Substitutability, Experience, and the Value Disparity:
Evidence from the Marketplace

John A. List[*]
Professor
University of Maryland

Prepared for the NBER Conference on Environmental Economics

22 April 2002

Abstract

　　Several experimental studies have recently provided strong evidence that the basic independence assumption, which is used in most theoretical and applied economic models to assess the operation of markets, is rarely appropriate. These results, which clearly contradict closely held economic doctrines, have led some influential commentators to call for an entirely new economic paradigm to displace conventional neoclassical theory. This paper refutes the generality of these experimental findings by going to a well-functioning marketplace and examining more than 350 individual decisions across various incentive compatible elicitation mechanisms. The data suggest that individuals with significant marketlike experience behave largely in accordance with neoclassical predictions: any observed WTA/WTP disparity amongst this group can be explained by neoclassical arguments. In light of these findings, I believe that we have discarded neoclassical explanations of the value disparity too quickly. More narrowly, these empirical results have important implications for stated valuation methods, such as contingent valuation.

JEL: Q21, Q26
Key words: WTP/WTA disparity, field experiment, marketlike experience

Correspondence: John A. List, The University of Maryland, AREC, 2200 Symons Hall, College Park, MD 20742-5535, email: Jlist@arec.umd.edu; website: http://www.arec.umd.edu/jlist/

I. Introduction

Economic modeling requires several fundamental assumptions. While most of the main tenets appear to be reasonably met, the basic independence assumption, which is used in most theoretical and applied economic models to assess the operation of markets, has been directly refuted in several experimental settings (see, e.g., Knetsch, 1989; Kahneman et al., 1990; Batemen et al., 1997).[1] Nevertheless, there is some evidence that suggests the observed discrepancy between the willingness to pay (WTP) for a commodity and the willingness to accept (WTA) compensation to sell the same commodity is merely a mistake made by inexperienced consumers and through time their behavior will be more in line with predictions from neoclassical models (e.g., Knez et al., 1985; Coursey et al., 1987; Brookshire and Coursey, 1987). Yet, critics contend that the available evidence is mixed at best, and overall the data do not conclusively support the underlying premise (Knetsch and Sinden, 1987; Knetsch et al., 2001).

The major goals of this study are twofold. First, I extend the results of List (2001b) by examining explicit statements of WTA and WTP across distinct elicitation mechanisms.[2] I do this by comparing WTA statements of value with "compensated" and "uncompensated" WTP values gathered via two incentive compatible institutions: a random nth price auction and the Becker-DeGroot-Marschak (1964) discrete choice auction. Standard neoclassical arguments (see, e.g., Hanemann, 1991) suggest previous experimental tests that have found a disparity between WTA and "uncompensated" WTP

---

[1] Early evidence of the anomaly was found in contingent valuation studies (see, e.g., Brookshire et al., 1980). The interested reader should see Horowitz and McConnell (2002a) for a nice literature review of WTA/WTP studies. As a mnemonic, throughout the paper I will refer to the difference in compensation demanded and willingness to pay as the endowment effect or the WTA/WTP disparity.

[2] Since impacts of alternative market institutions have been found to be nontrivial in other settings (see, e.g., Smith, 1982; Plott, 1986), a comparison of value statements across institutions seems worthwhile.

may actually have presented results consonant with neoclassical theory since these studies compared value estimates from different indifference curves (e.g., Knetsch and Sinden, 1987; Coursey et al., 1987), thus different marginal rates of substitution apply.[3] By comparing "compensated" WTP and WTA statements of value, I control for income and substitution effects and therefore provide a more appropriate comparison of WTA and WTP values. This distinction allows a cleaner test to distinguish between neoclassical and alternative theories. Second, I push the substitutability argument a bit harder by examining whether exogenous changes in substitutability influence the WTA/WTP disparity.

