# Elicitation formats and the WTA/WTP gap: A study of climate neutral foods

Andreas C. Drichoutis and Jayson Lusk and Valentina Pappa

Agricultural University of Athens, Oklahoma State University, University of Ioannina

8. May 2014

# Elicitation formats and the WTA/WTP gap: A study of climate neutral foods

Andreas C. Drichoutis*[1], Jayson L. Lusk†[2], and Valentina Pappa‡[3]

[1]Agricultural University of Athens
[2]Oklahoma State University
[3]University of Ioannina

First Draft: May 8, 2014

**Abstract:** We conduct a field valuation experiment where we vary the valuation method (contingent valuation vs. inferred valuation) as well as the payment format (dichotomous choice vs. payment card). Willingness-to-accept and willingness-to-pay valuations are elicited in a within-subjects design for foods with climate neutral labels. We find a similar gap for valuations elicited with the contingent or the inferred valuation method. However, we also find that the gap can be muted by using a payment card elicitation format.

**Keywords:** willingness to pay; willingness to accept; contingent valuation; inferred valuation; payment card; single bounded.

**JEL Classification Numbers:** D12; C93.

## 1 Introduction

Economic theory posits that with small income effects and many available substitutes, the economic value people place on a good should be independent of whether they own it (Hanemann, 1991). One of the more popular anomalies is the observed divergence between two measures of economic value that reflect ownership: willingness to pay (WTP) to obtain

*Lecturer, Department of Agricultural Economics, Agricultural University of Athens, Iera Odos 75, 11855, Greece, e-mail: adrihout@aua.gr.

†Regents Professor & Willard Sparks Endowed Chair, Department of Agricultural Economics, Oklahoma State University, Stillwater, OK 74078, USA, e-mail:jayson.lusk@okstate.edu.

‡Graduated student, Department of Economics, University of Ioannina e-mail: pappaval@gmail.com.

a good and the willingness to accept compensation (WTA) to forsake the good. The divergence between these two values, known as the WTA/WTP gap, is such a common empirical regularity that has led Horowitz and McConnell (2002) to state that "The pervasiveness of high WTA/WTP ratios and the wide variety of goods that have been used in the experiments have combined to sustain interest in WTA vs. WTP for roughly 30 years."

The Contingent valuation (CV) method has become the most popular method to measure WTA and WTP, and these values are used to valuing the benefits of a new goods, services, or amenities. Although the CV method was principally developed in environmental and transport economics, it has made considerable headway in the valuation of food products over the last decades (e.g., Buzby et al., 1998; Corsi, 2007). Most, if not all, CV studies are conducted in hypothetical contexts, particularly in environmental valuation studies where a real market with salient payments is difficult to establish and some number is considered better than no number at all (Carson, 2012; Haab et al., 2013; Kling et al., 2012).

While there have been many different studies exploring the determinants of the WTA/WTP gap in the context of CV studies, herein, we examine the effect of two largely unexplored factors. First, we explore how an alternative to the CV method, the Inferred valuation (IV) method[1], that was designed to mitigate social desirability bias can affect valuations and consequently the gap. Second, we explore how different elicitation formats can affect the propensity of individuals to report valuations that lead to a WTA/WTP gap.

There are several motivations for examining whether the WTA/WTP gap is affected by the valuation method. First, there are now several studies that show that when people are asked to predict other people's value, as in the IV method, they state a different value than their own. This would imply that valuations elicited under the CV method are different than valuations elicited under the IV method, which could significantly affect the gap.[2] For example, in van Boven et al. (2000), owners of an item overestimated buyers WTP and buyers underestimated owners WTA such that owners and buyers underestimated the endowment effect by 40%. In Loewenstein and Adler (1995), subjects underestimated WTA when they had to predict for themselves which resulted in underestimation of the impact of the endowment effect by 84%. Similarly, van Boven et al. (2003) find that buyers underestimated the endowment effect and submitted bids sub-optimally lower than owners' WTA.

[1] A term coined by Lusk and Norwood (2009a,b).

[2] Even if we assume a symmetric unidirectional effect of IV on WTA and WTP e.g., of $k$ magnitude, then the WTA/WTP gap under IV would be: $WTA_{IV}/WTP_{IV} = (WTA_{CV} - k)/(WTP_{CV} - k) \neq WTA_{CV}/WTP_{CV}$. On the other hand, if the effect is not additive but multiplicative then it would be: $WTA_{IV}/WTP_{IV} = (WTA_{CV} \cdot k)/(WTP_{CV} \cdot k) = WTA_{CV}/WTP_{CV}$, which is the only case where we would observe no effect on the WTA/WTP gap. Even worse, if the effect is not symmetric e.g., it reduces WTP by $k_1$ and WTA by $k_2$, then there is no way that the gap remains constant across methods. Thus, we would expect that the gap will differ across methods more of the times.

Lusk and Norwood (2009a,b) argued that the IV method generates valuations that are less likely to suffer from normative or moral response biases (such as social desirability bias), and they found that responses to the IV method better predicted actual shopping behavior than did those from a CV method. They also found that the IV method produced less hypothetical bias when social desirability was present. The authors showed that goods with normative dimensions are more prone to social desirability bias and thus the IV method is more effective in bridging the gap between the laboratory and field valuations. In contrast to Lusk and Norwood (2009a,b), Frederick (2012) found that subjects tend to overestimate predicted WTP in real and hypothetical valuations. Frederick (2012) also found that predicted WTA was not statistically different than own WTA. Note, that the products used in Frederick (2012) lack a strong social desirability dimension which may explain why his result is the opposite of that obtained in Lusk and Norwood (2009a,b). The results in Kurt and Inman (2013) also suggest a smaller WTA/WTP gap. They found that owners predict a lower WTA for other owners while buyers predict a higher WTP for other buyers. In the spirit of Lusk and Norwood (2009a,b), Pronin (2007) argues that people tend to recognize biases in human judgment except when that bias is their own which implies that predictions over other peoples' preferences should mitigate biases.

With respect to elicitation formats, in the CV literature it is generally reported that using different elicitation formats results in different valuations. The open-ended format was criticized by the National Oceanographic and Atmospheric Administration (Arrow et al., 1993) as providing 'erratic and biased' responses. The NOAA panel suggested the dichotomous choice (DC) format which became the favored approach for several years. Other alternatives that were not considered by the NOAA panel were later developed including the payment card (PC) format. Payment cards have made considerable headway in the valuation literature due to several results that showed it exhibits more desirable properties than DC (Reaves et al., 1999), less 'yeah-saying' at high bid amounts (Zhongmin et al., 2006) and results in more conservative estimates (Blaine et al., 2005). Due to all this properties, it is now considered by far the most common format for CV studies in the health economics literature (Smith and Sach, 2010).[3] Donaldson et al. (1997) argued in favor of the PC format due to its resemblance to every-day behavior (individuals 'shop around', observe different values for a good and choose the one that suits them most). As a result cognitive demand is potentially mitigated and the validity of the instrument is potentially increased.

Carson and Groves (2007) offer a typology of elicitation formats that starts with the

[3]It's not a surprise that the payment card format has found its way into food (Brummett et al., 2007; Hsu et al., 2009; Aizaki et al., 2011; Hu et al., 2011) as well as resource economics studies (Lienhoop and MacMillan, 2007; Brouwer et al., 2008; Solomon and Johnson, 2009; Simpson and Hanna, 2009) as well.

single binary discrete choice format as the most basic format.[4] The DC format only provides a discrete indicator of subject's valuation. Early approaches in the literature tried to capture subject's valuation in continuous terms (using open-ended questions) albeit in the cost of having to drop a significant number of observations due to non-responses or 'protest zeros'. The payment card was developed as an attempt to overcome these problems and since it can come close to achieving a valuation response in continuous terms, it can be thought of as the polar opposite of the DC format.

