Does the elicitation method impact the WTA/WTP disparity?*

Sarah Brebner[a] and Joep Sonnemans[a,b,c]
[a] CREED, University of Amsterdam
[b] Tinbergen Institute
[c] Corresponding author, j.h.sonnemans@uva.nl

February 2017

Forthcoming Journal of Behavioral and Experimental Economics

Abstract

The size of the Willingness To Accept (WTA)/Willingness To Pay (WTP) disparity is compared using the Becker–DeGroot–Marschak (BDM) and multiple price list (MPL) methods. A robust WTA/WTP disparity is found using both elicitation methods. The MPL elicitation method appears to result in a slightly larger effect compared with the BDM method, contradicting claims that misconceptions specific to the BDM method are a driving force of the WTA/WTP disparity.

Keywords: WTA/WTP gap, endowment effect, elicitation method

*Financial support from the Research Priority Area Behavioural Economics of the University of Amsterdam is gratefully acknowledged. We thank Michalis Drouvelis and Lizzie Dipple, the editor and two anonymous reviewers for their comments. An earlier version of this paper was titled "Does the elicitation mechanism impact the endowment effect?"

1. Introduction

In this study, we compare the size of the Willingness To Accept (WTA)/Willingness To Pay (WTP) disparity using two different elicitation methods: open-ended Becker–DeGroot–Marschak (BDM) and multiple price list (MPL). Both methods, which are explained in detail below, are incentive compatible. The motivation behind this study is threefold.

First, Plott and Zeiler (2005, 2011) and Cason and Plott (2014) argue that the WTA/WTP disparity can be caused by design issues and especially subject misconceptions of the BDM method. Thus, we compare the size of the WTA/WTP gap when using the BDM and MPL methods.

Second, a comparison of the BDM and MPL methods is interesting because these methods represent different decision-making processes, namely judgment (BDM) and choice (MPL). In a judgment task the respondent has to report a value. Examples of judgment tasks include: scoring on a rating scale; bidding in a closed bid auction; and reporting the WTP or WTA in a BDM context. In a choice task, two or more alternatives are presented to the respondent who selects the option they prefer. An MPL method presents a series of such choices to the participant. A decision maker with well-defined preferences is supposed to make essentially the same decision in each response mode (procedure invariance). Preference reversals are well-known violations of procedure invariance (Grether and Plott 1979, Seidl 2002), and are typically explained as being caused by different psychological processes for judgment and choice. The two response modes may trigger different heuristics[1].

Third, a comparison of the BDM and MPL methods is of practical methodological interest. Both methods are used frequently, including in contexts other than the WTA/WTP disparity, and it is therefore important to learn whether these elicitation methods may yield different valuations.

The key findings of our experiment are that the WTA/WTP disparity is found in both treatments. Contrary to expectations, the use of the simpler MPL method does not result in a smaller effect compared with the BDM method,

[1] Note that some heuristics, like anchoring and adjustment (Tversky and Kahneman 1975), are only relevant in judgment tasks, while others, like Salience Theory (Bordalo et al 2012) and the focusing model of Kőszegi and Szeidl (2013), are only applicable to choice tasks.

contradicting the argument that misconceptions specific to the BDM method are a driving force of the WTA/WTP disparity. The next section discusses the major theories about the disparity; section 3 examines the two elicitation methods we compare in this study; section 4 describes the experimental design; section 5 presents the results and section 6 concludes.

2. Literature

For a general overview of the enormous literature in this field we refer to the following recent overview articles and meta-analyses. Ericson and Fuster (2014) discuss the experimental evidence and available theories, their conclusion is in favor of reference point explanations; Zeiler (forthcoming) discusses the evidence for and against the different theories and concludes that at the moment "no one theory clearly rises to the top as the leading explanation"; Tunçel and Hammitt (2014) present a meta-analysis of experiments and surveys with different goods and Yechiam et al (2017) undertook a meta-analysis of experiments that use lotteries as goods. There is also a large literature on preference measurements and valuation methods in the field of environment and resource economics; for a recent overview see the literature section in Price et al (2017)[2].

Many different explanations for the WTA/WTP gap are proposed; we will discuss some of them below but, because of space limitations, we cannot be exhaustive, and we refer the interested reader to the overview articles mentioned above. The reader should keep in mind that human behaviour can have a multitude of causes and it is therefore unlikely that any one of the theories will be able to explain all phenomena in all settings.

