forceful effect on the frequency of participation. Participants aged between 18 - 26 years old will be able to participate more hours compared to the age group ranging from 66 to 85 mainly due to physical aptitude. Moreover, age is not so tightly related to education, as a 25 year may have the same education level as a 78-year-old, and therefore dissimilar results are not a worrisome indicator. Likewise, Brouwer et al. (2017) also find a greater negative impact of age on participation frequency. With respect to income, the negative relationship consistently arises again. This negative association is significant in Shen et al. (2019), which stands in contradiction to Brouwer et al. (2017). Lastly, it is important to highlight the fact that, the constant of the model shows a significant positive magnitude which manifestly indicates that respondents in the sample present an overall positive attitude towards participation in urban cleanups.

Concluding, it is worth mentioning that variables such as 'gender', 'beach litter annoyance', and 'beach cleanliness perception' are found to be very statistically significant in Brouwer et al. (2017) and Shen et al. (2019). On the contrary, 'gender', 'satisfaction with city cleanliness', and 'littering behavior annoyance' are not statistically significant in the present study. The lack of significance in the latter two could emerge from the fact that Brouwer et al. (2017) and Shen et al. (2019) surveyed respondents in person at different beach locations, and they report that litter was present at the beaches when interviews were conducted. It seems reasonable that those variables present a stronger correlation to WTC if litter is in sight. On the contrary, in the present study, participants completed the questionnaire online and had to mentally visualize existing litter in Zaragoza. For example, counterintuitively, approximately 67% of citizens that affirmed being satisfied with overall cleanliness in Zaragoza, stated afterwards a positive WTC. In terms of gender, as has been stated before, the fact that females crowd the sample (approximately 66% of respondents) surely stands in the way of statistical significance.

Table 8**: Estimated Probit and Tobit models exploring willingness to cleanup

Variables	**Probit model** (willingness to cleanup: 1 = yes)	**Tobit model** (max n. of hours willing to cleanup)
Satisfaction (0 -5)	.003	-.009
	(.016)	(.043)
Priority (0 - 5)	.048**	.109*
	(.021)	(.058)
Cleanup (0 - 5)	.058***	.185***
	(.015)	(.042)
Highly educated (1 = yes)	.087**	.048
	(.044)	(.121)

Age group (1 - 5)	-.073***	-.161***
	(.017)	(.046)
Middle-high income (1 = yes)	-.036	-.102
	(.052)	(.136)
Constant		1.073***
		(.336)
Observations	447	447

Standard errors in parenthesis.
****p<0.01; **p<0.05; *p<0.10.*

5.2 Willingness To Pay

The results of the CV market were computed using two different regression models. Respondents were asked if they were willing to contribute to the payment of a green tax destined to create a cleaning brigade of professionals that would maintain the city free of litter all year round. The elicitation of bids was conducted via a payment card ranging from 0 euros to 90 euros per year. The statistical procedure differs from the one employed in the time dimension because of two main reasons: first, estimating a binary Probit model would have resulted in a measurement of the determinants of protesting or not, which is not the objective of the study. Secondly, applying a Tobit model to the entire sample would bias the results because protest bids would be incorrectly treated as true zero bids. Consequently, an alternative strategy is exploited. In the first place, a Tobit model excluding zero bids is estimated, to intentionally incur in sample selection bias. Censoring protest bids is a common practice in CV studies because of its convenience, but it results in hugely biased estimations. As empirical proof, the model is computed. Second of all, the FIML sample selection model is employed. The use of FIML regression entails state-of-the-art statistical analysis of CV surveys and has been advocated in many influential empirical studies (see Strazzera et al., 2001; Strazzera et al., 2003; Brouwer and Martín-Ortega, 2012). To estimate it, all protest bids are assumed to be in reality, positive bids. Then the model infers a positive value for protest responses based on sample characteristics. By utilizing the FIML regression model, not only you are eradicating the need for censoring, but also you are more accurately measuring the aggregation of non-market benefits. Even though the latter feature was not exploited since the number of statistically significant variables is too scarce to appropriately estimate the mean WTP of the sample. It should be noted, that comparison of results with the literature will not be conducted in this section because previous CV studies on litter do not generally account for the existence of protesters. Only Abate et al. (2020) recognize their existence, but the authors decide to exclude them from the analysis. For this reason, coefficients may be biased in many of the studies. Signs of coefficients will be compared, as the likelihood of biasedness is lower.