The DC format has been favored due to its well known property of incentive compatibility. This is due to the Gibbard-Satterthwaite theorem (Gibbard, 1973; Satterthwaite, 1975) which states that for the case of more than two alternatives (i.e., non-DC formats) no non-dictatorial strategy-proof voting procedure exists.[5] This is just to say that any response format with at least three possible outcomes is subject to individual manipulation (i.e., it's not incentive compatible). This does not imply that any binary DC format is incentive compatible but that, by elimination, only a DC format could be incentive compatible, assuming subjects believe their response is consequential (meaning there is some probability the respondent's answer will actual influence the provision fo the good). While proponents of the DC format take this result about the incentive compatibility of the DC format as granted for any type of good, Carson et al. (1997) show that the DC format is not incentive compatible in the case of provision of a new private or quasi-public good. The incentive compatibility of the DC format can be restored for quasi-public or private goods only if the binary choice is between two different forms of the good, so that the valuation question represents a change in the good (Carson et al., 2001; Carson and Groves, 2007).

To address the issues discussed above, in this study we explore whether the WTA/WTP gap can be muted under the inferred valuation method. We also ask whether the DC format performs better than a payment card. As our valuation products we use climate neutral foods, which entail a significant social desirability dimension. We elicit consumers' WTP to exchange a conventional food product with a climate neutral counterpart as well as WTA to forsake the climate neutral food in favor of the conventional counterpart in a within-subjects design that varies the order of the valuation scenarios. We find that the use of IV rather than CV has virtually no effect in the WTA/WTP gap but that the elicitation method (DC vs. PC) can mute the gap. To our surprise, a WTA/WTP gap surfaces under the DC format but is muted with the payment card format. The next section describes our experimental design. We next present some descriptive results and econometric analysis and conclude in

[4] Carson et al. (1997) have shown that all of the commonly used preference elicitation formats can be seen as generalizations of the DC format.

[5] The theorem was formalized by Gibbard (1973) and Satterthwaite (1975) and noted in passing by Dummett and Farquharson (1961).

the final section.

## 2 Experimental design

The field experiment was carried out in supermarkets located in Ioannina, Greece. The experimenter (i.e., one of the authors) approached each participant and invited him/her to participate voluntarily. The experiment was conducted outside the main entrance of 9 different supermarket chain stores. We approached 1162 persons and got 512 agreements to participate in the survey and 650 refusals, resulting in a response rate of 44.06%.[6]

Our experimental design combines elements of a within- and a between-subject design (Charness et al., 2012). We elicit valuations for the premium of two food products that carry climate neutral labels.[7] The products chosen were a bottle (1 L) of olive oil and a 6-pack of eggs. Valuations for these products were elicited between subjects, that is, for half of the subjects we elicited valuations for olive oil and for the other half we elicited valuations for eggs. We specifically chose to value products with a climate neutral label because this label entail an environmental protection attribute which is largely socially desirable. This renders appropriate the use of the Inferred valuation method (IV) which allows us to compare valuations with the CV method. The valuation method treatment (i.e., inferred valuation vs. contingent valuation) was employed as part of a between subjects design. Therefore, for half of the subjects valuations were elicited with CV and for the other half valuations were elicited with IV. We also varied the elicitation method between subjects by eliciting valuations for half of the subjects using a payment card (PC) and for the other half using a dichotomous choice (DC) format.

To explore the WTA/WTP gap, we elicited valuations for WTP and WTA within subjects, that is, each person was asked both a WTP and a WTA valuation question. The order of the elicitation of valuations was counterbalanced to avoid any order effects (e.g., Harrison et al., 2005). Table 1 exhibits the experimental design followed for each of the two food products. Each cell of Table 1 was administered to 64 subjects (32 subjects for olive oil and

[6]Of the 650 refusals, 467 (71.8%) were females and 183 (28.2%) were males. These percentages are virtually the same with the gender composition of the sample that accepted to participate in the survey (73.4% females, 26.6% males). Thus, we can safely rule out selection effects due to gender.

[7]The climate-neutral label distinguishes organisations that offset their $CO_2$ emissions in *myclimate* carbon offset projects. The non-profit foundation *myclimate* is an international initiative with Swiss origins founded in 2002 as an ETH Zurich spinoff regarding voluntary carbon offsetting measures. It offers a comprehensive package of services for offsetting $CO_2$ emissions using an international network of project partners and representatives who act on behalf of *myclimate* in their countries. *myclimate* develops and supports projects around the world that directly reduce greenhouse gases. The foundation calculates the climate-impacting emissions produced by activities and products and provides consultation services to businesses regarding the optimal way to offset $CO_2$ according to the principle of "Avoid - Reduce - Offset".

32 subjects for eggs). The exact wording of the valuation questions for each treatment are reproduced in Appendix A. In the payment card format both WTP and WTA were elicited over the same payment card. The payment card for olive oil spanned a range up to €2.75 in 25 cents intervals (see Table B.1 in the Appendix). Similarly, the payment card for eggs spanned a range up to €1.10 in 10 cents intervals (see Table B.2 in the Appendix). Range of prices in the payment cards was decided after surveying actual market prices of conventional and organic olive oil and eggs, respectively. The bids in the DC format were varied over a range of prices that overlap with the payment cards. Therefore, the bids for olive oil were varied over {50, 100, 150, 200, 250} and for eggs were varied over {20, 40, 60, 80, 100}.

Table 1: Experimental design

| | | **Elicitation format** | |
|---|---|---|---|
| | | *Payment Card* | *Discrete choice* |
| **Method** | *Contingent valuation* | WTP then WTA | WTP then WTA |
| | | WTA then WTP | WTA then WTP |
| | *Inferred valuation* | WTP then WTA | WTP then WTA |
| | | WTA then WTP | WTA then WTP |

Besides standard demographic information, the questionnaire also elicited respondents' beliefs about the likelihood of hypothetical bias. Subjects were asked to indicate: i) given that people tend to exaggerate responses in hypothetical surveys when there is no real economic exchange, how likely it is that the same thing happened in this survey ii) how likely it is that other respondents will exaggerate their responses. Responses were recorded on a 5-point Likert scale anchored by extremely unlikely and extremely likely. In addition, given that the climate neutral products carry a socially desirable attribute (i.e., "protecting the environment") and subjects may be inclined to state a higher WTA/WTP, we also elicited the social desirability scale (SDS) of Stöber (2001). The final instrument is composed of 16 questions (e.g., "I always admit my mistakes openly and face the potential negative consequences", "I occasionally speak badly of others behind their back" etc.). Subjects were asked to state their level of agreement with each of the statements. The statements are then individually scored and summed to form a single index of social desirability. The higher the score on the SDS scale, the more likely it is the person falls prey to social desirability bias.

Two additional questions elicited subjects' price sensitivity and normative motivations. Subjects were asked to state on a 5-point Likert scale the importance of prices in their food decision making, anchored by not important at all and very important. In addition, they were asked to indicate their level of agreement to the statement "buying climate neutral olive oil [eggs] is the right thing to do" on a 5-point Likert scale anchored by totally disagree

and totally agree. Several demographic information were also recorded such as age, gender, education level, income level etc.

# 3 Results

Before we proceed with the analysis and results, we first provide some information regarding the demographic profile of our sample. The vast majority of the interviewees were females (73.4%). This is not as problematic as it may seem on first glance, given that primary shoppers are mainly females. For example, one study estimates that 75% of principal household shoppers in the US are females (Mediamark Research and Intelligence, 2009). Therefore, the gender composition of our sample is not representative of the population of the city but it might better represent the grocery shopping population. On average our subjects were 43.4 years old (S.D.=10.48, min=18, max=68) and had a household size of 3.6 (S.D.=1.06, min=1, max=6). The majority of our sample had a university degree (58.6%) followed by senior hi-school graduates (21.7%). Finally, the vast majority of our sample (85.7%) stated they were the major grocery shoppers of their household.