Reference dependent valuation and loss aversion. The original explanation for the WTA/WTP disparity (e.g. Thaler 1980, Kahnemann et al. 1990) is that buyers and sellers value the good by comparing it with a reference point. This reference point can be the status quo (for the seller (buyer) (not) owning the

[2] The focus of the methodological studies in this field are typically about the test-retest reliability of different measurements. We know only one study (Loomis 1990) that compares open ended contingent valuations with a multiple price list. Loomis (1990) uses both methods in an unincentivized survey, but does not report a direct comparison between the measured valuations. He finds that the test-retest reliability is about the same in both methods.

object) or it can be based upon expectations (e.g. Kőszegi and Rabin 2006). According to prospect theory there is a kink in the value function v at the reference point: losses loom larger than gains. $v(-x)=-\lambda v(x)$ if $x>0$ with $\lambda>1$ the loss aversion parameter. A seller who trades loses the object, but gains money, while the buyer gains the object but loses money (see Ericson and Fuster 2014 for a formal description). This explanation predicts a larger WTA/WTP gap for decision makers who are more loss-averse in choices under uncertainty[3].

Bad deal aversion. The decision maker may compare the price with some reference price and buying (selling) for a higher (lower) price than the reference price is considered to be a "bad deal". The reference price will typically not be a precise amount but an "imprecision interval" (Butler and Loomes 2007). To avoid bad deals the reported WTA and WTP values will be "on the safe side": high in the interval for a WTA and low for a WTP, and the size of the WTA/WTP gap will thus depend on the size of the imprecision interval. Okada (2010) shows that the WTA/WTP gap indeed increases with reference price uncertainty.

Direct effects of ownership. Ownership may influence the valuation of the good directly, even before the possibility of trade (and thus before loss aversion can play a role). Decision makers may form different beliefs about the good depending on ownership (e.g. Plott and Zeiler 2011). For example, an optimist who owns a lottery ticket will think it is more likely that his ticket will win than a ticket he does not own, which leads to a higher subjective expected value of a lottery ticket that is owned. Drouvelis and Sonnemans (2017) study this hypothesis in the case of strategic uncertainty. The endowed good in their experiment is the right to play a two-person game and they measure beliefs about the action of the other player. They find a positive correlation between optimism and the size of the WTA/WTP disparity. However, optimism cannot completely explain the gap, because a significant disparity is also found when the beliefs were the same in the WTA and WTP cases. Another direct effect of ownership may be emotional. Georgantzis and Navarro-Martínez (2010) argue that participants who receive the object (a bottle of wine in their experiment) may experience positive feelings, which in turn positively influence their

[3] The only study we know that tests this hypothesis is a field study by Gächter et al (2010). They find a positive correlation between the WTA/WTP gap and loss aversion in a risky choice task.

valuation, even before they are asked for a WTA price (and thus independent of loss aversion). They find that a significant part of the WTA–WTP gap is explained by participants' positive feelings produced by receiving and owning the target good.

Buy-low, sell-high heuristic. An economic experiment will be an unfamiliar setting for many participants and they bring into the lab the heuristics that work quite well in real life outside the lab. One such heuristic is to start negotiations as a buyer (seller) with a low (high) bid which may lead to a low WTP and high WTA valuation.

Misconceptions. In a series of influential studies, Plott and Zeiler (2005, 2011) argue that the WTA/WTP disparity is not a fundamental feature of human preferences, but is caused by design issues and especially subject misconceptions of the BDM method. Decision makers who do not understand the BDM method are more likely to use a buy-low, sell-high heuristic. Plott and Zeiler's work has led to several follow-up studies. Isoni et al (2011) are able to replicate the Plott and Zeiler (2005) findings for consumption goods (mugs) but not lotteries, concluding that differences in goods (money vs consumption goods) is most likely driving the difference between their results. Fehr et al (2015) are not able to replicate the Plott and Zeiler (2005) study.

Cason and Plott (2014) study mistakes and misunderstandings of the BDM method. In a class room setting students are asked to state a minimum selling price for a card that can be redeemed from the experimenter for $2. A posted price is drawn from a known interval and the students sell the card and receive the posted price if this posted price is higher than their minimum selling price. The authors find that a large majority of students report a WTA price larger than $2. After the students learn the outcome from the initial task, the task is repeated, and more valuations are at $2, but the majority still reports WTA prices that are too high. At least some of the students seem to think that, when they sell the card, they will get the minimum selling price instead of the posted price. Bartling et al (2015) find a statistically significant WTA/WTP disparity even for participants who demonstrate an understanding of the BDM method, contradicting claims that the gap results only from participants' misconceptions of the elicitation method.