In the first column of table 9, the results of the Tobit regression model are displayed, whereas, in the second, the FIML model results are shown. All results can be interpreted without further calculations because they are expressed as marginal effects and the dependent variable is in logarithmic form. As expected, coefficients in the Tobit model, are downwardly biased relative to the FIML's estimations: nearly 1% positive difference in the variable 'cleanup', 20% positive difference in the dichotomous variable of education, and 28% positive difference in terms of age. Despite this sizable disparity, the signs of the coefficients were correctly estimated in both models and are in accordance with studies reviewed in the CV literature on marine litter (see for

example Brouwer et al., 2017; Shen et al., 2019; Abate et al., 2020, and Smith et al., 1997). All studies reviewed, conclude that age, education, and membership in an environmental organization have a positive effect on WTP for the removal of litter. In the present study, all these variables present a positive correlation with WTP for improving baseline conditions in the city in terms of waste. Albeit, in this study, income, yet again presents a negative sign, which in this case, is even more shocking than in the contingent behavioral market. An explanation could be found again in the fact that middle-high income citizens are underrepresented in the sample (approximately 71% of the sample stated middle-low or low-income levels). Moreover, in the sample, respondents with higher income levels, have a larger tendency to protest or to state lower bids, which in combination with a low number of individuals in that category, seemingly disrupt the logical result of WTP increasing as a function of income.

In terms of the coefficients' sign and magnitude in the FIML model, we find a handful of reliable effects. Firstly, an additional point in 'cleanup' triggers a 7.3% increase in the stated WTP at the 5% level of statistical significance. Not surprisingly, the effect of 'cleanup' is lower than in time contributions (11%). Seems logical that cleanup behavior is more strongly associated with labor supply than with WTP. Notwithstanding this, the variable 'cleanup' is intrinsically linked to a pro-environmental attitude which outcomes a positive correlation to WTP for improving the cleanliness state of the city. Perhaps, controlling for whether participants contribute financially to an environmental organization would have presented a stronger correlation to WTP. For instance, Brouwer et al. (2017), Zambrano-Monserrate and Ruano (2020), and Abate et al. (2020) do measure the effect of 'membership in an environmental organization' on WTP for beach litter removal and find a very strong positive correlation. Regarding education, the effect is large. Being highly educated correlates with an increase of 48% in the stated monetary bid, at the 1% statistical significance level. As discussed before, academic education is, on a general basis, a precursor of high earnings and environmental consciousness. Accordingly, Abate et al. (2020), Tyllianakis and Ferrini (2021), and Zambrano-Monserrate and Ruano (2020) also find a sturdy positive correlation between education and WTP for the removal of marine litter. Moreover, an increase in the age group is translated on average, with an 8.8% higher monetary bid. This result is significant at the 10% level. The size and sign of the coefficient, validate the argumentations proposed in the time contributions analysis of results because, in terms of money, age proves to be a positive and increasing force. Older respondents are more willing to contribute with money than with time. Furthermore, age is closely associated with income, so it is a solid statement that older participants have lesser budget constraints and more spare income for extra monetary expenditures. Lastly, the constant is positive and significant across the models (both for time and money contributions), which unquestionably indicates that participants are willing to move away from the status quo and secure more clean environment, regardless of the measure proposed., even though the size of the constant is larger for money contributions (2.13 relative to 1.073 in the WTC scenario).

Once again, as in the dimension of time contributions, attitudes, perceptions, gender, and income are not statistically relevant and have not been included in the regression analysis of data. In the case of gender and income, statistical significance is probably neglected because of the unbalanced number of observations inside the different categories of those variables. As mentioned before, females and low-income households are excessively represented in the sample. Concerning attitudes and perceptions, only "cleanup" is found to be significant in the money dimension of the study, but not "satisfaction", "priority", and "butt". In the case of "satisfaction", a substantial share of "satisfied" individuals would contribute financially to the green tax, which is counterintuitive to logical reasoning. An individual that is satisfied with the level of urban cleanliness in Zaragoza should be expected to state a zero-bid response. It seems that there are unobserved factors that are affecting this relationship. "Priority" and "butt", as

discussed before, presented very high sample means (around 4.3 out of 5) which means that a majority of the sample stated having a pro-environmental attitude. As a result not enough variance was achieved, and seems reasonable that statistical significance is not found.

Table 9**: Estimated Tobit and FIML models exploring willingness to pay

Variables	**Tobit model** (WTP>0)	**FIML model** (including zero bids)
Satisfaction (0 - 5)	-.034	-.0029
	(.039)	(.048)
Priority (0 - 5)	.059	.030
	(.063)	(.067)
Cleanup (0 - 5)	.062*	.073*
	(.036)	(.042)
Highly educated (1 = yes)	.302***	.49***
	(.103)	(.12)
Age group (1 - 5)	.06	.088*
	(.037)	(.046)
Middle-high income (1 = yes)	-.04	-.19
	(.139)	(.14)
Constant	2.48***	2.13***
	(.158)	(.20)
Log likelihood	-373.217	-3890.896
Observations	299	447

Standard errors in parenthesis.
****p<0.01; **p<0.05; *p<0.10.*

5.3 Sample Mean Model

In the sample mean model, raw calculations are performed to investigate the proposed measure that outcomes the highest economic non-market benefits (more valuable in monetary terms). The sample means are calculated as:

$$\overline{WTP} = \frac{8{,}709}{447} = 19.48 \; ; \; \overline{WTC} = \frac{3{,}422}{447} = 7.65$$