## 3.1 Descriptive analysis

Given that climate neutral products carry a highly socially desirable attribute, we first examine whether the CV method generates different estimates than the IV method. Our hypothesis is that the IV method would better manage to mitigate social desirability by generating less exaggerated valuations. Figure 1 shows distributions of responses in the payment cards by product and valuation method. Overall, it looks as if there are more people stating a valuation in the rightmost cells of the payment card under the CV method than under the IV method. This implies a higher WTP and WTA valuation for CV than IV. Kolmogorov-Smirnov tests on whether the distributions of responses on payment cards cells are similar across CV and IV, fails to reject the null for olive oil (p-value for WTP responses=0.788, p-value for WTA responses=0.788) and rejects the null for the WTA distribution at the 10% level for eggs (p-value for WTP responses=0.104, p-value for WTA responses=0.066).

When we look at the DC data, Figure 2 and Figure 3 show no clear pattern as per whether CV and IV generate different valuations. A proportions test of whether percentages of affirmative responses are equal across methods fails to reject the null in all cases (Olive oil: p-value for WTP responses=0.465, p-value for WTA responses=0.720, Eggs: p-value for WTP responses=0.127, p-value for WTA responses=0.703). Overall, the evidence of whether IV can mitigate social desirability bias are marginally in favor of the IV method.