Given the vagaries of using completely uncontrolled field data, answering the above queries without using laboratory techniques to separate out the consequences of factors of primary interest from the host of simultaneously occurring stimuli would be quite difficult. Yet, rather than relying on gathering data in a laboratory environment, I depart from a traditional laboratory investigation by observing actual market behavior. This procedure provides me with an opportunity to observe behavior of agents that have endogenously selected into certain roles within the market, such as being an intense or casual consumer, while simultaneously making use of certain controls afforded in an experiment. Of course, this procedure may lead to vastly different results compared to an experiment where the roles are exogenously induced by the experimenter (e.g., some subjects are given experience while others are not), but it is my belief that a rigorous examination of behavior in an actual environment that our theory intends to explain is an important step in testing the validity of economic theory.

---

[3] A comparison of movements along different portions of the same indifference curve (e.g., Brookshire and Coursey, 1987) follows similar reasoning.

The empirical results are sharp. First, I find that market experience severely attenuates the endowment effect across both elicitation institutions. This result is in accord with List (2001b), and provides a robustness check of his main findings in much different institutional settings. Second, I find that neoclassical theory does an adequate job of organizing the data, as the comparative static predictions are generally met. For example, i) "compensated" WTP better approximates WTA than "uncompensated" WTP, and ii) the observed WTA/WTP disparity provides a more reasonable estimate of the effect of income changes on WTP for experienced consumers than inexperienced consumers. Third, I find some evidence that exogenous changes in substitutability influence the degree of the WTA/WTP disparity.

The remainder of this study proceeds as follows. Section II presents a theoretical background and the experimental design. Section III provides the empirical results. Section IV concludes with a summary and directions for future research.

II. Background and Experimental Design

Ever since the initial experimental findings that suggested mere ownership of a good may induce a reference-dependent preference structure, neoclassical models have been under heavy scrutiny, as their basic independence assumption is clearly violated in such scenarios. In addition to casting doubt on closely held economic doctrines, influential commentators have used the disparity findings to question the entire basis of welfare economics and common interpretations of indifference curves. Furthermore, evidence of a disparity between WTA and WTP statements of value has been viewed implicitly as a failure of survey methods (Zhao and Kling, 1999). In a related perspective, Knetsch (1990) notes the importance of the disparity findings as follows: "it

3

is likely that, among other implications, losses are understated, standards are set at inappropriate levels, policy selections are biased, too many environmentally degrading activities are encouraged, and too few mitigation efforts are undertaken."[4]

Before becoming immersed in the field experimental design, it is worthwhile to summarize intuitively the major theoretical predictions of neoclassical and endowment theory. Assume that an individual derives utility, $u = u(g, x)$ from consuming two goods, X and G, where X could be viewed as a composite commodity, quantities of which could be measured in money units, and G is a particular consumption good. Consider Figure 1, where perfect substitution exists between the two goods. In this case, neoclassical theory predicts WTA=WTP. This insight is best illustrated via an example of an experimental design that is commonly employed in the literature: subjects enter the lab and are randomly allocated into two groups: Group 1 and Group 2. Group 1 subjects are endowed at (g_1, x_0) in Figure 1, on indifference curve U_1, while Group 2 subjects are endowed at (g_0, x_0), on indifference curve U_0. Group 1 subjects are asked to state their minimum WTA (or Hicksian equivalent surplus) to move from g_1 to g_0, while Group 2 subjects are asked to state their maximum WTP (or Hicksian compensating surplus) to move from g_0 to g_1. By the nature of infinite substitutability between x and g, the fact that these movements occur along different indifference curves is irrelevant, WTA=WTP $(x_1 - x_0 = x_0 - x_z)$. This represents Hanneman's (1991) proposition 1.

Upon relaxing the perfect substitutability assumption, and assuming quasi-concave utility functions, however, these linear indifference curves become strictly convex to the origin, as presented in Figure 2. In this case, Randall and Stoll (1980)

---

[4] Interestingly, while the disparity has been found in a myriad of settings, many policymakers have been slow to make use of the findings.

demonstrate that the WTA/WTP disparity depends on the "price flexibility of income": the elasticity of the marginal valuation of g with respect to x. Randall and Stoll (1980) show that when the price flexibility of income is close to zero the WTA/WTP disparity will be small.

