Validating consistency of non-hypothetical experimental auction data: application of the adding-up test in a multi-unit setting.

Levan Elbakidze
Associate Professor
University of Idaho
lelbakidze@uidaho.edu

Rodolfo Nayga
Professor
University of Arkansas
rnayga@uark.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2015 AAEA Annual Meeting, San Francisco, CA, 2015.

Validating consistency of non-hypothetical experimental auction data: application of the adding-up test in a multi-unit setting.

1. Introduction

Stated preference based valuation methods have been widely used to estimate the values that communities attach to non-market goods and services like environmental quality, food safety, animal welfare, scenic beauty, etc. However, the debates about the merits and validity of contingent valuation and other stated preference based valuation techniques continue. For example, volume 26(4) of the 2012 *Journal of Economic Perspectives* contained several articles by well-respected experts in the field debating the validity of contingent valuation surveys (Hausman, 2012; Carson, 2012; Kling, Phaneuf, and Zhao, 2012). In his critical assessment of contingent valuation techniques, Hausman (2012) presented a case where respondents in a contingent valuation survey did not appear to act as rational agents, but rather displayed preferences which appeared not to be consistent with basic economic principles and assumptions needed for unbiased estimation of nonmarket values. Specifically, the data collected from a contingent valuation survey did not pass the "adding-up test" (Diamond, 1996; Diamond and Hausman, 1994). This outcome was interpreted as a sign of internal inconsistency of the estimates, and/or estimation methods, representing a departure from basic assumptions of microeconomic theory (Diamond, 1996). Such violations of economic theory significantly undermine the credibility and reliability of economic valuation studies and subsequent Benefit Cost Analysis (BCA). The violation detected in Hausman (2012) was judged as especially problematic because the case study was done in the context of a familiar and well defined good (i.e., cable TV). He argued that this problem can be worse for less defined goods and services like biodiversity, habitat sustainability, scenic beauty, health aspects of new food technologies, etc. The hypothetical nature of stated preference based studies was seen as the primary cause for the observed inconsistencies.

Due to the limitations and observed inconsistencies seen from findings of studies using stated preference methods, applied economists are increasingly using non-hypothetical valuation methods, especially for marketing and behavioral economics research, due to their incentive compatibility property vis-à-vis stated valuation methods (Gracia, Loureiro, and Nayga 2011a). One of the most common of these non-hypothetical valuation methods in the literature today is experimental auction. Researchers have used experimental auctions to estimate consumers' willingness to pay (WTP) for new products and product attributes using real products and real money to create an experimental setting where participants have a greater incentive to reveal their true values for the products being examined (Gracia et al. 2011b; Elbakidze et al. 2013; Rousu 2004; Corrigan and Rousu 2006). The reason for growth in their popularity is that these methods are expected to encourage participants to disclose their WTP more truthfully than hypothetical contingent valuation methods (Shogren et al. 1994). However, while several studies have used the "adding-up" test to examine the internal consistency of estimates obtained from stated preference based contingent valuation instruments (e.g., Veisten et al. 2004; Nunes and Schokkaert, 2003; Adamowicz, et al. 2011), scant information is available on whether *non-hypothetical* experimental auction mechanisms might have the same shortcoming as hypothetical contingent valuation methods with regards to the "adding-up" properties. To our knowledge, only the studies by Corrigan and Rousu (2006) and Loureiro, Gracia, and Nayga (2013) have examined related questions. Specifically, Corrigan and Rousu (2006) investigated valuations for a second unit of an auctioned good, when participants are endowed with the first unit. They found that endowing participants with a free good resulted in higher bids for subsequent units of the same good. On the other hand, Loureiro, Gracia, and Nayga (2013) evaluated the issue of scope test in experimental auctions using quasi-public goods. They assessed the valuation for a qualitative, rather than quantitative, improvement in the nature of a quasi-public good. They tested how WTP changes when differential attributes are added to a regular good in the context of experimental auctions. They conducted experimental auctions on four products, in particular, the regular version of the product and three enhanced versions of this product (i.e., cured ham). Their results generally suggest that valuations show sensitivity to scope of the good, particularly when using simultaneous bidding for all products.

The two studies described above have focused on estimating WTP for single units of goods. Limited attention has been devoted to explicit application and performance of such techniques in multiunit value elicitation formats (Elbakidze, et al. 2013; Loureiro, et al. 2013; Akaichi et al. 2012). The advantage of collecting bids for various quantities of products is that this format allows one to reflect potential changes in preferences across different quantities of products. In other words, diminishing marginal benefits can be accounted for. This approach allows for estimation of demand schedules rather than estimation of a point on a demand curve. WTP values estimated using single unit auctions can generate biased estimates of aggregate demand (Elbakidze et al. 2013) due to the implicit assumption that the individuals are interested in purchasing only one unit (Lusk and Schroeder, 2006). The non-hypothetical experimental auction designs in Elbakidze et al. (2013) and in Louriero, et al. (2013) reflect the effects of scope - meaning that WTP should change as the quantity of the good is made smaller or larger (Lew and Wallmo, 2011; Carson et al., 2001). However, the experimental designs in those studies are not appropriate to explicitly implement the "adding-up test" with and without free provision of a good as needed to examine the "income effect. Furthermore, Loureiro et al.

(2013) examined scope sensitivity of estimates obtained from non-hypothetical experimental auctions in the context of estimating WTP for qualitative improvements in quasi-public attributes of a given good without explicitly addressing the "adding-up" test. The number of goods is held constant while a qualitative attribute is varied across auctioned items.

We propose to contribute to this literature by explicitly examining the "adding-up" properties of estimates obtained from non-hypothetical multiunit experimental auctions. Specifically, our design examines the "adding up" properties for quantitative rather than qualitative variations across alternative bundles. Hence, our experimental design employs a multiunit value elicitation format, holding qualitative attributes constant, which allows for direct implementation of the "adding up test" in addition to assessment of the scope test.

The goal of this research is to advance the scientific foundations of tools used for estimating some of the non-market values often used as part of BCA-based policy evaluation and recommendations. Credibility and validity of valuation methods used for estimating non-market values are critical for making sure that the right policy recommendations are made based on proper economic analysis. Recent growth in the popularity of experimental economics in valuation literature calls for studies that validate and/or identify limitations and improve the techniques consistent with basic the economic principles and assumptions. This is important for robustness of policy analysis and recommendations. *Hence, in this study, we examine the internal consistency of non-hypothetical experimental auctions and associated data in the context of multiunit demand valuation*.