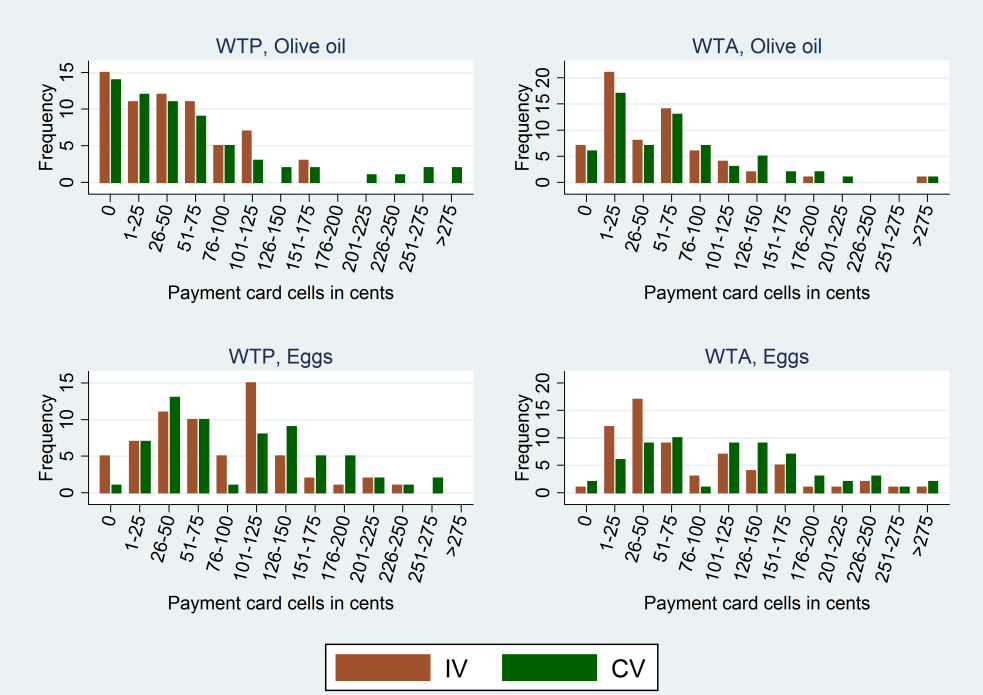


Figure 1: Distribution of responses in payment cards

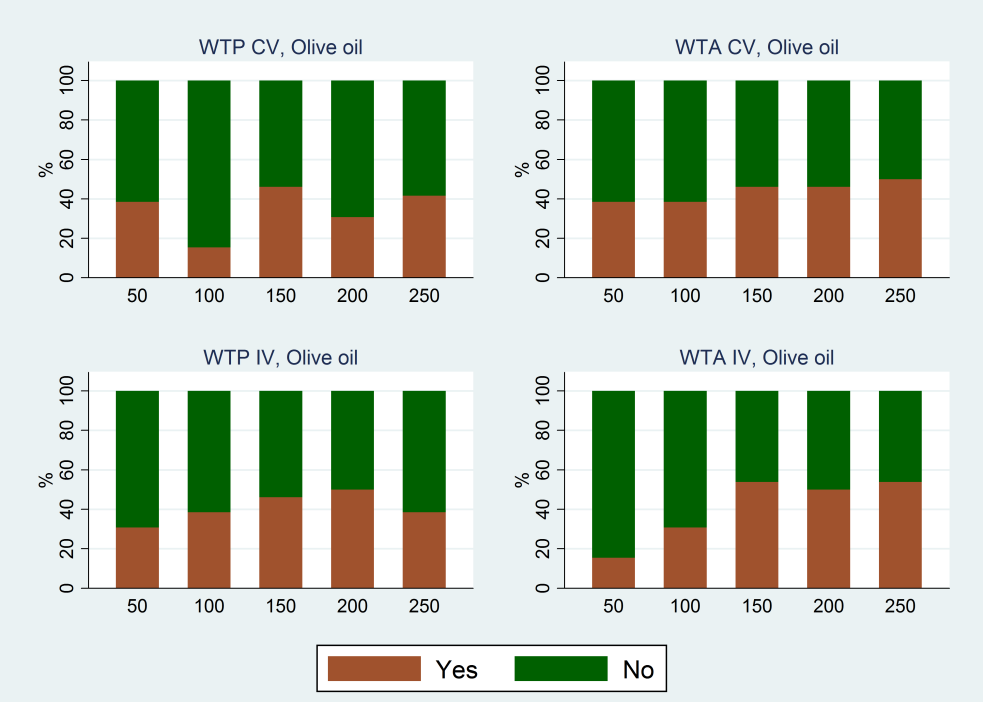


Figure 2: Percentage of Yes/No in the discrete choice format

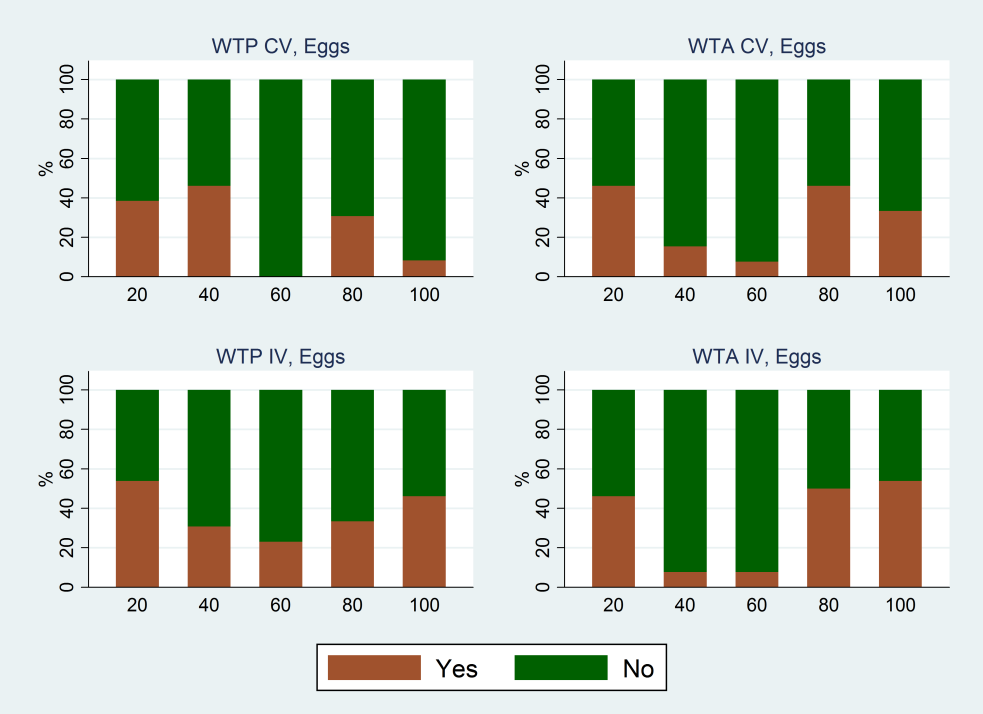


Figure 3: Percentage of Yes/No in the discrete choice format

Figures 4 and 5 show responses in the PC and DC format pooled over CV and IV. In the PC format the distribution of WTP responses roughly follows the distribution of WTA responses. Kolmogorov-Smirnov tests confirm that the distributions of responses on payment cards cells are similar across WTP and WTA (p-value for olive oil=0.788, p-value for eggs=0.951). Thus, there is no indication of a gap with the PC data. Figure 5 shows that the proportions of affirmative responses are roughly the same across WTP and WTA. However, an affirmative response to a WTP question provides a lower bound for the valuation while an affirmative response in a WTA question provides an upper bound for the elicited valuation. Therefore, the fact that distributions look alike implies a WTA/WTP gap. In essence, an absence of a gap would require that the proportion of 'Yes' in the WTP question should match the proportion of 'No' in the WTA question. On the contrary, a proportion test of whether the proportion of 'Yes' in the WTP question is equal to the proportion of 'No' in the WTA question, highly rejects the null (p-value for olive oil=0.001, p-value for eggs=0.000). Thus, there is an indication of a WTA/WTP gap with the DC data. [8]

[8] An issue that requires further exploration is that of price responsiveness in the DC format. This is important since it relates to construct validity of the valuation task. Figure 5 does not provide a clear picture on this issue. To gain a better insight, we regressed the discrete choice variable on the bid amount (probit regressions) separately for each product. For WTA, the respective coefficients of the bid amount are positive indicating that the likelihood of accepting the proposed bid increases with the proposed amount. For WTP, the coefficient for eggs is negative indicating that the likelihood of accepting the proposed amount

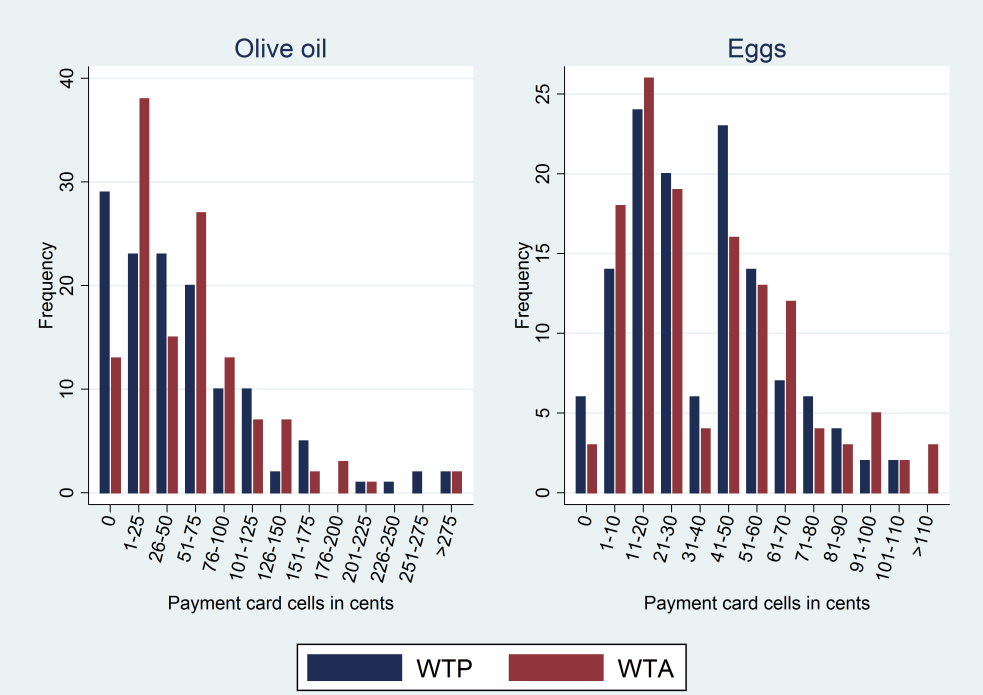


Figure 4: Distribution of responses in payment cards

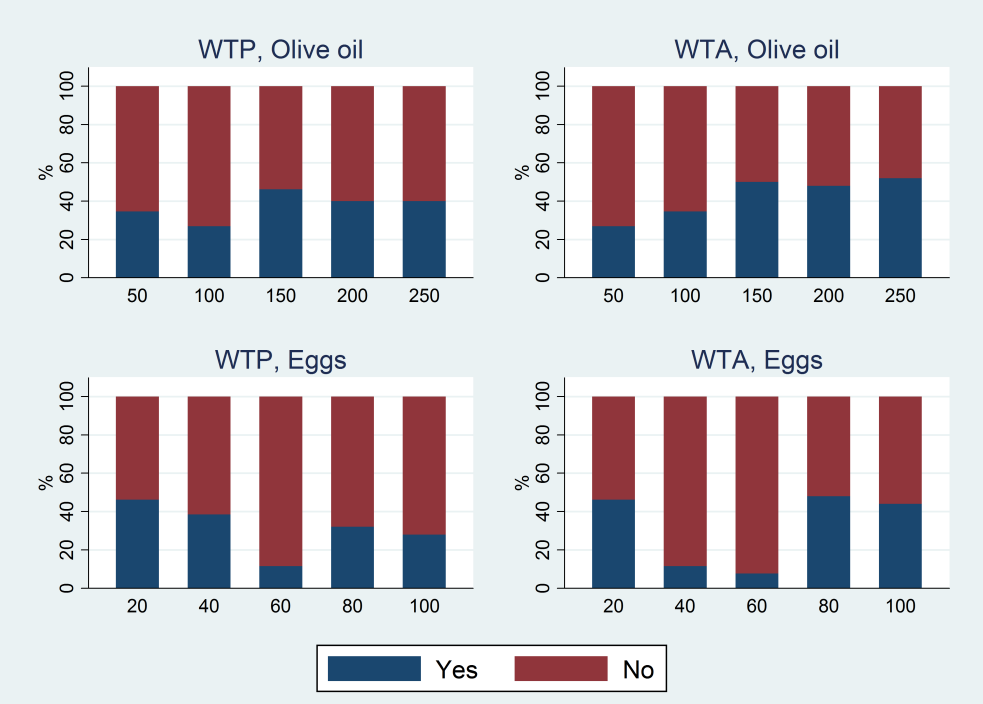


Figure 5: Percentage of Yes/No in the discrete choice format

## 3.2 Econometric analysis

To check whether the results obtained above hold under conditional analysis we estimated interval regression models with clustered standard errors at the individual level to account for the fact that each person provided two valuations responses (WTP and WTA). Table 2 shows coefficient estimates from several specifications. Models (1) to (3) present estimates with our basic treatment variables and models (4) to (6) add a series of demographics and attitudinal variables.[9] As indicated in Table 2, results are fairly robust to the addition of demographic/attitudinal variables. None of our main conclusions changes when we consider the models with or withour the demographic variables. We use several interaction terms between treatment variables to capture differential effects. Interaction terms only complicate slightly the interpretation of coefficients since we need to take into account that the effect of any of the interacted variables depends on the interaction terms as well. For example, the effect of the WTA variable would be $\frac{\partial bid}{\partial WTA} = b_{WTA} + b_{WTA \times CV} CV + b_{WTA \times PCard} PCard$. This expression can then be evaluated for all possible sets of values of $CV$ and $PCard$ i.e., $\{(CV, PCard) = (1,1), (1,0), (0,1), (0,0)\}$.