3. Elicitation methods

We will now introduce the two elicitation methods we use in this study and discuss possible advantages and disadvantages.

The BDM method is an incentive-compatible value elicitation method developed by Becker–DeGroot–Marschak (1964). Participants are asked to provide an offer for the good being valued. This is compared to a randomly drawn fixed price, which is used as the trading price. For example, if participants are making an offer to buy, they purchase the item only if their offer was higher than the fixed price. However, the price they pay is the randomly drawn fixed price, not their offer. A participant's dominant strategy is to offer exactly their value.

Disadvantages of the BDM method are that it is quite abstract and participants may misunderstand. With the BDM, two prices are mentioned: the *offer price* of the buyer or seller, and a randomly drawn *fixed price*. Any trade that takes place is for the *fixed price.* If the participant confuses these two prices and thinks that the trade will be at the *offer price*, as a buyer (seller) he will post an offer price that is lower (higher) for his value. As Cason and Plott (2014) note, this misconception is analogous to confusing a second-price auction with a first-price auction and this failure of game form recognition could be a cause of the WTA/WTP disparity.

Advantages of the BDM are, first, that (after a lengthy introduction and explanation) a single measurement costs little effort and time because the participant has only to report one number, and second, the valuation can be very precise (e.g. up to single cents).

In the MPL method participants are asked to make multiple independent decisions. For each decision, they are given a choice between a fixed amount and the good being valued[4]. The fixed amount varies for each decision and thus a valuation can be obtained from the point where the participant changes, or switches, to the other option. Decisions are incentivised by randomly selecting one decision for payment (only one decision is paid to avoid income effects).

[4] The MPL method is also commonly used in risk elicitation where participants are generally given a choice between two lotteries.

The main advantage of MPL are transparency and simplicity. There can be no misunderstanding over payment as only one "price" is mentioned (for the randomly selected decision that will be paid). The decision maker simply gets what she has chosen: either the lottery or the money. Of course, other kinds of confusion are in principle possible, but there is no obvious misunderstanding or mistake that would lead to a systematic WTA/WTP gap.

The disadvantages of the MPL method are that it requires much more time and effort from participants (for each measurement, multiple decisions must be made) and the method elicits only a value range rather than a point estimate (Anderson et al 2007). However, as noted by Anderson et al (2007), given controversies over the ability to elicit precise valuations, an interval response may be more appropriate (p.676). An MPL valuation fails if the participant switches more than once[5]. This can be caused by a mistake, a misunderstanding, or a participant who doesn't take the task seriously.

To the best of our knowledge, there is only one direct comparison of the MPL and BDM elicitation methods in relation to the WTA/WTP disparity. This is undertaken by Kahneman, Knetsch and Thaler (1990). They find that the WTA/WTP disparity is robust to elicitation methods using consumption goods. The methods they employ are MPL and a simple open-ended question asking for a valuation. In another experiment reported in the same paper, they also employ a version of BDM to address concerns regarding incentive incompatibility in the other experiments. They report similar findings across each of the experiments.

The experiment in this paper differs from that of Kahneman et al (1990) in several key ways. First, lotteries, not consumption goods, are used as the underlying good. Second, a test of the size of the treatment effect is included, not only a test of robustness of the WTA/WTP disparity. The existence of a treatment effect stemming from the elicitation method can provide evidence regarding the cause(s) of the WTA/WTP disparity. Third, the instructions are based on the Isoni et al (2011) instructions (which are based on the Plott and Zeiler (2005) instructions) to avoid subject misconceptions of the elicitation methods. Finally,

[5] Bruner (2011) found that the inclusion of instructions emphasizing the incentive compatibility of the payment rule reduced observed multiple switching behavior from 13.3% to 2.3% in probability-varied MPL and from 25.8% to 6.7% in reward-varied MPL (p.417)

only incentive-compatible elicitation methods are used and only in a controlled laboratory situation, while some of the experiments in Kahneman et al (1990) are classroom experiments without incentives. Remarkably, Kahneman et al (1990) find a WTA/WTP disparity of an enormous size (WTA two to three times as large as WTP) while more recent experiments using lotteries find a more modest disparity of typically between 10% and 50%.