It is important to emphasize that the objective of this study is not to defend or challenge hypothetical stated preference based contingent valuation methods. Instead, *our primary objective is to examine the use of the "adding-up" test as an instrument for assessing the* *reliability of stated preference based estimates of WTP*. We apply the "adding-up" test in the context of *non-hypothetical* multiunit experimental auction. The results of this project have relevance not only for non-hypothetical experimental auctions but also for the debate about the internal consistency of hypothetical stated preference based contingent valuation methods. For example, a confirmation of the adding up property of estimates obtained via a *non-hypothetical* auction technique can be a sign that a hypothetical nature of stated preference based valuation methods, along with mechanism for value elicitation, may indeed be a cause for potential internal inconsistency of obtained value estimates. On the other hand, failure to confirm the adding up property of the data collected via *non-hypothetical* experiments may indicate that the inconsistency of stated preference based value estimates may be due to reasons other than the hypothetical nature of the instruments. One could argue that it is not necessarily just the hypothetical nature of stated preference based valuation methods that causes this problem.

We also examine the role of the "income effect" (Hausman, 2012; Corrigan and Rousu, 2006) within the context of the "adding-up test" to validate the internal consistency of estimates obtained from stated preference based valuation methods. If the results indicate that the "income effect" can be present and can influence participant's behavior causing the "adding-up" property not to hold, then the appropriateness of using the "adding-up test" for assessing internal consistency of estimates obtained from stated preference based studies can be questioned. In this study we also control for whether the participants seem to behave as expected utility maximizers or not. By observing participants' choices in non-hypothetical gambles we examine the robustness of the results with respect to consistency with expected utility maximization based choices (Horowitz, 2006; Banerji and Gupta, 2014; Kurata et al. 2009). The results are important for appropriateness of using the "adding-up" test for the purposes of validating stated preference based contingent valuation studies or non-hypothetical experimental auction studies. The robustness of estimates from such studies is important because these estimates are normally used for welfare and benefit cost analyses with important marketing and policy implications for sustainability of coupled natural and human systems.

2. Multiunit experimental auctions

Shogren et al. (1994a) argue that an experimental lab provides a viable alternative to stated preference based hypothetical contingent valuation surveys. The advantage of techniques like experimental auctions is that real money and real goods are used to create valuation environments with tangible consequences and incentives. These methods are non-hypothetical in the sense that the participants are required to exchange real money for products depending on their performance during the experiments. As a result, such mechanisms provide the participants a motivation to disclose their willingness to pay (WTP) more truthfully (Cummings et al. 1995; Fox et al. 1998; List and Shogren 1998). Hence, as appropriate in many contexts, non-hypothetical experimental auctions (Corrigan and Rousu, 2006; Rousu et al., 2004; Louriero et al. 2013; Elbakidze et al. 2013) are increasingly used to estimate WTP. The participants in non-hypothetical experimental auctions are made aware that if they submit bids that are higher than their true WTP then they face the risk of having to pay too much for the item. They are also aware that if they do not bid high enough then they risk being outbid by other participants and may miss an opportunity to purchase the item at a price that could have been acceptable. Thus, the mechanism appears to be more incentive compatible than stated preference based contingent valuation methods, possibly reducing the potential for internal inconsistency of collected data. However, the method is still a lab-based technique which does not necessarily exclude the possibility that the participants may behave differently from real contexts outside of the lab.

Complications like the "top dog effect" (Shogren et al. 2001a), "anchoring effect" (Corrigan et al. 2012) and other issues may result in departure from internal consistency of the data to some degree.

As previously discussed, almost all applications of these techniques have focused on estimating WTP for single units of goods like various food items (Rousu et al. 2004; Corrigan et al. 2009; Masters and Sanogo, 2002; Shogren et al. 2001b). Limited attention has been devoted to application and performance of such techniques in multiunit value elicitation formats (Elbakidze, et al. 2013; Loureiro, et al. 2013; Akaichi et al. 2012). Hence, we propose to contribute to this literature by explicitly examining the "adding-up" properties of estimates obtained from non-hypothetical multiunit experimental auctions. Concerns about the effects of "demand reduction" on the efficiency of uniform price auctions in a multiunit auction formats have been documented (List and Lucking-Reiley, 2000; Engelbrecht-Wiggans, et al. 2006; Ausubel 2004). Only a handful of studies have applied experimental auction mechanisms for estimating WTP in multi-unit formats using auction designs where problems with "demand reduction" are not present (Elbakidze et al. 2013; Loureiro, et al. 2013; Rousu et al. 2008). Unlike traditional multiunit auctions (List and Lucking-Reiley, 2000; Engelbrecht-Wiggans, et al. 2006; Ausubel 2004), the studies by Elbakidze et al. (2013) and Rousu et al. (2008) design the auction mechanisms such that each submitted bid competes with others' bids for the same quantity of the good. In other words, the participants in these auctions, unlike traditional multi-unit auctions, do not submits bids for however many units they wish to purchase. Instead, each player bids for the same fixed quantities of products in each round and only the bids for corresponding quantities are used to determine the winner(s) in randomly determined quantity scenario. This kind of design eliminates a potential problem of demand reduction.

3. Approach

Non-hypothetical experimental auctions, like Vickrey uniform second (or third, fourth, etc.) price auctions (Dickinson and Baily, 2002; Corrigan et al , 2009; List and Shogren 1999) and random N-th uniform price auctions (Shogren et al., 2001a; Rousu et al. 2004; Shogren et al. 1994) have been used in various contexts. In a typical uniform price auction, the participants submit individual bids and all winners are expected to pay a price equal to fixed $(N-1)^{th}$ (2nd, or 3rd, or 4th, etc.) highest bid (Shogren et al. 1994; List and Shogren, 1999; Alfnes and Rickertsen, 2003) or randomly determined $(N-1)^{th}$ highest bid (Shogren et al. 2001a; Rousu et al. 2004) depending on the mechanism. In this study, we use the Becker-DeGroot-Marschak (1964) (BDM) auction. The BDM auction is theoretically equivalent to the second price auction, random nth price auction, and English auctions (Lusk, Feldkamp and Schroeder, 2004; Lusk and Rousu, 2007). However, various critical evaluations have appeared in the literature (Karni and Safra, 1987; Noussair, Robin and Ruffiex, 2004; Horowitz, 2006; Buckley et al., 2012). We use BDM mechanisms because it is easier to understand by the participants and because it allows for individual performance outcomes independent of performance of the other participants. Unlike uniform price auctions, the BDM mechanism also allows for heterogeneity of the binding price across winning participants. We rely on this feature of the BDM mechanism to examine the effect of expenditures in earlier rounds on the performance in subsequent rounds as explained below. In addition, it has been argued that the high bidding problem can be less severe in the BDM than in the Vickrey second price auction (Noussair, Robin and Ruffiex , 2004).

3.1. The adding-up test

Diamond and Hausman (1994) proposed the "adding-up test" as a means for validating the results of contingent valuation surveys to make sure that the survey responses, which are used for estimating non-market values, are based on stable well defined preferences, rather than based on "opinions invented on the fly" or "irrational preference instability" (Hausman, 2012). Diamond and Hausman (1994) argue that the adding-up test, as one of the tests of internal consistency of survey response data, can be used to assess the reliability and validity of contingent valuation survey data for non-market valuation. The adding-up test involves eliciting WTP for good X, WTP for good Y conditional on having been supplied X, and WTP for (X+Y). The adding-up test requires that the following equality be confirmed

$$WTP(X+Y)=WTP(X)+WTP(Y|X) \qquad (1)$$

If this equality is rejected, then Hausman (2012) and Diamond and Hausman (1994) argue that the data are not based on internally consistent preference structure and therefore cannot be used for valuation purposes.