Table 2: Interval regression estimates

| | Without demographics | | | With demographics | | |
|---|---|---|---|---|---|---|
| | Olive oil | Eggs | Pooled model | Olive oil | Eggs | Pooled model |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 123.844*** | 45.370*** | 59.799*** | 217.597*** | 55.237* | 96.131** |
| | (18.466) | (7.264) | (10.090) | (83.376) | (29.644) | (46.429) |
| Order | -36.036 | -20.205 | -31.644 | -19.591 | -17.891* | -27.732 |
| | (50.266) | (15.516) | (24.940) | (46.410) | (10.439) | (21.616) |
| WTA | 54.760*** | 34.833*** | 55.844*** | 54.327*** | 33.775*** | 55.970*** |

drops with the proposed amount but is positive for olive oil. Thus, more careful consideration is required when interpreting results for WTP for olive oil.

[9] The variable list is as follows. *Order* is a dummy that accounts for the order of the valuation questions, *WTA* is a dummy taking the value of 1 when WTA was elicited and 0 when WTP was elicited, *CV* is a dummy that takes the value of 1 for Contingent Valuation (0 for Inferred Valuation), *PCard* is a dummy indicating whether a PC format was used (=1) or a DC format (=0) and *Product* is a product specific dummy (1=olive oil, 0=eggs). *Shopper* is a dummy indicating whether the subject is the major grocery shopper for his/her household, *Gender* is a dummy for males, *Educ* are educational level dummies (the omitted category is "up to elementary school"), *Income* are income level dummies (the omitted category is when household's economic position is "below average or worse"). *Price sensitive* and *Normative* are dummies indicating whether the subject is price sensitive (or not) and whether subject thinks that buying climate neutral products is the right thing to do (as a proxy for normative motivations), respectively. The two *Hypothetical bias* variables are dummies indicating whether the respondent thinks it is likely or very likely that a) s/he exaggerated his/her valuation response b) others will exaggerate their responses. *SDS* is the social desirability scale variable.

| | | | | | | |
|---|---|---|---|---|---|---|
| | (12.587) | (5.539) | (7.416) | (12.342) | (5.375) | (7.379) |
| CV | 20.459 | 8.091 | 16.453 | 4.689 | 3.649 | 7.925 |
| | (54.335) | (16.966) | (27.211) | (50.124) | (13.032) | (24.382) |
| PCard | -89.575*** | -12.246 | -43.936*** | -90.184*** | -14.073* | -46.975*** |
| | (19.279) | (7.644) | (10.646) | (18.433) | (7.219) | (10.598) |
| WTA × CV | -0.201 | 8.664 | 5.897 | 0.358 | 8.115 | 5.329 |
| | (14.206) | (6.069) | (7.621) | (14.265) | (5.988) | (7.622) |
| CV × PCard | 31.257 | 16.832* | 24.420* | 27.227 | 16.240* | 27.051* |
| | (26.013) | (9.732) | (14.373) | (25.009) | (9.127) | (14.106) |
| WTA × PCard | -44.845*** | -36.096*** | -53.081*** | -43.403*** | -34.649*** | -52.405*** |
| | (13.885) | (6.065) | (7.877) | (13.799) | (5.904) | (7.845) |
| Product | | | 39.823*** | | | 45.723*** |
| | | | (5.757) | | | (5.793) |
| Shopper | | | | -5.313 | 4.476 | 3.106 |
| | | | | (13.676) | (5.787) | (8.414) |
| Age | | | | -0.257 | -0.476** | -0.572* |
| | | | | (0.665) | (0.232) | (0.336) |
| Gender | | | | -0.488 | 5.162 | 6.112 |
| | | | | (12.464) | (4.418) | (6.931) |
| Educ: junior hi-school | | | | 17.283 | 14.000 | 13.178 |
| | | | | (36.844) | (18.669) | (24.641) |
| Educ: senior hi-school | | | | -20.108 | 8.523 | -11.286 |
| | | | | (27.435) | (13.858) | (20.526) |
| Educ: some college or student | | | | -56.338* | 16.416 | -22.773 |
| | | | | (31.686) | (15.621) | (21.872) |
| Educ: university graduate or post-graduate studies | | | | -43.259 | 15.750 | -11.971 |
| | | | | (28.059) | (14.216) | (20.795) |
| Income: average | | | | 23.168 | -1.758 | 10.762 |
| | | | | (15.331) | (5.309) | (8.632) |
| Income: above average | | | | 42.133** | -4.517 | 11.911 |
| | | | | (18.739) | (6.017) | (10.294) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Income: good or very good | | | | 35.355 | 4.217 | 21.834 |
| | | | | (35.630) | (7.559) | (15.653) |
| Price sensitive | | | | -17.089 | -5.306 | -12.283** |
| | | | | (11.046) | (4.035) | (6.075) |
| Hypothetical bias likely | | | | -2.943 | 16.401*** | 8.106 |
| | | | | (22.093) | (5.369) | (9.753) |
| Hypothetical bias for others likely | | | | 3.581 | -6.294 | -1.468 |
| | | | | (11.773) | (3.928) | (6.255) |
| SDS | | | | -1.352 | -0.057 | -0.393 |
| | | | | (1.226) | (0.389) | (0.643) |
| Normative | | | | 26.122** | 6.514 | 19.459** |
| | | | | (13.173) | (6.146) | (8.427) |
| $ln(\sigma)$ | 4.503*** | 3.527*** | 4.232*** | 4.456*** | 3.462*** | 4.201*** |
| | (0.068) | (0.053) | (0.054) | (0.066) | (0.051) | (0.053) |
| $N$ | 512 | 512 | 1024 | 512 | 512 | 1024 |
| AIC | 1592.725 | 1687.152 | 3511.657 | 1593.821 | 1689.114 | 3506.733 |
| BIC | 1630.870 | 1725.297 | 3560.971 | 1729.447 | 1824.740 | 3669.472 |

Notes: Clustered standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table 3 shows Wald tests of linear combinations of coefficients for all interacted variables based on models (4) and (5) of Table 2. Figures 6, 7 and 8 graphically depict the effects of Table 3. The first four rows of Table 3 (associated effects are depicted in Figure 6) show the effect of the payment card format on elicited valuations. As evident, payment cards generate much lower valuations as compared to the discrete choice format. Most effects are highly significant as well.

The middle four rows of Table 3 (associated effects are depicted in Figure 7) show the effect of WTA as compared to WTP. Basically, the height of the bars in Figure 7 is evidence of a WTA/WTP gap. As shown, the gap is evident in most cases although Table 3 shows that it is significantly different from zero only when valuations are elicited using the DC format. For the payment card format, we fail to reject the null that the WTA/WTP gap is zero.