Concerning directly elicited money contributions, a total amount of 8,709 euros was hypothetically retrieved, which divided by the total size of the sample, gives us the sample mean

of 19.48 euros per year, meaning that the average respondent is willing to pay that amount every year. Concerning time contributions, participants are willing to cleanup an aggregate of 3,422 hours, which divided by 447 participants, yields an average of 7.65 hours per year. To transform stated WTC into WTP, back-of-the-envelope calculations have to be performed. The procedure is the following: compute the mean income group based on participant responses. The mean income group was the third, which comprises individuals that belong to a household that generates between 2,001 euros and 3,500 euros of monthly gross income. This age group provides a mid-point in the interval of 2,750 euros per month, which gives you an estimation of the average income inside the income category. Assuming that, generally speaking, households are formed by two working individuals, the average monthly gross wage per person in the sample is 1,375 euros. The calculation of monthly working hours is:

$$WH_{month} = (40 * 4) + (8 * 3) = 184$$

where there are 40 working hours per week, and a total of 4 weeks and 3 days in a month, which results in 184 working hours per month. To get the hourly gross wage then, divide 1,375 euros by 184 hours which equals 7.47 euros. That is the average gross hourly wage in the sample. Now, by multiplying 7.47 euros by the sample mean of 7.65 hours that participants are willing to cleanup per year, we get an individual contribution of approximately 57.20 euros per year. Then, if we aggregate welfare estimations:

$$WTC = 57.20 * 447 = 25{,}568.40€$$

$$WTP = 19.48 * 447 = 8{,}707.56€$$

we get a total of 25,568.4€ of non-market benefits measured via time contributions. In contrast with money contributions, the figure is substantially higher. A total of 8,709€ were hypothetically collected via the green tax. Overall and even though these are conservative and raw calculations, it is plain to see that time gifts outcomes the highest economic benefits based on citizens' willingness to contribute, compared to the green tax. It appears that individuals are much more willing to contribute with their time than with their money. This large difference in the worth of aggregated contributions could stem from four main reasons: firstly, several respondents may be subject to some form of mental accounting bias, in which respondents do not make the cognitive mathematical operations implicitly required in the proposed contingent market. As a result, they do not realize that they are contributing much more than they probably wanted. Secondly, participants in the sample highly value social, community-based, altruistic, and environmental activities and perceive cleanups as leisure activities. It appears that opportunity costs of, for instance, watching Netflix, are low if an initiative like this is in place. Thirdly, a portion of respondents may compute a quick cost-benefit analysis and determine that it is much more efficient and effective to directly employ citizens than collect public money and create supplementary cleaning brigades., and would contribute more time than money. To some extent, this could be considered a protest against current governmental practices. Lastly, inflation and energy prices are very high in Europe and Spain, which is having a profound effect on low-income citizens. As in the sample, the majority of respondents present low-income levels, it is reasonable to find that pro-environmental participants are much more willing to contribute time than money. Time is a powerful tool when tumultuous economic cycles strike and this fact should be harnessed by public authorities as a way to solve urban problems and reduce tax burdens.

6 Conclusions

Even though municipal waste only represents a tiny fraction of the total bulk of waste produced in the EU, its management constitutes one of the most complex issues in European cities. Vast amounts of hard-to-recycle residues are generated daily and the systems in place for appropriate treatment are not effective enough. One clear example is the accumulation of litter in urban environments that is caused by the improper disposal of residues. In this age, waste has become an inherent component of our socio-economic system, and litter is nowadays a ubiquitous feature of urban and natural landscapes, preposterously normalized and trivialized by citizens and public authorities. In this sense, it can be argued that urban cleanliness is a public amenity, subject to non-rival and non-excludable conditions, and threatened by the common's tragedy trap, which essentially means that, any good that is in the common ownership will inevitably be overexploited and massively degraded. This is what constitutes the litter predicament that threatens urban spaces in cities. However, while transforming entire waste management systems is a capital-intensive and costly task, litter removal practices are effective and based upon human determination and labor. As a first step towards paradigmatic change, the eradication of existing and future litter should be a priority if the intent is to break the vicious cycle of litter. With this purpose, stated preference techniques provide the optimal toolkit for environmental economists to calculate non-commercial benefits by utilizing consequential and incentive-compatible hypothetical policies that induce individuals to reveal their preferences and attitudes toward the improvement of a public and environmental good. If meticulously gauged and designed, CV and CB survey schemes may outline the tremendous benefits to be attained by eliminating litter from urban spaces and help policymakers to properly assess and implement measures that pass the test of a rigorous CBA aimed to tackle the litter predicament in cities. SP methods offer an ingenious instrument that exploits its double-edged hypothetical nature to envision desirable futures and measure the welfare changes that they would entail.