Hausman (2012) reports a prominent case study where the adding-up test failed. He uses the Contingent Valuation method to elicit consumers' willingness to pay for various bundles of cable television packages. Hausman finds that willingness to pay for bundle X plus willingness to pay for bundle Y given free provision of bundle X do not equal willingness to pay for bundle (X+Y). Thus, he concludes that the null hypothesis of the adding-up test is rejected. Based on this result the conclusion is made that the data are not internally consistent with basic economic assumptions. In Hausman (2012), as well as in Diamond and Hausman (1994), "income effect" is mentioned as a possible reason why adding-up test might not hold. However, in both papers the "income effect" is viewed as theoretically a possible explanation but empirically irrelevant because typical values associated with non-market goods being valued in contingent valuation surveys are very small relative to total income. Therefore, any potential "income effect" that can be observed will be negligibly small.

In our experiment, we revisit the above issues in a non-hypothetical experimental auction context. The above equation is tested in the context when participants are asked to disclose their willingness to pay for 1 and 2 units of the same product. The participants submit bids for (a) two units of the product (left hand side of equation 1), (b) one unit of the product (first component on the right hand side of equation 1), (c) one unit of the product given that they receive another unit for free regardless of the outcome (second component on the right hand side of equation 1, under free provision of one item), and (d) one unit of the product given that they had already purchased the one unit of the same product (second component on the right hand side of equation 1, having paid for another unit of same item). Hence two versions of equation (1) are examined - with free endowment, (a)=(b)+(c); and no endowment, (a)=(b)+(d).

Prior contingent valuation studies as well as experimental auction studies have treated total income, or total disposable income (Loureiro, et al. 2003), as the budget constraint which may influence individual behavior in contingent valuation surveys and experimental auction experiments. Hausman (2012) and Diamond and Hausman (1994) argue that since the expenditure on the good being valued generally comprises only a small proportion of total income, endowment with the first unit cannot be expected to cause a disparity between the left hand side and the right hand side of the equation above via the "income effect". Similar argument appeared in Corrigan and Rousu (2006). An alternative way to conceptualize a representative participant's behavior is that the participant may pre-allocate some small proportion of his or her disposable income to such nonessential, low priority, activities as voluntary contributions to public goods (like environmental quality projects or expenditures in exercises like non-hypothetical experimental auctions). If the participants' allocation of disposable income for such activities is small (which is likely the case as most participants in such studies rarely have large slacks in their incomes) then the endowment with the first unit (X) may have a non-trivial "income effect" for spending behavior in the experimental auction. If this is true, then for WTP estimates suitable for testing the "adding-up" property, the rejection of the above hypothesis does not necessarily indicate that the data show signs of "irrational preference instability" and therefore cannot be used for benefit-cost analysis. Instead, the reason might be the "budget effect", which may be present if participants allocate only a small portion of their total income for expenditures on public goods and non-essentials like purchases in laboratory experimental auctions.

In this study, we theorize that the use of aggregate income (monthly, annual, etc.) may not be appropriate for representing the budget (or income) effect on the behavior of participants in WTP experiments or contingent valuation surveys with multiple bundles of goods. Instead, we ask whether the "slack" income (after subtracting fixed expenses on housing, food, transportation, other regular expenditures, entertainment, savings, etc.) may be more appropriate to capture the "income effect" of endowing the participants with a free unit of a good when eliciting WTP for a second good. Only a small portion of the disposable income might be relevant for possibly being used for spending in WTP experiments or as contributions to funds for provision of non-market goods, as estimated via contingent valuation surveys. In such case, endowment with the first unit in the adding-up type of tests may make a difference through the "budget effect".

Corrigan and Rousu (2006) examine the effects of an initial endowment on WTP. In their study the initial endowment is the first unit of a good which the participants receive for free in addition to the compensation they receive for participating in the experiment ($5). Corrigan and Rousu proceed to test whether WTP for the second unit, when endowed with the first unit for free, equals to the difference between WTP for two units and WTP for one unit without endowment. This is essentially a form of the adding-up test proposed by Diamond and Hausman (1994). Corrigan and Rousu find that WTP for the second unit, when endowed with the first for free, exceeds WTP for the second unit implied by the difference between WTP for two units and WTP for first unit. Their explanations for such "abnormality" include the "top dog effect" and the "reciprocity" effect. They also consider the potential of the income effect but reject this rationale because very small increase (0.22%) in monthly income (represented by the value of the first unit endowed for free) would not be expected to produce such a significant (64%) increase in participants' WTP for the second unit. Similar to most of the literature on estimation of WTP, Corrigan and Rousu use total income (monthly) as the relevant measure of budget which may affect the behavior of participants. Total income is treated as the budget constraint which can affect individual behavior in these experiments. In this study we conjecture that if the participants consciously or subconsciously allocate only a small proportion of their disposable total income for activities like in-kind contributions to funds for public goods, or expenditures in experimental auctions, then the value of endowment with the first unit may not be negligibly small relative to the total budget that participants might allocate for such activities. We ask whether the total income effect, as measured via total disposable income, may differ from a potential effect of a budget which experiment participants and survey respondents might be willing to allocate for expenditures on non-market goods and WTP experiments.

We test whether the self-reported monthly income has a statistically significant effect on the bids. We also consider the effect of non-refundable expenditures in the first round of the experiment on the size of the bids in the subsequent round. In other words, does having paid for the first item affect the value of bids they submit for the second item? Further, we test whether the self-reported expenditures on public goods and charity have a statistically significant association with submitted bid values. If we find that expenditures on the first item, depending on whether or not the first item is provided for free, affect bids for the second item then it must be true that small changes in income have an effect on bidding behavior controlling for decreasing marginal utility of additional purchased units. Such a result, despite previous arguments that very small changes in income should not significantly affect bidding behavior (Corrigan and Rousu, 2006), would imply that the effect of endowment with the first item on bidding for the second item may not be insignificant even if the effect on total income is small. In such case the "adding-up test" can be rejected not necessarily because the estimates are internally inconsistent, but because "income effect" of endowment, according to the structure of the "adding-up test", can cause identity in equation (1) not to hold.