4. Design

We employ a between-subjects comparison of the two treatments: BDM and MPL elicitation methods. Participants are randomly allocated to a treatment.
The experiment is computerized and programmed in PHP, MySQL and JavaScript. To ensure that sessions are comparable, all instructions to participants (other than verbal instructions regarding entry and seating procedures) are included in the computer program. The instructions can be found in the downloadable Appendix. Instructions for the BDM treatment largely replicate the Isoni et al (2011) instructions which are, in turn, based on the Plott and Zeiler (2005) instructions with modifications to accommodate a computerized instead of a paper-and-pen implementation. The MPL instructions are formulated to align with the BDM instructions as much as possible to ensure differences in the treatments are based only on the differences in the elicitation methods themselves.

Participants are asked for their valuations of each lottery and are instructed that there are both buying and selling tasks. In the buying task, they do not own the lottery and a WTP is elicited using the allocated method. In the selling task, they are informed that they do own the lottery and a WTA is elicited using the allocated method. For each treatment, both WTP and WTA is elicited for each item from each participant. This enabled a within-subject measurement of WTA/WTP.

The goods selected to be valued are lotteries as shown in Table 1. Only lotteries with strictly positive outcomes are included, to ensure that lotteries have similar complexity. To enable a within-subject measurement of WTA/WTP whilst avoiding participants being asked to value exactly the same lottery twice, pairs of lotteries are used. Each pair consists of a WTP lottery (valued in the

buying task) and a WTA lottery (valued in the selling task), with the outcomes in the WTP lottery differing from each of the outcomes in the corresponding WTA lottery by €1. In half the lotteries, each WTP outcome is €1 higher than the corresponding WTA outcome; in the other lotteries WTP outcomes are €1 lower than the WTA outcomes (the adjustment for each pair is also indicated in Table 1). Assuming constant absolute risk aversion, expected utility theory implies that for each lottery pair WTP – WTA = €1 where the WTP outcomes are higher than the WTA outcomes or WTP – WTA = – €1 where the WTP outcomes are lower (Fehr et al 2015 p. 122). This approach has previously been used, including by Isoni et al (2011), Fehr et al (2015) and Drouvelis and Sonnemans (2017).

Table 1: Lotteries used for value elicitation[6]

Lottery Number	WTA Lottery	Lottery Number	WTP Lottery	Relationship between WTP and WTA
1	(€2, 0.6; €4, 0.4)	12	(€1, 0.6; €3, 0.4)	WTP = WTA - 1
3	(€2.5, 0.8; €7, 0.2)	14	(€1.5, 0.8; €6, 0.2)	WTP = WTA - 1
5	(€2, 0.6; €6.5, 0.4)	16	(€3, 0.6; €7.5, 0.4)	WTP = WTA + 1
7	(€7, 0.5; €2, 0.5)	18	(€6, 0.5; €1, 0.5)	WTP = WTA - 1
9	(€1.5, 0.7; €3.5, 0.3)	20	(€2.5, 0.7; €4.5, 0.3)	WTP = WTA + 1
11	(€5, 0.7; €2, 0.3)	2	(€4, 0.7; €1, 0.3)	WTP = WTA - 1
13	(€2.5, 0.5; €0.5, 0.5)	4	(€3.5, 0.5; €1.5, 0.5)	WTP = WTA + 1
15	(€2, 0.7; €6, 0.3)	6	(€1, 0.7; €5, 0.3)	WTP = WTA - 1
17	(€5, 0.8; €2.5, 0.2)	10	(€6, 0.8; €3.5, 0.2)	WTP = WTA + 1
19	(€5, 0.6; €1, 0.4)	8	(€6, 0.6; €2, 0.4)	WTP = WTA + 1

Notes: The lottery number indicates the order that lotteries are displayed for valuation during the experiment. WTP and WTA lottery valuations are interleaved with corresponding WTA and WTP lottery valuations separated by a number of rounds. The notation (€2, 0.6; €4, 0.4) indicates a lottery with a 60% chance of winning €2 and a 40% chance of winning €4.

Similar to Drouvelis and Sonnemans (2017), WTP and WTA elicitation rounds are interleaved for a given treatment[7]. The order in which lotteries are shown ensured that there are at least six rounds between WTA and WTP elicitation for related lotteries. This is done to reduce the likelihood that participants would recognise similarities between related lotteries, thereby impacting their

[6] Lotteries 2, 4, 6 and 12 are used in the experiment by Isoni et al (2011), albeit they use pounds rather than euros (their large-scale lotteries 11, 13, 14 and 12, respectively). However, in our lotteries 4, 6 and 12 we add (instead of subtract) 1 euro for the related WTA lotteries to avoid 0 euro outcomes. For the other six lottery pairs, we use outcomes between but excluding 0 and 8 euros and probabilities 0.3–0.7, 0.4–0.6 or 0.5–0.5.