In the present work, two main SP methodologies have been employed. The CV method explores the non-market benefits of litter removal derived by citizens and the driving factors behind a zero or positive WTP for urban cleanliness. In this dimension, monetary contributions are elicited and utilized to measure economic benefits. Secondly, the CB scheme is applied to measure the disposition of citizens to contribute with their time to the removal of litter and to study the determinants of a zero or positive WTC. To maximize the validity of results, bias should be minimized in the survey scheme via incentive compatibility and consequentialism. In this regard, state-of-the-art procedures are applied to the creation and presentation of the contingent market using a parallel design strategy for both time and monetary contributions. CB methods are extensively unexplored and the sources and magnitude of biases remain unknown. Thus, the extent to which stated bids trigger similar biases compared to stated frequencies is still an empirical question. For the present research, an analogous strategy is utilized in both methodologies. In this regard, follow-up studies can and should be conducted to explore the causes and consequences of hypothetical biases in CB methods to grasp the gap between stated and actual behavior. Moreover, the econometric treatment of responses is still subject to debate and protestors comprise an almost physiological component of contingent surveys. Advanced techniques have been proposed in the literature, and have been successfully utilized in the present work by first, identifying protest responses, and then employing an adequate econometric treatment to the analysis of results. Once again, protest bids and selectivity bias in CB surveys have not been addressed by the empirical literature, and their identification and treatment are unknown. As opposed to bias minimization, a differing strategy is adopted for the treatment of monetary and time results. Concerning results, once corrected

for the implicit selectivity bias that would arise from excluding protestors, monetary results suggest that a pro-environmental attitude and behavior, education, and age positively correlate with an increasing WTP for urban cleanliness. The sample mean WTP is found to be 19.48€ per year, which gives a total amount of 8,709€ in economic benefits derived by participants from the provision of urban cleanliness every year. In the time dimension of the study, a pro-environmental attitude and behavior, and high education positively correlate to the binary decision to participate or not in a public cleanup, and to the maximum number of participations hours. Age presents a negative linear correlation with WTC. With respect to the sample mean hours that citizens are WTC in a year, an unforeseeably large figure of 7.65 hours is retrieved. If time is converted into money, via wages, participation in cleanups would provide an economic benefit in the sample of 25,568.40€. Even though many different reasons could be enunciated to explain this result, it unmistakably points to the fact that citizens are hugely inclined to contribute with their time, rather than with their money, to the removal of existing litter in the city of Zaragoza.

Overall, the present work contributes with a rigorous application of the SP methods to the context of litter and waste pollution, accounting for a wide set of factors such as protest behavior, that are not generally controlled for in the literature. Secondly, a great number of knowledge gaps have been identified in the field of CB methods and time contributions that should be explored in future studies if the objective is to reach a general theory to able biases in survey schemes. Thirdly, two sound policy proposals are conducted in the study, which are supported by the majority of the sample, and that should be taken seriously if litter is to eradicated from the city of Zaragoza. Fourthly, despite the raw welfare computations, it is very evident that time contributions in the form of participation in cleanups deliver large economic benefits, and in the context of dire economic conditions, citizens' direct participation and implication should be harnessed to tackle long-standing urban challenges.

7 Bibliography

Aadland, D., Caplan, A.J., 2006. Cheap talk reconsidered: New evidence from CVM. Journal of Economic Behavior & Organization 60, 562–578. https://doi.org/10.1016/j.jebo.2004.09.006

Abas, M.A., Hassin, N.H., Hambali, K.A., Karim, M.F.A., Hussin, H., Ismail, L., Fitriani, N., 2021. Public satisfaction and willingness to pay (WTP) for better solid waste management services in the rural area of Kelantan, Malaysia. IOP Conf. Ser.: Earth Environ. Sci. 756, 012083. https://doi.org/10.1088/1755-1315/756/1/012083

Abate, T.G., Börger, T., Aanesen, M., Falk-Andersson, J., Wyles, K.J., Beaumont, N., 2020. Valuation of marine plastic pollution in the European Arctic: Applying an integrated choice and latent variable model to contingent valuation. Ecological Economics 169, 106521. https://doi.org/10.1016/j.ecolecon.2019.106521

Adam, I., 2021. Tourists' perception of beach litter and willingness to participate in beach clean-up. Marine Pollution Bulletin 170, 112591. https://doi.org/10.1016/j.marpolbul.2021.112591

Adamowicz, W., Louviere, J.J., Williams, M., 1994. Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities. Journal of Environmental Economics and Management 26, 271–292. https://doi.org/10.1006/jeem.1994.1017

Alberini, A., Zanatta, V., Rosato, P., 2007. Combining actual and contingent behavior to estimate the value of sports fishing in the Lagoon of Venice. Ecological Economics 61, 530–541. https://doi.org/10.1016/j.ecolecon.2006.05.003

Andreoni, J., 2006. Chapter 18 Philanthropy, in: Kolm, S.-C., Ythier, J.M. (Eds.), Handbook of the Economics of Giving, Altruism and Reciprocity, Applications. Elsevier, pp. 1201–1269. https://doi.org/10.1016/S1574-0714(06)02018-5

Andreoni, J., G. Gale, W., Karl Scholz, J., 1996. Charitable contributions of time and money. University of Wisconsin-Madison Working Paper.