3.2. Experimental design

Each experiment session consists of 3 rounds (rounds A, B, and C). One of the rounds is then randomly chosen as the binding round at the end of the session. The winners in the binding round (and only in the binding round) are expected to purchase the items they win at the binding price(s) in the binding round. In each round, the participants submit competitive sealed bids for one or two units of a familiar candy bar. Two of the rounds (rounds B and C) have one auction each and one of the rounds (round A) consists of two auctions. In the round with two auctions (round A), the participants bid for one candy bar at a time in each auction. The participants are aware that if this round is chosen as binding then they are expected to purchase the items they win in either or both of the auctions in this round. After submitting a bid for one candy bar in the first auction, the participants find out if they won in that auction before submitting a bid for another item in the second auction in this round. Hence, the participants know the outcome of the first auction in this round before submitting a bid in the second auction. In the other two rounds (rounds B and C), the participants bid for either two candy bars as a bundle (round C) or for one candy bar (round B) after being told that if this round is chosen as binding at the end of the session then they will receive one candy bar for free regardless of their performance in this auction/round. Hence in this round, the participants are bidding for the second item after being endowed with a free first item. The order of the three rounds varies across 6 treatments[1] and 18 sessions with 3 sessions for each of the treatments. The order of treatments is randomized across 18 sessions.

We also examine whether behavior consistent with expected utility maximization has an effect on the outcome of the study. Several studies have examined the role of expected utility maximization in BDM auctions (Horowitz, 2006; Banerji and Gupta, 2014; Kurata et al. 2009). As part of the experiment sessions all participants are asked to choose two non-hypothetical gambles from two sets of four gambles where they have a chance of winning $0, $1, $2, $7, or $12 combined. The choices of participants are checked for consistency with expected utility based choices.

3.2.1. Steps in a representative experiment session

In this section we provide a sequence of activities from treatment 1 for illustration purposes. The steps in other treatments are similar and differ only in terms of order of rounds across treatments.

1. Upon arrival the participants receive $10 broken into coins and $1 bills and are told that they will also receive $5 upon conclusion of the session. The participation payment is broken into small bills and coins to facilitate bidding and transactions. They also receive the packets containing unique numerical IDs (which allows for anonymity of collected

[1] Six treatments according to sequence of rounds A, B, and C: (A,B,C), (A,C,B), (B,A,C), (B,C,A), (C,A,B), (C,B,A).

data), the instruction materials, auction mechanism explanations, bidding forms (including practice and actual rounds), and socio-demographic survey.

2. The participants fill out and return the consent forms after a brief preview of the activities in the session.
3. The participants are asked not to turn pages until instructed to do so throughout the session.
4. It is explained to the participants that their names will not appear on any documentation except for the consent form. Since the consent forms are collected separately from upcoming bid forms and the survey instrument, the anonymity of their responses and performance in the experiment are guaranteed.
5. The experimenter clearly explains the BDM auction mechanism with examples.
6. The experimenter explains to the participants and demonstrates that in BDM auctions, it is always in their best interest to submit bids equal to their maximum willingness to pay.
7. The participants fill out and return quizzes designed to test their understanding of BDM mechanism. The quiz involves answering two multiple choice questions (one in the context of underbidding and one in the context of overbidding) and one open ended question asking how much would the participant pay (if any) to obtain the auctioned item under a particular submitted bid and binding price scenario. After the participants mark their answers for the two questions, the experimenter will discloses the correct answers for the first two questions before the participants answer the third question.
8. The participants are reminded that prior to submitting their bids in each round they will not know what is the binding price. The price is disclosed after the bids are submitted. The participants are told that binding prices have been randomly determined and have been

placed in labeled sealed envelopes included in their packets. The envelopes are labeled according to the bidding round order including practice round.

9. The participants take part in a practice round and find out the outcome. They record their hypothetical bid for a hypothetical candy bar on the practice round bidding sheet. Next they are asked to open the envelop labeled as “Practice Round”, extract the predetermined binding price, and record it in the box next to their submitted bid in this round. Next, they are asked to circle “yes” or “no” next to the question “Did you win in this round”?
10. The experimenter informs the participants that they will be taking part in three non-hypothetical, non-practice rounds. The participants are also told that one of the rounds will have two bidding auctions. The other two rounds will have only one bidding auction each. In total there will be four auctions. This structure is displayed in front of the room. The experimenter explains that in each auction the winners will be determined according to BDM auction mechanism as seen in steps 4, 5, and 6, and in the practice round.
11. The participants are told that either round 1, round 2, or round 3 will be chosen as binding according to a random draw at the end of the session. The winners in the randomly chosen round will be expected to purchase the items using the binding price from that round. If the round with two auctions is selected as binding then the winners in both and either of the auctions will be expected to purchase the item(s) they win.

The following steps correspond to treatment 1(round order A, B, C). Other treatments differ only in terms of order of rounds.

12. *Round A: Auction 1)* – the participants record their bids for the first unit of the item. The participants are then asked to extract the binding price from an envelope labeled "Round A, Auction 1" to see if they win or not. Experimenter records individual bids.
13. *Round A: Auction 2)* – the participants record bids for the second unit of the item knowing that if they are the winners in either or both auctions in round 1 then they will be purchasing the units they win in either or both auctions they win if round 1 is chosen as binding. After marking their bids the participants are asked to extract the binding price from envelopes labeled "Round A, Auction 2". The experimenter records individual bids.
14. *Round B* –The participants are told that if round B is determined to be binding then all participants will receive one unit for free regardless of their preference or bids. Hence, the subjects are submitting bids for the second unit in this round. Before the bidding, the participants are reminded that only one of the three rounds will be randomly determined as binding. After recording their bids the participants are asked to extract the binding price from the envelope labeled "Round B". The experimenter records the bids.
15. *Round C* – the participants bid for two units as a bundle after being told that if round 3 is selected as binding then the winners in this round will be purchasing 2 units for the binding price in this round. The participants are asked to extract the binding price from the envelope labeled "Round C". The experimenter records the bid values.
16. The experimenter randomly determines and announces the binding round (round A, B, or C) according to randomly drawn card labeled A, B, or C.
17. The participants are asked to make two choices between two gambles: choose between gamble 1A and gamble 1B, and between gamble 2A and 2B. All gambles are non-hypothetical. Depending on their choices they win \$0, \$1, or \$6 in each gamble.

a. Choose between gamble 1A (get $1 with certainty) or 1B (80% chance of getting $1, 15% chance of getting $0, and 5% chance of getting $6)
b. Choose between gamble 2A (80% chance of getting 0; 20% chance of getting $1) or gamble 2B (95% chance of getting $0; 5% chance of getting $6)

18. All participants fill out the socio-demographic survey.
19. The participants turn in their surveys and draw cards from stacks labeled 1A, 1B, 2A and 2B corresponding to their choices made in step 17. The cards in respective bins are labeled with $0, $1, or 6$ corresponding on probabilities for particular bins.
20. The winners exchange money for the items they win, and receive the remainder $5.
21. The participants are asked not to discuss the experiment with anyone over the next few weeks.
22. The experimental session concludes.

3.2.2. Hypotheses

Hypotheses 1, 2, and 3 examine the adding-up test conditional on whether or not the first unit is provided for free when bidding for the second unit. The objective of these hypotheses is to empirically examine the reasoning that "income effect" from endowment should be negligible in these types of experiments (Diamond and Hausman 1994; Hausman 2012; Corrigan and Rousu 2006). Hypothesis 1 applies the adding-up test as proposed by Hausman (2012) where WTP for the second item is elicited given that the first item is provided free of charge.