[7] To avoid confusion and assist participants in differentiating the tasks, different colours are used to indicate the different tasks in the same way as in Drouvelis and Sonnemans (2017). See the downloadable appendix for example screen shots.

valuations/choices. Lotteries are presented in the same order to all participants across both treatments.

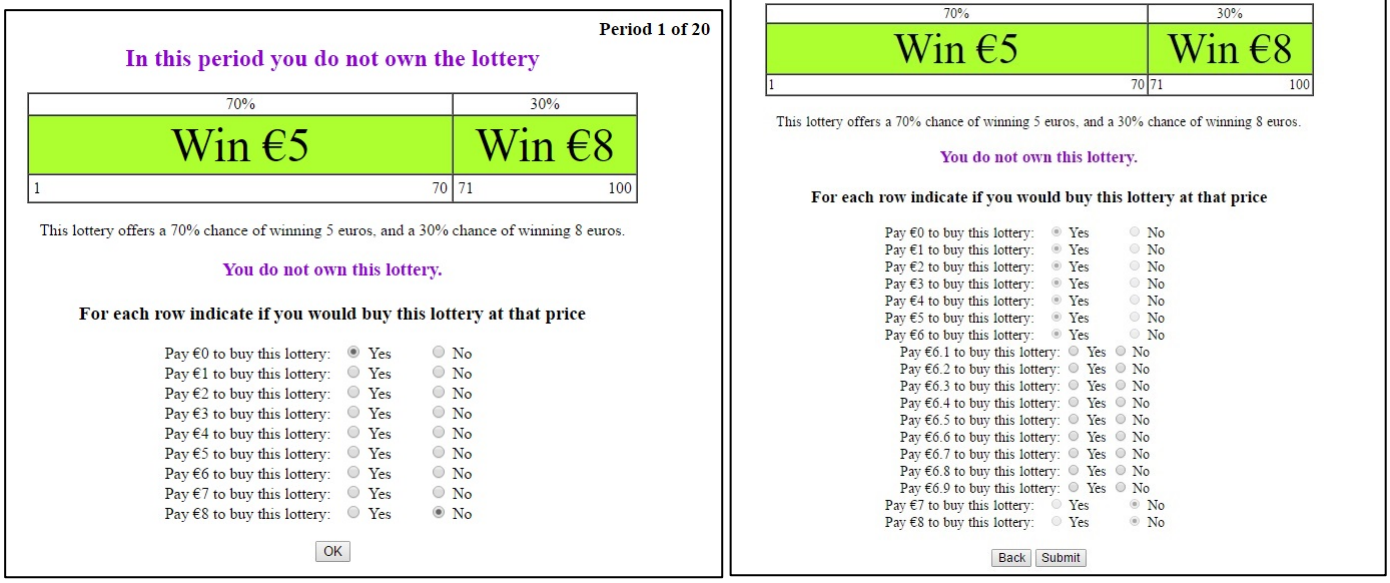
In the BDM treatment, participants are asked to enter either a WTP or a WTA into a text box. Offers are limited to the range €0 to €8 inclusive, with participants being able to specify prices to the nearest cent. To determine the outcome (in the event this decision was selected for payment), offers are compared to a fixed offer drawn from a uniform distribution over the range €0 to €8 (specified to the nearest cent). For WTP, participants who indicate an offer equal to or higher than the fixed offer buy the item at the fixed price. For WTA, participants who indicate an offer equal to or lower than the fixed offer sell the item and receive the fixed price. Fixed offers and the results for each participant are announced only after the conclusion of the experiment, to control for learning.

In the MPL treatment, participants are presented with a list of choices between the lottery and a fixed value (see Figure 1). For each price, the participant has to choose whether they would buy/sell the lottery at the stated price. Choices are elicited using two-stage MPL. In the first stage, choices are made between €0 and €8 in €1 increments. This range is selected to ensure consistency with the BDM elicitation method. The second stage choices depend on the participant's selection from the first stage and are used to increase precision without overly burdening participants. The second stage choices are 10¢ increments displayed between the prices where the user switched from "Yes" (i.e. willing to buy/sell) to "No" (i.e. not willing to buy/sell). See Figure 1 for an example of the MPL screens and a basic explanation of the functionality. To further reduce the burden on participants in the MPL treatment, the first and last options are pre-selected as defaults (although participants can change these selections). This is explained in the instructions to participants. The computer program does not allow multiple switching, instead raising an error message and indicating to the user that their choices are not consistent and to ask for assistance if they do not understand. Only one user requested such an explanation.