Hanemann (1991) later demonstrated that the "price flexibility of income" is analytically equivalent to the ratio of the ordinary income elasticity of demand for the good to the Allen-Uzawa elasticity of substitution between the good and the numeraire. Thus, for low elasticity of substitution values, the "price flexibility of income" is large, suggesting that the ratio WTA/WTP is also large. Hanemann (1991) argues, therefore, that the large WTA/WTP disparities that have been observed in the literature can be reconciled with neoclassical theory.

Examining Figure 2, and measuring the WTA (WTP) for changes in G from g_1 to g_0 (g_0 to g_1), illustrates this result.[5] In this case, WTA $= x_1 - x_0 > x_0 - x_z =$ WTP; due to income (shifts in the indifference curve) and substitution (curvature of the indifference curves) effects, neoclassical theory predicts value divergences.[6] In practice, it would appear that the substitution effect would hold more prominence in shaping the value disparity than the income effect. The intuition behind this thought is contained in the Engel aggregation condition, which requires that the income elasticities of demand for g and x, weighted by their respective budget shares, sum to one. While this certainly limits

---

[5] Note that if one assumes utility is quasi-convex in g then a corner-solution emerges rather than an interior solution. In this case indifference curves are concave and the substitution effect becomes positive, rather than negative. In any event, the curvature of the indifference curves could merely be thought of as a reflection of convexity of the expenditure function in g.

[6] Recall that for price changes the difference between the compensating and equivalent variation depends simply on the income effect.

the size of the income effect associated with g, the substitution effect can range from zero to infinity.[7]

I should note that insights on the disparity are identical if one considers movements along *different* segments of the *same* indifference curve. One can glean this by choosing any point along either convex indifference curve in Figure 2 and moving equidistant (in g) along the indifference curve in opposite directions. Instead of both income and substitution effects causing the divergence, in this case the divergence is caused purely by the convexity of the indifference curve. Depending on the structure of the utility function, the divergence observed along different segments of the same indifference curve can be greater or smaller than differences observed when moving along different indifference curves. A majority of the WTA/WTP experimental studies cited in Horowitz and McConnell (2002a) fall into one of these two categories, either having subjects moving along different indifference curves or different portions of the same indifference curve.

As discussed in Morrison (1997), the simple intuition behind Hanneman's (1991) neoclassical arguments can be tested by appropriately compensating Group 2 subjects so they state their maximum WTP to move from g_0 to g_1 along the <u>same</u> indifference curve as Group 1 subjects. In Figure 2, this can be achieved simply by endowing Group 2 subjects at point (g_0, x_1) and asking these subjects to state their maximum WTP to move from g_0 to g_1. Comparing these "compensated" WTP values with WTA statements will

---

[7] Care should be taken not to dismiss income effects out of hand, however. Under certain scenarios, the spacing of the indifference curves is quite important. While in Figure 2, I have implicitly assumed a unitary income elasticity by making the indifference curves for U_0 and U_1 parallel displacements, if I assumed an income elasticity greater than 1, or that the marginal rate of substitution increases along any ray from the origin, the WTA/WTP disparity would increase. The opposite occurs, of course, if g is an inferior good—the marginal rate of substitution decreases along any ray from the origin.

yield a relatively clean test of whether the indifference curve is reversible, as neoclassical theory predicts. This, of course, represents the intuition of List (2001b), who examined trading patterns of sports memorabilia on the floor of a sportscard show, and trading patterns of collector pins in a market constructed by *Walt Disney World* at the *Epcot Center* in Orlando, FL.