H1, (Adding-up test, free first item):
$H_{o:}$ (Bids in the first auction of round A) + (bids in the round B) = (Bids in round C)
$H_{a:}$ the above equality does not hold

Hypothesis 2 addresses the adding-up property in a context where the first item is not provided free of charge when eliciting WTP for the second item. This specification of the adding up test is designed to capture the "budget effects" that may be present after purchasing the first item. To test hypothesis 2 only the bids submitted by the participants who win in the first round of round A are used.

H2 (Adding-up test, no free items)

$H_{o:}$ *(Bids in the first auction of round A) + (Bids in the second auction round A) = (Bids round C).*

$H_{a:}$ *The above equality does not hold*

Hypothesis 3 is another way of analyzing the difference in participants' bids for the second item depending on whether the subjects pay for the first unit or obtain the first unit for free. The data of only those who win in the first auction of round A is used. The bids of these subjects from the second auction of round A and the bids from round B are compared directly. If the bids from these two auctions are found to be not statistically different then no "income effect" due to paying for the first item must be present.

$H3_{o:}$ *(Bids in the second auction of the round A) = (bids in the round B)*

$H3_{a:}$ *The above equality does not hold*

Hypothesis 3 is intended to provide an additional way of looking at the "budget" effect on bidding. Specifically we are interested in the effect of the amount of money spent on the first unit on bids for the second unit. The data of only the winners in the first auction of round A will be used

$H4_{o:}$ *The amount spent in the first auction of the round A has a statistically significant effect on the difference between bids in the second auction of round A and the bids in round B.*

H4a: not $H3_o$

In addition to the above hypothesis we also examine the effects of self-reported total monthly income, self-reported expenditures on voluntary contributions to public goods and charities, and other socio-demographic variables on bidding behavior.

3.3. **Data**

One hundred and ninety one subjects were recruited to take part in the experiment. Each participant was assigned to one of 18 sessions. Each session represents one of six treatments based on the order of rounds A, B, and C. Hence, each treatment was repeated in three sessions randomly ordered across 18 sessions. Table 1 provides summary statistics for participants across treatments and for combined data. Approximately 80 percent of participants were undergraduate students ranging from 17 to 60 years in age. Average household size was 2.5 individuals with average household income falling within 3 to 4 thousand dollars per month. Average individual expenditures on non-essentials (which excludes expenditures on housing, household necessities, utilities, tuition, various debt and loan payments, food, clothing, transportation, medical bills) is $165 per month. Participants were asked to report their guess for the price of the auctioned candy bar. Average guess was $1.65. About 87 percent of the sample correctly answered 4 out of 4 practice questions. 97 percent of the sample answered at least 3 out of 4 questions correctly. Approximately half of the participants chose lotteries at the end of the session in a manner consistent with expected utility maximization.

4. Results

Table 2 provides summary statistics of bids across subgroups of participants as well as combined sample. Mean bids and associated standard deviations are reported for all participants, expected utility maximizers (subjects whose choices of lotteries were consistent with expected utility maximization-based behavior) and non-expected utility maximizers (subjects whose choices of preferred lotteries were not consistent with expected utility maximization based behavior). The table also separates bids submitted by those who won in the first auction of round A from those who did not win in the first auction of round A. Summary statistics of subsamples in which participants correctly answered 4 out of 4 practice questions, and at least 3 out of 4 practice questions, are also reported. In all cases, bids for two items as a bundle exceed bids for one unit consistent with the scope test (Carson 1997). Bids for the second item when endowed with first for free (Round B) are less than the bids for the first item (Auction 1 in Round A) in all cases except for those subjects who did not win in the first auction of round A[2]. It is also interesting that the average bids for the second item when endowed with first for free (Round B) are consistently greater than bids in the auction 2 of round A regardless of whether they won or did not win in the first auction of round A. The reason for this result maybe the reciprocity effect where the participants may wish to repay the experimenter for the free candy (Corrigan and Rousu, 2006). Bids in auction 2 of round A are generally less than bids in auction 1 of round A. This appears to be particularly true for those who win in auction 1 of round A, consistent with decreasing marginal value. However, this is not necessarily true for those who lose in auction 1 of round A. It is possible that those who lose in auction 1 of round A adjust their bids upward to increase chances of winning. For those who win in auction 1 of round A, bids for two items (round C) appear to be less than either the sum of the bids in auction 1 and auction 2 of round A

[2] Total of 21 individuals who lost in the first auction of round A bid more for the second item when endowed with first for free (Round B) then for the first item (auction 1 of round A). Of these individuals 18 also lost the second auction of round A. 13 of these subjects were in treatments 1,2, or 4 where round A preceded round B. It is possible that these individuals adjusted their bids upward after losing in round A. It is also worth noting that the majority of the individuals who lost in auction 1 of round A bid less in round B then in the first auction of round A as one would expect.

or the sum of bids in auction 1 of round A and bids in round B (bids for second unit with first free). This is not necessarily true for those who did not win in auction 1 of round A. In fact, bids in auction 1 of round A plus bids in round B appear to equal bids in round C. However, bids in auction 2 of round A (which are also bids for the first item since they lost in auction 1 of round A) plus bids in round B exceed bids in round C.

Skewness and Kurtosis (D'Agostino et al. 1990) as well as Shapiro-Wilk (Royston 1992) tests strongly reject the hypotheses that bids from the four auctions are normally distributed[3]. Therefore, nonparametric sign tests are used for comparison of matched data (Snedecor and Cochran, 1989). Table 3 provides p-values for respective hypothesis listed in column one. The results suggest that bids for two units as a bundle (round C) tend to be less than the bid for the first unit (round A, auction 1) plus bids for the second unit when first is provided for free (round B) except for those participants who lose in the first auction of round A. For those participants who lost in auction 1 of round A (i.e. they are not expected to purchase the first item) there is no statistical evidence to reject the null hypothesis that the bid for two items as a bundle equals the bid for the first item plus bid for the second item when first is provided for free. This result holds for those subjects whose choice for participation in lotteries are consistent with expected utility maximization as well as for those whose choices are not consistent with expected utility maximization. There is also no statistical evidence to reject the null hypothesis that for those who lost in auction 1 of round A the median of the difference between the bid for two items as a bundle and the bid for the item in auction 2 of round A plus the bid for the second item in round B given first item is free is zero. However, according to the sign test, these participants' bids in

[3] Test results are available upon request.

second auction of round A were greater than in the first auction of round A confirming that perhaps after losing the first auction, the participants may have adjusted their bids upward[4].

In hypothesis 2, generally the bids for two units as a bundle tend to be less than the sum of bids for the first item (round A, auction 1) and for the second item (round A, auction 2) under no endowment. This result, similar to hypothesis one, does not necessarily hold for those subjects who lost in the first auction of round A. For these participants the equality seems to hold. However, this specification of the adding up test for the losers in auction 1 of round A is incorrect because we are comparing a bid for two units with the sum of the bids for one unit and for second unit which is submitted knowing that the bid for first unit is when the bid for first unit is not binding.