Ansink, E., Weikard, H.-P., Withagen, C., 2019. International environmental agreements with support. Journal of Environmental Economics and Management 97, 241–252. https://doi.org/10.1016/j.jeem.2018.02.001

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., 1993. Report of the NOAA Panel on Contingent Valuation 67.

Auten, G.E., Sieg, H., Clotfelter, C.T., 2002. Charitable Giving, Income, and Taxes: An Analysis of Panel Data. American Economic Review 92, 371–382. https://doi.org/10.1257/000282802760015793

Bateman, I.J., Cole, M., Cooper, P., Georgiou, S., Hadley, D., Poe, G.L., 2004. On visible choice sets and scope sensitivity. Journal of Environmental Economics and Management 47, 71–93. https://doi.org/10.1016/S0095-0696(03)00057-3

Bateman, I.J., Langford, I.H., Jones, A.P., Kerr, G.N., 2001. Bound and path effects in double and triple bounded dichotomous choice contingent valuation. Resource and Energy Economics 23, 191–213. https://doi.org/10.1016/S0928-7655(00)00044-0

Bauer, T.K., Bredtmann, J., Schmidt, C.M., 2013. Time vs. money — The supply of voluntary labor and charitable donations across Europe. European Journal of Political Economy 32, 80–94. https://doi.org/10.1016/j.ejpoleco.2013.06.006

Bénabou, R., Tirole, J., 2006. Incentives and Prosocial Behavior. American Economic Review 96, 1652–1678. https://doi.org/10.1257/aer.96.5.1652

Bouma, J.A., Koetse, M.J., 2019. Mind the Gap: Stated versus Revealed Donations and the Differential Role of Behavioral Factors. Land Economics 95, 225–245. https://doi.org/10.3368/le.95.2.225

Boyle, K.J., Bishop, R.C., 1988. Welfare Measurements Using Contingent Valuation: A Comparison of Techniques. American Journal of Agricultural Economics 70, 20–28. https://doi.org/10.2307/1241972

Boyle, K.J., Bishop, R.C., Welsh, M.P., 1985. Starting Point Bias in Contingent Valuation Bidding Games. Land Economics 61, 188–194. https://doi.org/10.2307/3145811

Boyle, K.J., Johnson, F.R., McCollum, D.W., 1997. Anchoring and Adjustment in Single-Bounded, Contingent-Valuation Questions. American Journal of Agricultural Economics 79, 1495–1500. https://doi.org/10.2307/1244370

Brouwer, R., Hadzhiyska, D., Ioakeimidis, C., Ouderdorp, H., 2017. The social costs of marine litter along European coasts. Ocean & Coastal Management 138, 38–49. https://doi.org/10.1016/j.ocecoaman.2017.01.011

Brouwer, R., Martín-Ortega, J., 2012. Modeling self-censoring of polluter pays protest votes in stated preference research to support resource damage estimations in environmental liability. Resource and Energy Economics 34, 151–166.

Calia, P., Strazzera, E., 2001. A sample selection model for protest responses in contingent valuation analyses. Statistica 61, 473–85.

Cappellari, L., Ghinetti, P., Turati, G., 2011. On time and money donations. The Journal of Socio-Economics 40, 853–867. https://doi.org/10.1016/j.socec.2011.08.013

Carmi, N., 2019. On Social Distress, Littering and Nature Conservation: The Case of Jisr A-Zarka. Coastal Management 47, 347–361. https://doi.org/10.1080/08920753.2019.1598223

Carson, R.T., 2012. Contingent Valuation: A Practical Alternative When Prices Aren't Available. Journal of Economic Perspectives 26, 27–42. https://doi.org/10.1257/jep.26.4.27

Carson, R.T., Flores, N.E., Meade, N.F., 2001. Contingent Valuation: Controversies and Evidence. Environmental and Resource Economics 19, 38.

Carson, R.T., Groves, T., 2007. Incentive and informational properties of preference questions. Environ Resource Econ 37, 181–210. https://doi.org/10.1007/s10640-007-9124-5

Champ, P.A., Moore, R., Bishop, R.C., 2009. A Comparison of Approaches to Mitigate Hypothetical Bias. Agricultural and Resource Economics Review 38, 166–180. https://doi.org/10.1017/S106828050000318X

Chen, B., Qi, X., 2018. Protest response and contingent valuation of an urban forest park in Fuzhou City, China. Urban Forestry & Greening 29, 68–76.