The MPL elicitation method results in an implied range for the participants' valuation. The mid-point of this range is used as the valuation in the

subsequent analysis. The elicitation range of €0 to €8 is chosen so that participants valuations should fall within this range (all lottery outcomes are in the range €0.50 to €7.50) rather than on or above a boundary, which would lead to an unbounded valuation range.[8]

Figure 1: example screenshot for the initial screen (left) and the second stage decision (right) for the MPL elicitation method for an example WTP lottery



Notes:
The first and last decisions have a default selection (which the user can change if desired). After making a choice for all decisions (i.e. every row), the "OK" button can be pressed to display the second stage decisions.
The second stage decisions depend on the selection in the first stage. In the example above, the participant switched between €6 and €7 so the second stage shows values between these two amounts.

The experiment was undertaken at the Center for Research in Experimental Economics and political Decision-making (CREED) laboratory in Amsterdam with 92 participants (37 female (40%), 55 male (60%), all participants were students: economics or business 49%; other social sciences 21%) who earned an average of €12.40 in approximately 1 hour. One participant was excluded from all analyses because she was aware of the topic and key research question of the experiment (she was an intern at CREED). A further participant was excluded from all analysis as it

[8] This appeared to be successful as no participants indicated either that they would sell/buy at any price or that they would not sell/buy at any price (i.e. all participants had a switching point within the elicitation range). However, the preselection of default answers at either end of the range was also strongly suggestive against this behaviour.

did not appear that she correctly understood the task[9]. The BDM treatment group consisted of 44 participants (19 female and 25 male) whilst the MPL treatment group consisted of 46 participants (16 female and 30 male).

5. Results

Table 2 shows the main results. To compare the WTA and WTP valuations, we correct the WTP reports in line with the last column of table 1. In the BDM treatment group, we find significantly higher WTA than WTP in 7 of the 10 lotteries (two-sided Wilcoxon $p<0.05$) and no effect in the other three lotteries. For the MPL treatment we find significantly higher WTA than WTP in all 10 lotteries.

For all lotteries, the mean WTA (WTP) is somewhat lower (higher) in the BDM treatment than in the MPL treatment. Thus, in all lotteries the average gap between WTA and WTP is larger when using MPL. However, this difference is not significant for 6 of the 10 lottery pairs, marginally significant ($p<0.10$) for two lotteries and significant for two lotteries ($p<0.05$). Over all the lotteries we do not find a difference in the gap between the two elicitation methods using a non-parametric (Mann–Whitney) test.

The valuation of a lottery, averaging over the WTA and WTP measurements, does not differ between the two elicitation methods. Five lotteries are valued on average a few cents more in the BDM and for the other five lotteries the difference is the other way around; over all the lotteries the difference in average valuation is less than €0.0007.

Another way to analyse the data is to count how often the WTA/WTP disparity occurs in each treatment. Note, however, that due to the precision of the MPL elicitation method (to the nearest 10¢), we have to use ranges for this comparison. Thus, a subject was classified as displaying a WTA/WTP disparity for a particular lottery if, for that lottery WTA > WTP + 0.10. Cases with WTA < WTP - 0.10 are coded as an anti- WTA/WTP disparity, and all cases where WTA differed from WTP by 10 cents or less are coded as neutral. This coding was adopted for both treatments to make the comparison fair. Table 3 shows that the WTA/WTP disparity

[9] She reported only extremely low WTP values of €0.05 (8 times) and €0.95 (2 times) noting that all lotteries have a minimum payoff of 1 euro or more (see table 1). She was in the MPL treatment.

is more often observed in the MPL treatment than in the BDM treatment[10], but the difference is not statistically significant.