In hypothesis 3 bids for second item when first is provided for free (round B) generally tend to be greater than bids for the second item when first is not provided for free (round A, auction 2). In other words, subjects are generally willing to pay more for the second item if the first one is provided for free than when the first item is not free. This result holds for subjects who win the auction for the first item (round A, auction 1) but not for subjects who do not win in the first item. The latter group's bids for the second item when the first item is provided for free tend to equal the bids in the second auction of round A. However, for this group bids in auction 2 of round A are bids for their first item because they did not win in the first auction of round A. Hence, for this group the value of the second item when the first item is provided for free is actually the value of the first item. For the former group the bids for the second item when first is free (round B) tend to be greater than bids for the second item after paying for the first item (auction 2, round A). The reason could be either income effect, due to not having to pay for the

[4] Test statistics are not reported in table 3 but are available upon request.

first item, and/or reciprocity effect (Corrigan and Rousu, 2006). These results hold for expected utility maximizers at 10% level of significance and for non-expected utility maximizers.

Hypotheses 1 and 2 generally suggest that regardless of whether the subjects paid for the first unit (round A) when bidding for second or not (round B), bids for two units as a bundle are less than bids for the first plus bids for the second. This difference seems more statistically pronounced in hypothesis 1 than 2. Specifically, for those who win in the first auction of round A at 1% significance level, hypothesis 1 suggests that bids for two items are less than bids for first plus bids for second. However, hypothesis 2 suggests that at 1% significance level we would fail to reject the null. Similar conclusions are reached for winners in expected utility maximizer and non-utility maximizer groups separately. Potential difference across hypothesis 1 and 2 would be due to differences between bids for the second item when first is free (round B) and bids for second when first is not free. Hypothesis 3 generally indicates that bids for the second item when the first item is free do not equal bids for second when the first item is not free. This is particularly true for those who win in the first item (auction 1 of round A).

5. Conclusions

The non-parametric sign test results indicate that in this experiment the adding up condition generally does not hold. Participants who bid for and win the first item tend to bid more for the second item if the first item is free than if the first item is not free. Income and/or reciprocity effects could lead to such behavior. The income effect would imply that the participants are willing to pay less for the second item if they have to pay for the first item than if they do not have to pay for the first item because of expended allowance on the first item. The reciprocity effect would imply that the participants wish to repay the experimenter for providing the first item for free by bidding high for the second item.

The presence of either the income or the reciprocity effects can raise questions about the use of the adding up test for validation of data obtained through non-hypothetical value elicitation mechanisms based on BDM auction format. However, the adding up condition in this experiment does not hold regardless of whether the first item is provided for free (at 1% level of significance) or the participants pay for the first item (at 5% level of significance). In this sense the results suggest that the data in this experiment are not internally consistent. Hence, even in non-hypothetical value elicitation formats the adding-up test can uncover departures from internal consistency assumptions of the obtained data. Therefore, it is also possible that the violations of the adding-up tests in hypothetical contingent valuation surveys are likely to be caused by reasons other than just the hypothetical nature of the survey tool.

References

Adamowicz, W., D. Dupont, A. Krupnick, and J. Zhang, 2011, Valuation of cancer and microbial disease risk reductions in municipal drinking water: An analysis of risk context using multiple valuation methods, *Journal of Environmental Economics and Management* 61:213–226

Akaichi, F., R.M. Nayga, Jr., and J. Gil, 2012, On the Use of Multi-Unit Auctions in Measuring Consumers' Willingness to Pay for Products, *Advances in Consumer Research*, 39: 311-317.

Ausubel, M. L.2004. An Efficient Ascending-Bid Auction for Multiple Objects, *American Economic Review*, 94(5):1452-1475

Banerji, A. and N. Gupta, 2014, "Detection, Identification, and Estimation of Loss Aversion: Evidence from an Auction Experiment", *American Economic Journal: Microeconomics*, 6(1):91-133

Becker, G., DeGroot, M. and J. Marschak. 1964. "Measuring utility by a single-response sequential method." *Behavioral Science*. 9: 226 – 236.

Buckley, N.J., Cuff, K., Hurley, J., McLeod, L., Nuscheler, R. and D. Cameron. 2012. "Willingness-to-Pay for Parallel Private Health Insurance: Evidence from a Laboratory Experiment." *Canadian Journal of Economics*. 45(1): 137 – 166.

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., Flores, N.E., Meade, N.F., 2001. Contingent Valuation: Controversies and Evidence. *Environmental and Resource Economics*, 19:173-210

Carson, Richard T. 1997. "Contingent Valuationand Tests of Insensitivity to Scope." In Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods, edited by R. J. Kopp, W. Pommerhene, and N. Schwartz, 127–163. Boston: Kluwer.

Corrigan J. R., and M. C. Rousu, 2006. The Effect Of Initial Endowments in Experimental Auctions, *American Journal of Agricultural Economics*, 88(2): 448–457

Corrigan, J.R., A.C. Drichoutis,J. Lusk, R.M. Nayga, Jr., and M. Rousu, 2012. Repeated Rounds with Price Feedback in Experimental Auction Valuation: An Adversarial Collaboration, *American Journal of Agricultural Economics*, 94(1), 97-115.

Corrigan, J. R., D. P. T. Depositario, R. M. Nayga, Jr., X. Wu, T. P. Laude, 2009. Comparing Open-ended Choice Experiments and Experimental Auctions: An Application to Golden Rice, *American Journal of Agricultural Economics* 91(3):837–853

Cummings, R.G., G.W. Harrison, and E.E. Rutstrom, 1995. Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible? *American Economic Review*, 85:260-266.

D'Agostino, R. B., A. J. Belanger, and R. B. D'Agostino Jr. 1990. A suggestion for using powerful and informative tests of normality. *American Statistician* 44: 316–321.

Diamond P. A. Testing the Internal Consistency of Contingent Valuation Surveys, *Journal of Environmental Economics and Management,* 30(1996):337-347 26(4):43–56

Diamond P. A. and J. A. Housman, 1994. Contingent Valuation: Is Some Number Better than No Number? *Journal of Economic Perspectives,* 8(4):45–64

Dickinson L. D. and D. Bailey. 2002. Meat Traceability: Are U.S. Consumers Willing to Pay for it? *Journal of Agricultural and Resource Economics*, 27(2):348-364

Elbakidze, L, R. M. Nayga Jr., H. Li, C. McIntosh, 2014. Value Elicitation for Multiple Quantities of a Quasi-Public Good Using Open Ended Choice Experiments and Uniform Price Auctions, *Agricultural Economics*, 45(2):253-265

Engelbrecht-Wiggans, R., J. A. List, and D. H. Reiley, 2006. Demand Reduction in Multi-unit Auctions with Varying Numbers of Bidders: Theory and Evidence from a Field experiment, *International Economic Review*, 47(1):203-231

Fox, J. A., J. F. Shogren, D. J. Hayes, and J. B. Kliebenstein, 1998. CVM-X: Calibrating Contingent Values with Experimental Auction Markets*, American Journal of Agricultural Economics,* 80:455-65.