The experimental test constructed below represents the value elicitation analogs of these trading tests in the marketplace. If preferences are defined over changes in consumption, then one could envision a kink for Group 1 subjects at point (g_1, x_0); this kink would cause the indifference curve for Group 1 subjects to pivot from U_1 to U_1^E, as displayed in Figure 3. Under neoclassical arguments, WTA should be measured as $x_1 - x_0$, but assuming the endowment effect induces the indifference curve to pivot, WTA becomes $x_2 - x_0 > x_1 - x_0$, which illustrates the agent's aversion to sacrifice a good once it becomes part of her endowment. The WTA/WTP disparity would be exacerbated if a similar kink formed for Group 2 subjects, which would pivot their indifference curves counterclockwise from point (g_0, x_1). This seems unlikely if X represents money; nevertheless, this is the general intuition behind the endowment effect, which implies that a good's value increases once it becomes part of an individual's endowment.

Following Sugden (1999) and Horowitz and McConnell (2002b), one can push the neoclassical explanation a bit further by exploiting the relationship between the WTA/WTP disparity and the effect of income on WTP. One advantage of such an exercise is that it provides a further use of the observed WTA/WTP ratios and permits an examination of whether the observed ratios are consistent with economic intuition. To add structure to the argument, we continue with Figure 2 and assume that the initial

endowment is (g_0, x_0). As previously suggested, the individual's "uncompensated" WTP to move from g_0 to g_1 is defined as follows:[8]

$$(1) \qquad u(g_1, x_0\text{-WTP}(x_0)) = u(g_0, x_0),$$

where $\text{WTP}(x_0) = x_0 - x_z$. Similarly, "compensated" WTP is given as follows:

$$(2) \qquad u(g_1\text{-WTP}(x_1), x_1) = u(g_0, x_1),$$

where $\text{WTP}(x_1) = x_1 - x_0$. WTA is defined analogously:

$$(3) \qquad u(g_0, x_0\text{+WTA}) = u(g_1, x_0).$$

Using the intuition that $x_1 - x_0 = \text{WTA}$, and these simple definitions of WTA and WTP, the following relationship is derived:

$$(4) \qquad u(g_1, x_0\text{+WTA–WTP}(x_0\text{+WTA})) = u(g_0, x_0\text{+WTA}) = u(g_1, x_0).$$

Equality of the first and third expressions implies $x_0 + \text{WTA} - \text{WTP}(x_0\text{+WTA}) = x_0$, or $\text{WTP}(x_0 + \text{WTA}) = \text{WTA}$. A first order approximation of $\text{WTP}(x_0 + \text{WTA})$ yields WTA $\sim \text{WTP} + \text{WTA} \dfrac{\partial WTP}{\partial y}$. Rewriting this leads to the basic relationship derived by Sugden (1999) and Horowitz and McConnell (2002b):

$$(5) \qquad \frac{\partial WTP}{\partial y} \approx 1 - \frac{WTP}{WTA}$$

This simple formula allows an empirical analysis of whether the observed WTA/WTP disparity provides a reasonable estimate of the effect of income changes on WTP. Given that most studies in the literature report WTA values that exceed WTP by factors greater than two, these numbers suggest that if an agent was instantly endowed with $100 (say he received a nice gift from his grandmother), he would spend at least $50 more on the good

---

[8] This mathematical formulation closely follows Sugden (1999) and Horowitz and McConnell (2002b).

in question. These sorts of insights have led most scholars to remark that the WTA/WTP ratio is much higher than neoclassical theory would predict.

Experimental Design

The construct laid out above provides a rich environment to examine predictions associated with the endowment effect and neoclassical theory. As aforementioned, Figure 3 provides the intuition behind the theory of the endowment effect. If subjects instantaneously place an "attachment value" on a good once it becomes part of their endowment, then a kink for Group 1 subjects at point (g_1, x_0) causes the indifference curve for Group 1 subjects to pivot and rotate clockwise. This leads to point (g_0, x_1) being on a lower indifference curve, which induces WTA>WTP. Thus, if the endowment effect is important, it should create a wedge between WTA and WTP values even if compensated WTP values are elicited from Group 2 subjects. Alternatively, whether indifference curves are linear or convex to the origin, neoclassical theory predicts no differences between compensated WTP and WTA, but provides an explanation for deviations between uncompensated WTP and WTA values.