Collins, A.R., Rosenberger, R.S. (Eds.), 2007. Protest Adjustments in the Valuation of Watershed Restoration Using Payment Card Data. Agricultural and Resource Economics Review. https://doi.org/10.22004/ag.econ.44706

Cummings, R.G., Taylor, L.O., 1999. Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. American Economic Review 89, 649–665. https://doi.org/10.1257/aer.89.3.649

Desvousges, W.H., Gable, A.R., Dunford, R.W., Hudson, S.P., 1993. Contingent valuation: the wrong tool to measure passive-use losses. Choices 8.

Duncan, B., 1999. Modeling charitable contributions of time and money. Journal of Public Economics 72, 213–242. https://doi.org/10.1016/S0047-2727(98)00097-8

Directive (EU) 2018/851 of the European Parliament and of the Council of 30 May 2018 amending Directive 2008/98/EC on waste (Text with EEA relevance), 2018. , OJ L. http://data.europa.eu/eli/dir/2018/851/oj/eng

Eastman, L.B., Núñez, P., Crettier, B., Thiel, M., 2013. Identification of self-reported user behavior, education level, and preferences to reduce littering on beaches – A survey from the SE Pacific. Ocean & Coastal Management 78, 18–24. https://doi.org/10.1016/j.ocecoaman.2013.02.014

Englin, J., Cameron, T.A., 1996. Augmenting travel cost models with contingent behavior data: Poisson regression analyses with individual panel data. Environ Resource Econ 7, 133–147. https://doi.org/10.1007/BF00699288

Eurobarometer, 2022. Standard Eurobarometer 97 - Summer 2022 - Eurobarometer survey. https://europa.eu/eurobarometer/surveys/detail/2693

Feldman, N.E., 2010. Time Is Money: Choosing between Charitable Activities. American Economic Journal: Economic Policy 2, 103–130. https://doi.org/10.1257/pol.2.1.103

Fox, J.A., Shogren, J.F., Hayes, D.J., Kliebenstein, J.B., 1998. CVM-X: Calibrating Contingent Values with Experimental Auction Markets. American Journal of Agricultural Economics 80, 455–465. https://doi.org/10.2307/1244548

Gächter, S., Renner, E., 2010. The effects of (incentivized) belief elicitation in public goods experiments. Exp Econ 13, 364–377. https://doi.org/10.1007/s10683-010-9246-4

Gowdy, J.M., 2004. The Revolution in Welfare Economics and Its Implications for Environmental Valuation and Policy. Land Economics 80, 239–257. https://doi.org/10.2307/3654741

Grijalva, T.C., Berrens, R.P., Bohara, A.K., Shaw, W.D., 2002. Testing the Validity of Contingent Behavior Trip Responses. American Journal of Agricultural Economics 84, 401–414. https://doi.org/10.1111/1467-8276.00306

Haab, T.C., Interis, M.G., Petrolia, D.R., Whitehead, J.C., 2013. From Hopeless to Curious? Thoughts on Hausman's "Dubious to Hopeless" Critique of Contingent Valuation. Applied Economic Perspectives and Policy 35, 593–612. https://doi.org/10.1093/aepp/ppt029

Halstead, J.M., Lindsay, B.E., Brown, C.M., 1991. Use of the Tobit model in contingent valuation: Experimental evidence from the Pemigewasset Wilderness Area. Journal of Environmental Management 33, 79–89. https://doi.org/10.1016/S0301-4797(05)80049-0

Halstead, J.M., Luloff, A.E., Stevens, T.H., 1992. Protest Bidders in Contingent Valuation. Northeastern Journal of Agricultural and Resource Economics 21, 160–169. https://doi.org/10.1017/S0899367X00002683

Hausman, J., 2012. Contingent Valuation: From Dubious to Hopeless. Journal of Economic Perspectives 26, 43–56. https://doi.org/10.1257/jep.26.4.43

Hardin, G., 1968. The Tragedy of the Commons. Science 162, 1243–1248. https://doi.org/10.1126/science.162.3859.1243

Herriges, J.A., Shogren, J.F., 1996. Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning. Journal of Environmental Economics and Management 30, 112–131. https://doi.org/10.1006/jeem.1996.0008

Holmes, T.P., Kramer, R.A., 1995. An Independent Sample Test of Yea-Saying and Starting Point Bias in Dichotomous-Choice Contingent Valuation. Journal of Environmental Economics and Management 29, 121–132. https://doi.org/10.1006/jeem.1995.1035

Huffman, K.T., Grossnickle, W.F., Cope, J.G., Huffman, K.P., 1995. Litter Reduction: A Review and Integration of the Literature. Environment and Behavior 27.

Jacquemet, N., James, A., Luchini, S., Shogren, J.F., 2017. Referenda Under Oath. Environ Resource Econ 67, 479–504. https://doi.org/10.1007/s10640-016-0023-5

Jeon, Y., Herriges, J.A., 2010. Convergent Validity of Contingent Behavior Responses in Models of Recreation Demand. Environ Resource Econ 45, 223–250. https://doi.org/10.1007/s10640-009-9313-5

Johnston, R.J., Boyle, K.J., Adamowicz, W. (Vic), Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., Vossler, C.A., 2017. Contemporary Guidance for Stated Preference Studies. Journal of the Association of Environmental and Resource Economists 4, 319–405. https://doi.org/10.1086/691697

Jorgensen, B.S., Syme, G.J., 2000. Protest responses and willingness to pay: attitude toward paying for stormwater pollution abatement. Ecological Economics 33, 251–265.