Gracia, A., M. Loureiro, and R.M. Nayga, Jr. 2011a. Are Valuations from Non-Hypothetical Choice Experiments Different from those of Experimental Auctions?, *American Journal of Agricultural Economics*, 93: 1358-1373.

Gracia, A., Loureiro, M. L., Nayga, R. M. Jr. 2011b. Valuing an EU Animal Welfare Label using Experimental Auctions, *Agricultural Economics*, 42(6):669-677

Hausman J. A. 2012. Contingent Valuation: From Dubious to Hopeless, *Journal of Economic Perspectives*, Diamond P. A. and J. A. Housman, 1994. Contingent Valuation: Is Some Number Better than No Number? *Journal of Economic Perspectives*, 8(4):45–64

Horowitz, J.K. 2006. "The Becker-DeGroot-Marschak mechanism is not necessarily incentive compatible, even for non-random goods." *Economics Letters*. 93(1): 6 – 11.

Karni, E. and Z. Safra. 1987. Preference Reversals' and the Observability of Preferences by Experimental Methods, *Econometrica*. 55(3): 675 – 85.

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

Kurata, H., H. Izawa, and M. Okamura, 2009, Non-expected utility maximizers behave as if expected utility maximizers: An experimental test, *The Journal of Socio-Economics*, 8(4):622–629

Lew, D.K. and K. Wallmo, 2011. External Tests of Scope and Embedding in Stated Preference Choice Experiments: An Application to Endangered Species Valuation, *Environmental and Resource Economics*, 48:1-23

List, A. J. and D. Lucking-Reiley, 2000. Demand Reduction in Multiunit Auctions: Evidence from a Sportscard Field Experiment, *American Economic Review*, 90(4):961-972

List, J.A. and J.F Shogren, 1998. Calibration of the Differences between Actual and Hypothetical Valuations in a Field Experiment, *Journal of Economic Behavior and Organization* 37:193-205.

List, J.A. and J.F Shogren. 1999. Price Information and Bidding Behavior in Repeated Second-Price Auctions, *American Journal of Agricultural Economics,* 81(4):942-949

Loureiro, M. L., A. Gracia, R. M. Nayga Jr., 2013. Do Experimental Auction Estimates Pass the Scope Test? *Journal of Economic Psychology,* 37:7-17

Lusk, J.L., Feldkamp, T. and T.C. Schroeder. 2004. "Experimental Auction Procedure: Impact on Valuation of Quality Differentiated Goods." *American Journal of Agricultural Economics*. 86(2): 389 – 405

Lusk, J.L., and T.C. Schroeder, 2006. Auction Bids and Shopping Choices. *Advances in Economic Analysis and Policy*, 6(1) Art.4.

Lusk, J.L. and M. Rousu. 2007. "Market price endogeneity and accuracy of value elicitation mechanisms." In J. A. List, ed., *Using Experimental Methods in Environmental and Resource Economics*. Cheltenham: Edward Elgar. P. 20.

Masters, W. and D. Sanogo, 2002. Welfare Gains from Quality Certification of Infant Foods: Results from a Market Experiment in Mali, *American Journal of Agricultural Economics* 84(4): 974–989

Noussair, C., Robin, S. and B. Ruffiex. 2004. "Do Consumers Really Refuse to Buy Genetically Modified Food?" *The Economic Journal*. 114: 102 – 120.

Nunes, P.A.L.D. and E. Schokkaert, 2003. Identifying the Warm Glow Effect in Contingent Valuation, *Journal of Environmental Economics and Management*, 45:231–245

Rousu, M. C., W. E. Huffman, J. F. Shogren, and A. Tegene, 2004. Estimating the Public Value of Conflicting Information: The Case of Genetically Modified Foods, *Land Economics* 80(1):125-135

Rousu, M. C., R. H. Beach, and J. R. Corrigan, 2008. The Effects of Selling Complements and Substitutes on Consumer Willingness to Pay: Evidence from a Laboratory Experiment, *Canadian Journal of Agricultural Economics* 56:179–194

Royston, P. 1992. Approximating the Shapiro–Wilk W-test for non-normality. *Statistics and Computing* 2: 117–119.

Shogren, F. J., M. Margolisa, C. Koob, and J. A. List, 2001a. A Random Nth-price Auction, *Journal of Economic Behavior and Organization*, 46(4):409-421

Shogren, J. F., J. A. Fox, D. J. Hayes, and J. B. Kliebenstein, 1994a. Bid Sensitivity and the Structure of the Vickrey Auction, *American Journal of Agricultural Economics*, 76: 1089-1095

Shogren, J. F., S. Y. Shin, D. J. Hayes, and J. B. Kleibenstein, 1994b, Resolving Differences in Willingness to Pay and Willingness to Accept, *American Economic Review*, 84(1):255-270

Snedecor, G. W., and W. G. Cochran. 1989. Statistical Methods. 8th ed. Ames, IA: Iowa State University Press.

Veisten, K. H. F. Hoen, and J. Strand, 2004. Sequencing and the Adding-up Property in Contingent Valuation of Endangered Species: Are Contingent Non-Use Values Economic Values? *Environmental and Resource Economics* 29:419-433

Table 1. Participant summary statistics (means and standard deviations in brackets)

	Treatments						Combined
	1	2	3	4	5	6	
	N=34	N=30	N=30	N=31	N=37	N=29	N=191
Gender (Male =1)	0.56	0.47	0.40	0.58	0.30	0.59	0.477
	(.5)	*(.51)*	*(.5)*	*(.5)*	*(.46)*	*(.5)*	*(.5)*
Age	21.94	25.93	24.40	22.39	22.54	22.76	23.26
	(4.73)	*(10.19)*	*(8.07)*	*(7.93)*	*(5.89)*	*(7.1)*	*(7.4)*
Undergrad (yes=1)	0.85	0.73	0.63	0.81	0.84	0.90	0.79
	(.36)	*(.45)*	*(.49)*	*(.4)*	*(.37)*	*(.31)*	*(.4)*
Monthly expenditure on non-essentials ($)	205.15	180.33	133.10	140.00	172.58	154.48	165.36
	(292.94)	*(184.01)*	*(124.06)*	*(149.36)*	*(184.47)*	*(196.11)*	*(196.7)*
Individuals in household	2.44	2.00	2.30	2.87	2.54	2.93	2.51
	(1.85)	*(1.2)*	*(1.44)*	*(1.41)*	*(1.39)*	*(1.96)*	*(1.58)*
Household income ($)[a]	6.15	6.70	6.17	7.13	6.46	5.24	6.32
	(5.57)	*(5.11)*	*(4.59)*	*(5.57)*	*(5.62)*	*(4.85)*	*(5.22)*
Guess market price	1.67	1.53	1.90	1.59	1.64	1.59	1.65
	(.63)	*(.61)*	*(.68)*	*(.98)*	*(.65)*	*(.75)*	*(.73)*
Practice: 4 out of 4	0.97	0.63	0.97	0.94	0.81	0.90	0.87
	(.17)	*(.49)*	*(.18)*	*(.25)*	*(.4)*	*(.31)*	*(.33)*
Practice 3 out of 4	1.00	0.87	0.97	1.00	1.00	0.97	0.97
	(.)	*(.35)*	*(.18)*	*(.)*	*(.)*	*(.19)*	*(.17)*
Lottery choices consistent with EU	0.44	0.66	0.53	0.45	0.43	0.65	0.52
	(.5)	*(.48)*	*(.5)*	*(.5)*	*(.5)*	*(.48)*	*(.5)*