Finally, a remark about the robustness of the WTA/WTP gap when valuing lotteries. Kahneman et al (1990, page 1328) expects an endowment effect only for nonmonetary items that can be utilized by the sellers. Lotteries with money prizes have a well-defined expected value and risk, so the WTA and the WTP should only depend on the individual's risk-attitude. Yechiam et al (2017) report a meta-analysis of studies of the WTA/WTP disparity with lotteries. These studies differ in many aspects: within or between subjects, different elicitation methods, incentivized or not. Of the 20 incentivized experiments, 19 show larger WTA than WTP prices, and this difference is statistically significant in 11 studies. With our study, we add another two confirming observations to the literature.

Table 2: Mean valuations of the lotteries using the BDM and MPL methods

Lottery pair	BDM (N=44)			MPL (N=46)			p-value comparison effect in BDM and MPL[b]
	WTA Mean (SD)	WTP Mean (SD)	p-value[a]	WTA Mean (SD)	WTP Mean (SD)	p-value[a]	
1 & 12	2.87 (0.52)	2.55 (0.33)	0.001	2.94 (0.48)	2.44 (0.41)	0.000	0.247
3 & 14	3.66 (1.00)	3.19 (0.55)	0.012	3.80 (0.93)	3.13 (0.67)	0.000	0.166
5 & 16	3.92 (0.78)	3.42 (0.60)	0.000	4.22 (0.94)	3.15 (0.82)	0.000	0.215
7 & 18	4.26 (0.80)	4.01 (0.85)	0.250	4.63 (0.96)	3.72 (0.97)	0.001	0.086
9 & 20	2.04 (0.29)	1.98 (0.23)	0.594	2.28 (0.52)	1.82 (0.47)	0.001	0.021
11 & 2	3.76 (0.74)	3.74 (0.64)	0.647	4.00 (0.69)	3.46 (0.66)	0.001	0.016
13 & 4	1.46 (0.45)	1.23 (0.38)	0.017	1.59 (0.52)	1.05 (0.52)	0.001	0.194
15 & 6	3.19 (0.65)	2.90 (0.55)	0.011	3.55 (1.02)	2.70 (0.61)	0.000	0.098
17 & 10	4.19 (0.59)	3.82 (0.64)	0.000	4.28 (0.58)	3.71 (0.81)	0.002	0.981
19 & 8	3.33 (0.71)	2.78 (0.73)	0.000	3.43 (0.79)	2.40 (1.09)	0.000	0.633

Notes: a: two-sided Wilcoxon signed-rank; b: two-sided Mann-Whitney comparing WTA-WTP in each treatment. If we test instead on WTA/WTP the p-values are very similar and give the same results at standard significance levels.

Table 3: Occurrence of the WTA/WTP disparity

[10] The numbers in table 3 can be compared with the data of Isoni et al (2011). Using the same classification criteria, we find 47.0% WTA/WTP disparity, 34.7% neutral and 18.3% anti-WTA/WTP disparity in their data.

	BDM (N=44)	MPL (N=46)	p-value
WTA/WTP disparity	42.7%	53.9%	0.111
Neutral	41.4%	34.6%	0.145
Anti - WTA/WTP disparity	15.9%	11.5%	0.078

Notes: A pair of valuations is coded as a WTA/WTP disparity if WTA – WTP>€0.1, as an anti-WTA/WTP disparity if WTA – WTP<-€0.1 and as neutral in all other cases. Treatment effects are tested using a Mann-Whitney test with observations on the level of individuals. For each individual, we count the number of cases where a WTA/WTP disparity is observed, the number of neutral cases and the number of anti-WTA/WTP cases.

6. Conclusion

This study is inspired by the suggestions in recent literature that the Willingness To Accept (WTA)/ Willingness To Pay (WTP) disparity may be attributed to confusion about the Becker–DeGroot–Marschak (BDM) method (Plott and Zeiler 2005, 2011). We measure the WTA/WTP disparity using the BDM and Multiple Price List (MPL) methods. With both elicitation methods, we find a WTA/WTP disparity, and we find that this effect is somewhat stronger in the MPL treatment. We conclude that the WTA/WTP disparity is not solely caused by characteristics or misconceptions specific to the BDM method.

The two elicitation methods produce approximately the same valuations (means and standard deviations), which is good news for experimentalists. The main advantage of MPL over BDM is the ease of explanation. The main disadvantage of MPL is that it takes more time and effort from the respondent: in our setup, each valuation requires participants to click 16 times, while the BDM method only asks participants to type a single number. The number of mouse clicks needed depends on the precision desired. When only few valuations with a relatively low resolution are needed, MPL seems to be a practical choice, but when participants have to make many valuations, using BDM (after a thorough explanation) makes more sense.