The last four rows of Table 3 (associated effects are depicted in Figure 8) show the effect

Table 3: Wald tests for conditional marginal effects from models (4) and (5) of Table 2

| | | Olive oil | | | Eggs | | |
|---|---|---|---|---|---|---|---|
| | | Coef. | $\chi^2$ | p-value | Coef. | $\chi^2$ | p-value |
| $\frac{\partial bid}{\partial PCard}$ | CV=1, WTA=1 | -106.359*** | 30.580 | 0.000 | -32.483*** | 22.460 | 0.000 |
| | CV=1, WTA=0 | -62.956*** | 9.830 | 0.002 | 2.166 | 0.100 | 0.748 |
| | CV=0, WTA=1 | -133.586*** | 55.240 | 0.000 | -48.722*** | 43.360 | 0.000 |
| | CV=0, WTA=0 | -90.184*** | 23.940 | 0.000 | -14.073* | 3.800 | 0.051 |
| $\frac{\partial bid}{\partial WTA}$ | CV=1, PCard=1 | 11.282 | 0.810 | 0.368 | 7.241 | 2.180 | 0.140 |
| | CV=1, PCard=0 | 54.684*** | 21.190 | 0.000 | 41.890*** | 55.610 | 0.000 |
| | CV=0, PCard=1 | 10.924 | 0.810 | 0.368 | -0.874 | 0.030 | 0.857 |
| | CV=0, PCard=0 | 54.327*** | 19.380 | 0.000 | 33.775*** | 39.490 | 0.000 |
| $\frac{\partial bid}{\partial CV}$ | WTA=1, PCard=1 | 32.274 | 0.440 | 0.509 | 28.004** | 5.830 | 0.016 |
| | WTA=1, PCard=0 | 5.046 | 0.010 | 0.920 | 11.764 | 0.730 | 0.393 |
| | WTA=0, PCard=1 | 31.916 | 0.420 | 0.515 | 19.889* | 3.510 | 0.061 |
| | WTA=0, PCard=0 | 4.689 | 0.010 | 0.926 | 3.649 | 0.080 | 0.780 |

Notes: * p<0.1, ** p<0.05, *** p<0.01

of the CV method as compared to the IV method. The positive coefficients indicate that in general IV elicits lower valuations than CV, albeit the effect is statistically different from zero only in the case of eggs where valuations are elicited using a payment card.

Overall, our econometric results are in accordance with the descriptive analysis presented above.

# 4 Conclusions

In this study we opted to test the robustness of the WTA/WTP gap under two basic experimental manipulations in a field valuation experiment of climate neutral foods. In one of our treatments we varied the valuation method between the classic Contingent valuation method and the more recent Inferred valuation method. Although we had good reasons to expect that a WTA/WTP gap could be muted with the inferred valuation we found that both the CV and IV methods generated a similar gap.

Our second treatment evaluated two different payment formats with respect to the WTA/WTP gap. We compared two polar opposite formats i.e., the Dichotomous choice format, which has been favored in the valuation literature for many years, with the Payment card format, which has made considerable headway over the last few years. We find that the PC format is able to mute the valuation gap between WTA and WTP, but not the DC format.