Jorgensen, B.S., Syme, G.J., Bishop, B.J., Nancarrow, B.E., 2000. Protest Responses in Contingent Valuation. Ecological Economics 33, 20.

Kahneman, D., Knetsch, J.L., 1992. Valuing public goods: The purchase of moral satisfaction. Journal of Environmental Economics and Management 22, 57–70. https://doi.org/10.1016/0095-0696(92)90019-S

Khanna, J., Sandler, T., 2000. Partners in giving:: The crowding-in effects of UK government grants. European Economic Review 44, 1543–1556. https://doi.org/10.1016/S0014-2921(99)00007-0

Kirchgässner, G., 2010. On minimal morals. European Journal of Political Economy, Special Issue on Ethics and Economics 26, 330–339. https://doi.org/10.1016/j.ejpoleco.2010.03.003

Kling, C.L., Phaneuf, D.J., Zhao, J., 2012. From Exxon to BP: Has Some Number Become Better than No Number? Journal of Economic Perspectives 26, 3–26. https://doi.org/10.1257/jep.26.4.3

Koop, S.H.A., van Leeuwen, C.J., 2017. The challenges of water, waste, and climate change in cities. Environ Dev Sustain 19, 385–418. https://doi.org/10.1007/s10668-016-9760-4

Kotchen, M.J., Reiling, S.D., 1999. Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? Another Comment. Land Economics 75, 478–482. https://doi.org/10.2307/3147192

Kriström, B., 1997. Spike Models in Contingent Valuation. American Journal of Agricultural Economics 79, 1012–1023.

Lavers, J.L., Bond, A.L., 2017. Exceptional and rapid accumulation of anthropogenic debris on one of the world's most remote and pristine islands. Proceedings of the National Academy of Sciences 114, 6052–6055. https://doi.org/10.1073/pnas.1619818114

Lindhjem, H., Navrud, S., 2011. Are Internet surveys an alternative to face-to-face interviews in contingent valuation? Ecological Economics, Special Section - Governing the Commons: Learning from Field and Laboratory Experiments 70, 1628–1637. https://doi.org/10.1016/j.ecolecon.2011.04.002

Lo, A.Y., Jim, C.Y., 2015. Protest response and willingness to pay for culturally significant urban trees: Implications for Contingent Valuation Method. Ecological Economics 114, 58–66.

Loomis, J., 2011. What's to Know About Hypothetical Bias in Stated Preference Valuation Studies? Journal of Economic Surveys 25, 363–370. https://doi.org/10.1111/j.1467-6419.2010.00675.x

Loomis, J., Gonzalez-Caban, A., Gregory, R., 1994. Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? Land Economics 70, 499–506. https://doi.org/10.2307/3146643

Loomis, J., Santiago, L., 2013. Economic Valuation of Beach Quality Improvements: Comparing Incremental Attribute Values Estimated from Two Stated Preference Valuation Methods. Coastal Management 41, 75–86. https://doi.org/10.1080/08920753.2012.749754

Louviere, J.J., Hensher, D.A., Swait, J.D., 2000. Stated Choice Methods: Analysis and Applications. Cambridge University Press.

Lucrezi, S., Digun-Aweto, O., 2020. "Who wants to join?" Visitors' willingness to participate in beach litter clean-ups in Nigeria. Marine Pollution Bulletin 155, 111167. https://doi.org/10.1016/j.marpolbul.2020.111167

McFadden, D., 2017. Stated preference methods and their applicability to environmental use and non-use valuations. Contingent Valuation of Environmental Goods 153–187.

Menchik, P.L., Weisbrod, B.A., 1987. Volunteer labor supply. Journal of Public Economics 32, 159–183. https://doi.org/10.1016/0047-2727(87)90010-7

Meyerhoff, J., Liebe, U., 2006. Protest beliefs in contingent valuation: Explaining their motivation. Ecological Economics 57, 583–594.

Milon, J.W., 1989. Contingent valuation experiments for strategic behavior. Journal of Environmental Economics and Management 17, 293–308.

Mitchell, R.C., Carson, R.T., 1989. Using Surveys to Value Public Goods: The Contingent Valuation Method, 1st Edition. ed. RFF Press, New York. https://doi.org/10.4324/9781315060569

Muñoz-Cadena, C.E., Lina-Manjarrez, P., Estrada-Izquierdo, I., Ramón-Gallegos, E., 2012. An Approach to Litter Generation and Littering Practices in a Mexico City Neighborhood. Sustainability 4, 1733–1754. https://doi.org/10.3390/su4081733

Murphy, J.J., Allen, P.G., Stevens, T.H., Weatherhead, D., 2005. A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation 13.