References

Anderson, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2007). Valuation using multiple price list formats. *Applied Economics*, *39*(6), 675-682.

Bartling, B., Engl, F., & Weber, R. A. (2015). Game form misconceptions are not necessary for a willingness-to-pay vs. willingness-to-accept gap. *Journal of the Economic Science Association*, 2015, Vol.1(1), 72-85.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, *9*(3), 226-232.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience Theory of Choice Under Risk. *The Quarterly journal of economics*, *127*(3), 1243-1285.

Bruner, D. M. (2011). Multiple switching behaviour in multiple price lists. *Applied Economics Letters*, *18*(5), 417-420.

Butler, D. J., & Loomes, G. C. (2007). Imprecision as an account of the preference reversal phenomenon. *The American Economic Review*, *97*(1), 277-297.

Cason, T. N., & Plott, C. R. (2014). Misconceptions and game form recognition: Challenges to theories of revealed preference and framing. *Journal of Political Economy*, *122*(6), 1235-1270.

Drouvelis, M., & Sonnemans, J. (2017). The Endowment Effect in Games. *European Economic Review* 94, 240-262.

Ericson, K. M., & Fuster, A. (2014). The Endowment Effect. *Annual Review of Economics*, 2014, Vol.61(1), 555-579.

Fehr, D., Hakimov, R., & Kübler, D. (2015). The willingness to pay–willingness to accept gap: A failed Replication of Plott and Zeiler, *European Economic Review*, 78 (2015) 120-128.

Gächter, S., Johnson, E. J., & Herrmann, A. (2010). *Individual-level loss aversion in riskless and risky choices* (No. 2010-20). CeDEx discussion paper series.

Ganzach, Y. (1995). Attribute scatter and decision outcome: Judgment versus choice. *Organizational Behavior and Human Decision Processes*, *62*(1), 113-122.

Georgantzís, N., & Navarro-Martínez, D. (2010). Understanding the WTA–WTP gap: attitudes, feelings, uncertainty and personality. *Journal of Economic Psychology*, *31*(6), 895-907.

Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, *69*(4), 623-638.

Isoni, A., Loomes, G., & Sugden, R. (2011). The Willingness to Pay-Willingness to Accept Gap, the Endowment Effect, Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Comment. *American Economic Review*, 2011, Vol.101(2), 991-101.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of political Economy*, *98*(6), 1325-1348.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133-1165.

Kőszegi, B., & Szeidl, A. (2013). A model of focusing in economic choice. *The Quarterly Journal of Economics*, *128*(1), 53-104.

Loomis, J. B. (1990). Comparative reliability of the dichotomous choice and open-ended contingent valuation techniques. *Journal of environmental economics and management*, *18*(1), 78-85.

Okada, E. M. (2010). Uncertainty, risk aversion, and WTA vs. WTP. *Marketing Science*, *29*(1), 75-84.

Price, J., Dupont, D., & Adamowicz, W. (2017). As Time Goes By: Examination of Temporal Stability Across Stated Preference Question Formats. *Environmental and Resource Economics*, *68*, 643–662.

Plott, C. R., & Zeiler, K. (2005). The Willingness to Pay-Willingness to Accept Gap, the Endowment Effect, Subject Misconceptions, and Experimental Procedures for Eliciting Valuations. *American Economic Review*, 2005, Vol.95(3), 530-545.

Plott, C. R., & Zeiler, K. (2011). The Willingness to Pay-Willingness to Accept Gap, the Endowment Effect, Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Reply. *American Economic Review*, 2011, Vol.101(2), 1012-1028.

Seidl, C. (2002). Preference Reversal. *Journal of Economic Surveys*, Vol.16(5), 621-655.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39-60.

Tunçel, T., & Hammitt, J. K. (2014). A new meta-analysis on the WTP/WTA disparity. *Journal of Environmental Economics and Management*, *68*(1), 175-187.

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141-162). Springer Netherlands.

Yechiam, E., Ashby, N. J., & Pachur, T. (2017). Who's Biased? A Meta-Analysis of Buyer-Seller Differences in the Pricing of Lotteries. *Psychological bulletin* [Epub ahead of print] doi: 10.1037/bul0000095.

Zeiler, K. (forthcoming) What explains observed reluctance to trade? a comprehensive literature review *Research Handbook on Behavioral Law and Economics* Retrieved from: http://www.bu.edu/law/working-papers/what-explains-observed-reluctance-to-trade-a-comprehensive-literature-review/