Our result that a payment card format can mute the WTA/WTP gap echoes the con-

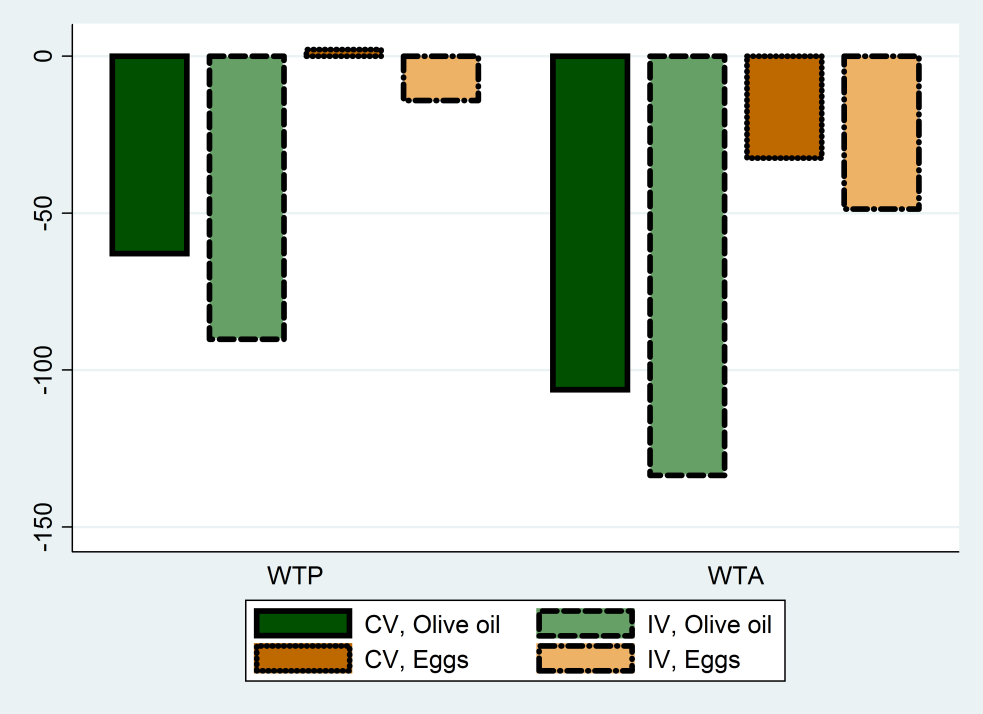


Figure 6: Effect sizes of Payment Card compared to Discrete Choice format

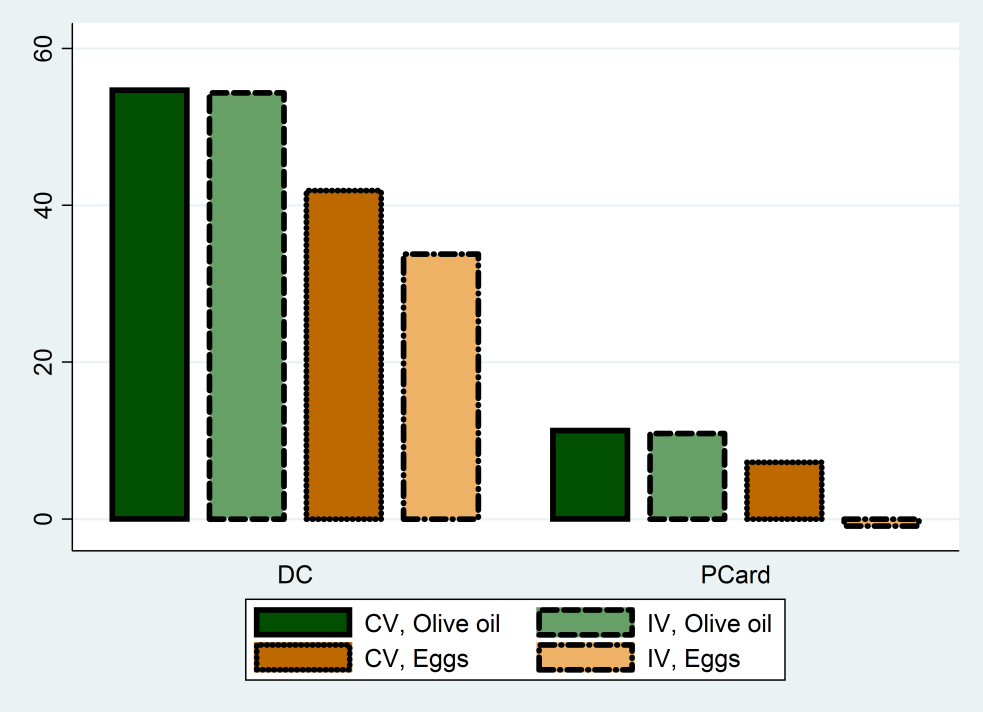


Figure 7: Effect sizes of Willingness-to-Pay compared to Willingness-to-Accept

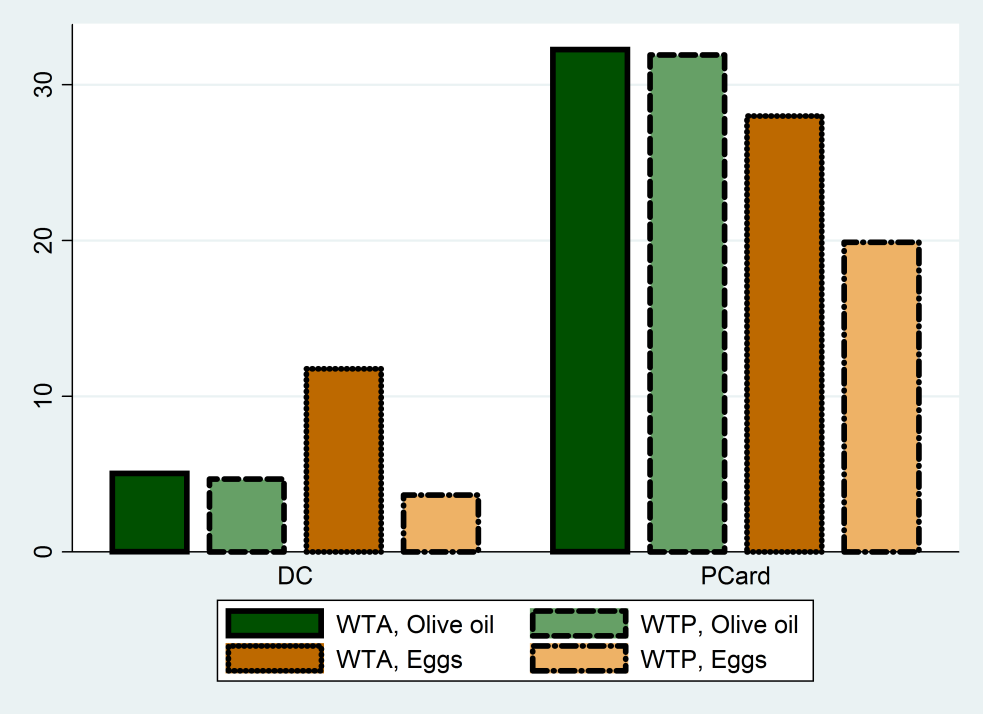


Figure 8: Effect sizes of Contingent Valuation compared to Inferred Valuation

clusion in Plott and Zeiler (2005) that observed gaps do not reflect a fundamental feature of human preferences given that the gap was not observed across all our experimental treatments. However, we do not have a definite answer as per what accounts for the gap across the different payment formats. Our study casts serious doubts in the general applicability of the DC format for valuation studies. More research studies that will try to explore the merits of the PC format are indeed warranted.

# References


Aizaki, H., Y. Nakashima, K. Ujiie, H. Takeshita, and K. Tahara (2011). Intention to access food risk information through internet-enabled mobile phones: The role of critical thinking. *Applied Economics Letters 18*(11), 1005–1009.

Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner, and H. Schuman (1993). Report of the noaa panel of contingent valuation. *Federal Register 58*(4601-4614).

Blaine, T. W., F. R. Lichtkoppler, K. R. Jones, and R. H. Zondag (2005). An assessment of household willingness to pay for curbside recycling: A comparison of payment card and referendum approaches. *Journal of Environmental Management 76*(1), 15–22.

Brouwer, R., P. van Beukering, and E. Sultanian (2008). The impact of the bird flu on public willingness to pay for the protection of migratory birds. *Ecological Economics 64*(3), 575–585.

Brummett, R. G., J. Rodolfo M. Nayga, and X. Wu (2007). On the use of cheap talk in new product valuation. *Economics Bulletin 2*(1), 1–9.

Buzby, J. C., J. A. Fox, R. C. Ready, and S. R. Crutchfield (1998). Measuring consumer benefits of food safety risk reductions. *Journal of Agricultural and Applied Economics 30*(1), 69–82.

Carson, R. T. (2012). Contingent valuation: A practical alternative when prices aren't available. *Journal of Economic Perspectives 26*(4), 27–42.

Carson, R. T., N. E. Flores, and N. F. Meade (2001). Contingent valuation: Controversies and evidence. *Environmental and Resource Economics 19*(2), 173–210.

Carson, R. T. and T. Groves (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics 37*(1), 181–210.

Carson, R. T., T. Groves, and M. J. Machina (1997). Stated preference questions: Context and optimal response. *Paper presented at the National science foundation preference elicitation symposium, University of California, Berkeley*.

Charness, G., U. Gneezy, and M. A. Kuhn (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization 81*(1), 1–8.

Corsi, A. (2007). Ambiguity of measured wtp for quality improvements when quantity is unconstrained: a note. *European Review of Agricultural Economics 34*(4), 501–515.

Donaldson, C., R. Thomas, and D. J. Torgerson (1997). Validity of open-ended and payment scale approaches to eliciting willingness to pay. *Applied Economics 29*(1), 79–84.

Dummett, M. and R. Farquharson (1961). Stability in voting. *Econometrica 29*(1), 33–43.

Frederick, S. (2012). Overestimating others willingness to pay. *Journal of Consumer Research 39*(1), 1–21.

Gibbard, A. (1973). Manipulation of voting schemes: A general result. *Econometrica 41*(4), 587–601.

Haab, T. C., M. G. Interis, D. R. Petrolia, and J. C. Whitehead (2013). From hopeless to curious? Thoughts on hausman's dubious to hopeless critique of contingent valuation. *Applied Economic Perspectives and Policy 35*(4), 593–612.

Hanemann, W. M. (1991). Willingness to pay and willingness to accept: How much can they differ? *American Economic Review 81*(June), 635–647.

Harrison, G. W., E. Johnson, M. M. McInnes, and E. E. Rutström (2005). Risk aversion and incentive effects: Comment. *The American Economic Review 95*(3), 897–901.

Horowitz, J. K. and K. E. McConnell (2002). A review of WTA/WTP studies. *Journal of Environmental Economics and Management 44*(3), 426–447.

Hsu, J. L., H.-Y. Chen, C. M.-L. Chen, and W.-H. Chang (2009). Sustainability of fresh fruit certifications: Willingness to pay measures. *Journal of Sustainable Agriculture 33*(6), 690–700.

Hu, W., T. A. Woods, S. Bastin, L. J. Cox, and W. You (2011). Assessing consumer willingness to pay for value-added blueberry products using a payment card survey. *Journal of Agricultural and Applied Economics 43*(2), 243–258.

Kling, C. L., D. J. Phaneuf, and J. Zhao (2012). From exxon to bp: Has some number become better than no number? *Journal of Economic Perspectives 26*(4), 3–26.

Kurt, D. and J. J. Inman (2013). Mispredicting others' valuations: Self-other difference in the context of endowment. *Journal of Consumer Research 40*(1), 78–89.

Lienhoop, N. and D. MacMillan (2007). Valuing wilderness in iceland: Estimation of wta and wtp using the market stall approach to contingent valuation. *Land Use Policy 24*(1), 289–295.

Loewenstein, G. and D. Adler (1995). A bias in the prediction of tastes. *The Economic Journal 105*(431), 929–937.

Lusk, J. and B. F. Norwood (2009a). Bridging the gap between laboratory experiments and naturally occurring markets:an inferred valuation method. *Journal of Environmental Economic and Management 58*(2), 236–250.

Lusk, J. L. and B. F. Norwood (2009b). An inferred valuation method. *Land Economics 85*(3), 500–514.

Mediamark Research and Intelligence (2009). Despite decades of gains in the workforce, women still the predominant household shoppers. http://www.gfkmri.com/PDF/MRIPR_111209_HouseholdShoppers.pdf. Last accessed on April 9, 2013.

Plott, C. R. and K. Zeiler (2005). The willingness to pay-willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. *The American Economic Review 95*(3), 530–545.

Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences 11*(1), 37–43.

Reaves, D. W., R. A. Kramer, and T. P. Holmes (1999). Does question format matter? valuing an endangered species. *Environmental and Resource Economics 14*(3), 365–383.

Satterthwaite, M. A. (1975). Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory 10*(2), 187–217.

Simpson, S. N. and B. G. Hanna (2009). Willingness to pay for a clear night sky: Use of the contingent valuation method. *Applied Economics Letters 17*(11), 1095–1103.

Smith, R. D. and T. H. Sach (2010). Contingent valuation: What needs to be done? *Health Economics, Policy and Law 5*(01), 91–111.

Solomon, B. D. and N. H. Johnson (2009). Valuing climate protection through willingness to pay for biomass ethanol. *Ecological Economics 68*(7), 2137–2144.

Stöber, J. (2001). The social desirability scale-17 (sds-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment 17*(3), 222–232.

van Boven, L., D. Dunning, and G. Loewenstein (2000). Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. *Journal of Personality and Social Psychology 79*(1), 66–76.

van Boven, L., G. Loewenstein, and D. Dunning (2003). Mispredicting the endowment effect:

Underestimation of owners selling prices by buyers agents. *Journal of Economic Behavior & Organization 51*(3), 351–365.

Zhongmin, X., J. Loomis, Z. Zhiqiang, and K. Hamamura (2006). Evaluating the performance of different willingness to pay question formats for valuing environmental restoration in rural china. *Environment and Development Economics 11*(5), 585–601.

# A Appendix: Valuation questions

[This is an English translation of the original text written in Greek]

Activities in the everyday life of modern people have an effect on the environment by contributing to the **climate change** phenomenon. More specifically, the average Greek causes approximately 12.7 tones of greenhouse gases per year. For this reason, some companies have certified their products as **climate neutral** in order to contribute in the fight against climate change. For a product to be certified as climate neutral one must first calculate its **carbon footprint** (that is, the quantity of $CO_2$ emissions produced during its production process) and then **offset the $CO_2$** emissions by funding carbon offset projects equivalent to the $CO_2$ emissions. $CO_2$ emissions are calculated in Greece by the *Center for Sustainability and Excellence* which is an international sustainability consulting, coaching, and training firm (with offices in Chicago, Athens and Brussels) and the offseting project is implemented by the Swiss non-profit organization *myclimate*.

1. **[Contingent valuation]**

   **[WTP elicitation]**

   Assume you have a free coupon which you can redeem for 1 liter of conventional extra virgin olive oil (a six-pack of conventional eggs) [experimenter shows picture C.1 or C.3 depending on the product treatment].

   [Payment card treatment]

   What is the maximum you would be willing to pay to exchange this coupon with a coupon which you can redeem for 1 liter of olive oil (a six-pack of eggs) like the one in the picture [experimenter shows picture C.2 or C.4 depending on the product treatment]? This olive oil (six-pack of eggs) is a climate neutral product and compensates for $CO_2$ emissions that are produced during the production process.

   [Discrete choice treatment]

   Would you be willing to pay €X to exchange this coupon with a coupon which you can redeem for 1 liter of olive oil (a six-pack of eggs) like the one in the picture [experimenter shows picture C.2 or C.4 depending on the product treatment]? This olive oil (six-pack of eggs) is a climate neutral product and compensates for $CO_2$ emissions that are produced during the production process.

   **[WTA elicitation]**

   Assume you have a free coupon which you can redeem for 1 liter of extra virgin olive

oil (a six-pack of eggs) [experimenter shows picture C.2 or C.4 depending on the product treatment]. This olive oil (six-pack of eggs) is a climate neutral product and compensates for $CO_2$ emissions that are produced during the production process.

[Payment card treatment]

What is the minimum you would be willing to accept to exchange this coupon with a coupon which you can redeem for 1 liter of conventional olive oil (a six-pack of conventional eggs) like the one in the picture [experimenter shows picture C.1 or C.3 depending on the product treatment]?

[Discrete choice treatment]

Would you be willing to accept €X to exchange this coupon with a coupon which you can redeem for 1 liter of conventional olive oil (a six-pack of conventional eggs) like the one in the picture [experimenter shows picture C.1 or C.3 depending on the product treatment]?

2. **[Inferred valuation]**

**[WTP elicitation]**

Assume that an average consumer has a free coupon which s/he can redeem for 1 liter of conventional extra virgin olive oil (a six-pack of conventional eggs) [experimenter shows picture C.1 or C.3 depending on the product treatment].

[Payment card treatment]

What is the maximum s/he would be willing to pay to exchange this coupon with a coupon which s/he can redeem for 1 liter of olive oil (a six-pack of eggs) like the one in the picture [experimenter shows picture C.2 or C.4 depending on the product treatment]? This olive oil (six-pack of eggs) is a climate neutral product and compensates for $CO_2$ emissions that are produced during the production process.

[Discrete choice treatment]

Would s/he be willing to pay €X to exchange this coupon with a coupon which s/he can redeem for 1 liter of olive oil (a six-pack of eggs) like the one in the picture [experimenter shows picture C.2 or C.4 depending on the product treatment]? This olive oil (six-pack of eggs) is a climate neutral product and compensates for $CO_2$ emissions that are produced during the production process.

**[WTA elicitation]**

Assume that an average consumer has a free coupon which s/he can redeem for 1 liter of extra virgin olive oil (a six-pack of eggs) [experimenter shows picture C.2 or C.4 depending on the product treatment]. This olive oil (six-pack of eggs) is a climate neutral product and compensates for $CO_2$ emissions that are produced during the production process.

[Payment card treatment]

What is the minimum s/he would be willing to accept to exchange this coupon with a coupon which s/he can redeem for 1 liter of conventional olive oil (a six-pack of conventional eggs) like the one in the picture [experimenter shows picture C.1 or C.3 depending on the product treatment]?

[Discrete choice treatment]

Would s/he be willing to accept €X to exchange this coupon with a coupon which s/he can redeem for 1 liter of conventional olive oil (a six-pack of conventional eggs) like the one in the picture [experimenter shows picture C.1 or C.3 depending on the product treatment]?

# B Appendix: Payment cards

Table B.1: Payment card for olive oil

| | | | |
|---|---|---|---|
| **1.** | 0 cents | **8.** | 1.51 - 1.75 € |
| **2.** | 1 - 25 cents | **9.** | 1.76 - 2.00 € |
| **3.** | 26 - 50 cents | **10.** | 2.01 - 2.25 € |
| **4.** | 51 - 75 cents | **11.** | 2.26 - 2.50 € |
| **5.** | 76 cents - 1.00 € | **12.** | 2.51 - 2.75 € |
| **6.** | 1.01 - 1.25 € | **13.** | >2.75 € |
| **7.** | 1.26 - 1.50 € | | |

# C Appendix: Pictures

Table B.2: Payment card for eggs

| | | | |
|---|---|---|---|
| **1.** | 0 cents | **8.** | 61 - 70 cents |
| **2.** | 1 - 10 cents | **9.** | 71 - 80 cents |
| **3.** | 11 - 20 cents | **10.** | 81 - 90 cents |
| **4.** | 21 - 30 cents | **11.** | 91 cents - 1.00 € |
| **5.** | 31 - 40 cents | **12.** | 1.01 €- 1.10 € |
| **6.** | 41 - 50 cents | **13.** | > 1.10 € |
| **7.** | 51 - 60 cents | | |


Figure C.1: Picture of 1 liter of conventional extra virgin olive oil


Figure C.2: Picture of 1 liter of climate neutral extra virgin olive oil


Figure C.3: Picture of a 6-pack of conventional eggs


Figure C.4: Picture of a 6-pack of climate neutral eggs