Murphy, J.J., Stevens, T.H., n.d. Is Cheap Talk Effective at Eliminating Hypothetical Bias in a Provision Point Mechanism? 17.

Novotny, T.E., Lum, K., Smith, E., Wang, V., Barnes, R., 2009. Cigarettes Butts and the Case for an Environmental Policy on Hazardous Cigarette Waste. International Journal of Environmental Research and Public Health 6, 1691–1705. https://doi.org/10.3390/ijerph6051691

Noy, C., 2008. Sampling Knowledge: The Hermeneutics of Snowball Sampling in Qualitative Research. International Journal of Social Research Methodology 11, 327–344. https://doi.org/10.1080/13645570701401305

Parker, C., Scott, S., Geddes, A., 2019. Snowball sampling. SAGE Publications Ltd.
Schuhmann, P.W., Bass, B.E., Casey, J.F., Gill, D.A., 2016. Visitor preferences and willingness to pay for coastal attributes in Barbados. Ocean & Coastal Management 134, 240–250. https://doi.org/10.1016/j.ocecoaman.2016.09.020

Shen, M., Mao, D., Xie, H., Li, C., 2019. The Social Costs of Marine Litter along the East China Sea: Evidence from Ten Coastal Scenic Spots of Zhejiang Province, China. Sustainability 11, 1807. https://doi.org/10.3390/su11061807

Smith, V.K.K., Zhang, X., Palmquist, R.B., 1997. Marine Debris, Beach Quality, and Non-Market Values. SSRN Journal. https://doi.org/10.2139/ssrn.31975

Smith, W., Chang, C., 2002. Shipping the good apples out: a note on contributions of time and money. Economics Bulletin 10, 1–14.

Spash, C.L., 2006. Non-Economic Motivation for Contingent Values: Rights and Attitudinal Beliefs in the Willingness To Pay for Environmental Improvements. Land Economics 82, 602–622. https://doi.org/10.3368/le.82.4.602

Stock, J.H., Watson, M.W., 2015. Introduction to econometrics 3rd ed. Pearson Education.

Strazzera, E., Genius, M., Scarpa, R., Hutchinson, W.G., 2001. The Effect of Protest Votes on the Estimates of Willingness to Pay for Use Values of Recreational Sites. SSRN Journal. https://doi.org/10.2139/ssrn.293695
Strazzera, E., Scarpa, R., Calia, P., Garrod, G.D., Willis, K.G., 2003. Modeling zero values and protest responses in contingent valuation surveys. Applied Economics 35, 133–138. https://doi.org/10.1080/0003684022000015900

"The Litter Challenge." Clean Europe Network, 27 July 2015, http://www.cleaneuropenetwork.eu/en/the-litter-challenge/auw/

Tyllianakis, E., Ferrini, S., 2021. Personal attitudes and beliefs and willingness to pay to reduce marine plastic pollution in Indonesia. Marine Pollution Bulletin 173, 113120. https://doi.org/10.1016/j.marpolbul.2021.113120

Venkatachalam, L., 2004. The contingent valuation method: a review. Environmental Impact Assessment Review 24, 89–124. https://doi.org/10.1016/S0195-9255(03)00138-0

Vossler, C.A., Doyon, M., Rondeau, D., 2012. Truth in Consequentiality: Theory and Field Evidence on Discrete Choice Experiments. American Economic Journal: Microeconomics 4, 145–171. https://doi.org/10.1257/mic.4.4.145

Whitehead, J.C., Pattanayak, S.K., Van Houtven, G.L., Gelso, B.R., 2008. Combining revealed and stated preference data to estimate the non market value of ecological services: An assessment of the state of the science: Combining revealed and stated preference data. Journal of Economic Surveys 22, 872–908. https://doi.org/10.1111/j.1467-6419.2008.00552.x

Whitehead, J.C., Phaneuf, D.J., Dumas, C.F., Herstine, J., Hill, J., Buerger, B., 2010. Convergent Validity of Revealed and Stated Recreation Behavior with Quality Change: A Comparison of Multiple and Single Site Demands. Environ Resource Econ 45, 91–112. https://doi.org/10.1007/s10640-009-9307-3

Wright, K., 2001. Generosity vs. Altruism: Philanthropy and Charity in the United States and United Kingdom. VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations 12, 399–416. https://doi.org/10.1023/A:1013974700175

Xie, L., Adamowicz, W., Kecinski, M., Fooks, J.R., 2022. Using economic experiments to assess the validity of stated preference contingent behavior responses. Journal of Environmental Economics and Management 114, 102659. https://doi.org/10.1016/j.jeem.2022.102659

Zambrano-Monserrate, M.A., Ruano, M.A., 2020. Estimating the damage cost of plastic waste in Galapagos Islands: A contingent valuation approach. Marine Policy 117, 103933. https://doi.org/10.1016/j.marpol.2020.103933