[a]Household monthly income was reported and coded in intervals: 0 –(less than $500), 1 – ($500-$1,000), 2 – ($1,000-$,500), 3 – ($1,500 – $2,000), 4 – ($2,000-2,500), …, 14 – (more than 6,500$).

Table 2. Summary statistics of bids

	All data			EU Maximizers			Non EU Maximizers		
	Combined	Auction 1		Combined	Auction 1		Combined	Auction 1	
		Winners	Losers		Winners	Losers		Winners	Losers
					All data				
N	191	144	47	100	77	23	91	67	24
Round A									
Auction 1	1.22	1.48	0.42	1.24	1.47	0.45	1.20	1.49	0.38
	(.71)	*(.6)*	*(.24)*	*(.64)*	*(.53)*	*(.22)*	*(.77)*	*(.68)*	*(.27)*
Auction 2	0.93	1.08	0.48	0.95	1.08	0.50	0.92	1.08	0.47
	(.63)	*(.65)*	*(.28)*	*(.6)*	*(.61)*	*(.28)*	*(.67)*	*(.7)*	*(.28)*
Round B	1.09	1.24	0.61	1.08	1.19	0.72	1.10	1.31	0.50
	(.79)	*(.74)*	*(.77)*	*(.81)*	*(.72)*	*(.99)*	*(.78)*	*(.76)*	*(.48)*
Round C	2.10	2.44	1.03	2.19	2.50	1.18	1.99	2.39	0.89
	(1.17)	*(1.06)*	*(.83)*	*(1.23)*	*(1.11)*	*(1.09)*	*(1.1)*	*(1.)*	*(.44)*
					With participants who answered all 4 practice questions correctly				
N	166	129	37	90	69	21	76	60	16
Round A									
Auction 1	1.25	1.48	0.46	1.23	1.46	0.46	1.28	1.50	0.46
	(.69)	*(.61)*	*(.24)*	*(.64)*	*(.53)*	*(.23)*	*(.76)*	*(.69)*	*(.27)*
Auction 2	0.94	1.06	0.51	0.95	1.09	0.47	0.93	1.03	0.57
	(.62)	*(0.63)*	*(.27)*	*(.61)*	*(.61)*	*(.27)*	*(.63)*	*(.67)*	*(.26)*
Round B	1.12	1.25	0.67	1.10	1.21	0.74	1.15	1.30	0.58
	(.8)	*(.74)*	*(.85)*	*(.83)*	*(.74)*	*(1.03)*	*(.77)*	*(.76)*	*(.54)*
Round C	2.13	2.43	1.07	2.19	2.50	1.20	2.05	2.35	0.91
	(1.18)	*(1.07)*	*(.9)*	*(1.25)*	*(1.13)*	*(1.12)*	*(1.08)*	*(.99)*	*(.47)*
					With participants who answered 3 out of 4 practice questions correctly				
N	185	142	43	100	77	23	85	65	20
Round A									
Auction 1	1.24	1.48	0.44	1.24	1.47	0.45	1.24	1.50	0.42
	(.7)	*(.61)*	*(.24)*	*(.64)*	*(.53)*	*(.22)*	*(.77)*	*(.69)*	*(.27)*
Auction 2	0.94	1.08	0.50	0.95	1.08	0.50	0.94	1.07	0.51
	(.63)	*(.65)*	*(.27)*	*(.6)*	*(.61)*	*(.28)*	*(.67)*	*(.7)*	*(.27)*
Round B	1.10	1.24	0.62	1.08	1.19	0.72	1.12	1.31	0.50
	(.8)	*(.74)*	*(.8)*	*(.81)*	*(.72)*	*(.99)*	*(.79)*	*(.76)*	*(.52)*
Round C	2.12	2.45	1.04	2.19	2.50	1.18	2.03	2.39	0.88
	(1.18)	*(1.06)*	*(.86)*	*(1.23)*	*(1.11)*	*(1.09)*	*(1.11)*	*(1.)*	*(.44)*

Table 3. Non parametric sign test p-values.

	Combined data			EU Maximizers			Non EU Maximizers		
		Auction A1			Auction A1			Auction A1	
Hypothesis	All	Winners	Losers	All	Winners	Losers	All	Winners	Losers
Hypothesis 1 (Diamond-Hauseman, 1994)									
Ho: Median (C –A1-B)=0									
Ha: Median (C –A1-B)>0	1.000	1.000	0.564	0.994	0.992	0.849	0.992	0.999	0.339
Ho: Median (C –A1-B)=0									
Ha: Median (C –A1-B)<0	0.001	0.000	0.564	0.011	0.015	0.304	0.015	0.002	0.798
Ho: Median (C –A1-B)=0									
Ha: Median (C -A1-B)<>0	0.001	0.000	1.000	0.021	0.030	0.607	0.030	0.003	0.678
Hypothesis 2									
Ho: Median (C-A1-A2)=0									
Ha: Median (C-A1-A2)>0	0.995	0.990	0.885	0.961	0.943	0.849	0.975	0.969	0.820
Ho: Median (C-A1-A2)=0									
Ha: Median (C-A1-A2)<0	0.009	0.016	0.196	0.062	0.091	0.304	0.042	0.056	0.324
Ho: Median (C-A1-A2)=0									
Ha: Median (C-A1-A2)<>0	0.017	0.031	0.392	0.124	0.182	0.607	0.085	0.111	0.648
Hypothesis 3									
Ho: Median (B-A2)=0									
Ha: Median (B-A2)>0	0.001	0.000	0.364	0.071	0.063	0.500	0.002	0.001	0.407
Ho: Median (B-A2)=0									
Ha: Median (B-A2)<0	1.000	1.000	0.757	0.957	0.965	0.696	0.999	1.000	0.760
Ho: Median (B-A2)=0									
Ha: Median (B-A2)<>0	0.001	0.001	0.728	0.142	0.126	1.000	0.004	0.002	0.815

A1 is bids in the first auction of round A (bids for one item with no endowment), A2 is bids in the second auction of round A (bids for second item contingent on performance in A1 with no endowment), B is bids in round B (bids for second item with first provided for free), C bids in round C (bids for two items).