The field experimental treatments attempt to create an environment that allows an investigation into these theories while simultaneously providing a test of whether market experience induces behavior more in line with neoclassical predictions. Figure 4 illustrates the general experimental design. Points U, C, and A are the initial endowments. Subjects endowed at Point U (denoting uncompensated) are asked to state their maximum willingness to pay to move from Point U to Point D; this is termed uncompensated WTP. Subjects endowed at Point C (denoting compensated) are asked to

state their maximum willingness to pay to move from Point C to Point A; this dollar value represents compensated WTP. Finally, subjects endowed at Point A (denoting accept) are asked to state their minimum willingness to accept to move from Point A to Point C, which represents WTA.

This simple experimental design is used for several different treatment types. Before presenting the various treatments, I should note that I was careful in choosing goods the individual would actually consume, rather than put up for trade or sale immediately after the transaction. During the exit interview, more than 90% of the subjects stated that they planned to consume the good in question (e.g., keep it for their own collection).[9] The good I use is a variety of Kansas City Royals game ticket stubs dated June 14, 1996. The stubs were issued for admission to the baseball game in which Cal Ripken Jr. broke the world record for consecutive games played. I was able to obtain numerous copies of the ticket stubs because I attended the event. Given that I rarely see this particular piece of memorabilia on the market, market value is difficult to determine, but I have seen a quality piece fetch as much as $40.

The sportscard marketplace is a natural venue where the good's attributes play an integral role in its valuation. This link is perhaps best illustrated via an example. Making use of the grading system of a well-known third-party, Professional Sports Authenticators (PSA), consider the value difference between a PSA graded (Gem Mint 10) Ken Griffey Jr. 1989 *Upper Deck* baseball card and an identical Griffey card, but one which has slight

---

[9] This is potentially important because some evidence suggests WTA and WTP are roughly equivalent for securities−Kahneman et al. (1990, p. 1328) note: "there are some cases in which no endowment effect would be expected, such as when goods are purchased for resale rather than for utilization." One explanation for this conjecture is that subjects dealing with resale goods do not allow themselves to get "attached to the good" because it will soon leave their portfolio. The reported results are qualitatively the same if I exclude those agents that stated they would trade or sell the commodity.

fraying on one corner—a Mint 9 card. The Gem Mint version of the Griffey Jr. card readily sells for $2700, whereas the identical Griffey Jr. card with slight corner fraying will only garner between $125 and $350. This is mainly driven by perceived scarcity and available substitutes: as of April 14, 2002 approximately two percent of all Griffey cards were graded PSA 10 (44,358 have been graded), while about 30 percent received grades of PSA 9.

To provide exogenous variation in quality across the various Ripken ticket stubs, a PSA representative and I hand-graded more than 500 ticket stubs to provide enough goods that varied across important attributes. In trying to maintain consistency with PSA guidelines and standards, we graded 180 of the stubs as "near mint" (corresponding to a PSA grade of 7) and 150 as "mint" (corresponding to a PSA grade of 9), with the remaining stubs generally grading below near mint. I then placed them in protective stubholders with the grade clearly marked. With the variation in stub quality in mind, I turn to a summary of the various treatments, which correspond to Figure 4:

Treatment A subjects:

WTA1: Subjects are endowed with a "near mint" ticket stub and asked to state their minimum willingness to accept to sell the stub. Subjects are reminded that the stub is graded Near Mint and that an examination of typical grading distributions reveals that only about 10% of *graded* items actually grade *below* near mint.

WTA2: Subjects are endowed with a "mint" ticket stub and asked to state their minimum willingness to accept to sell the stub. Subjects are reminded that the stub is graded Mint and that typical distributions of grading reveal that only about 2% of *graded* items actually grade *above* mint.

Treatment C subjects:

WTPC1:   Subjects are endowed with the average WTA value from treatment WTA1 and asked to state their maximum willingness to pay for a "near mint" ticket stub. Subjects are reminded that the stub is graded Near Mint and that an examination of typical grading distributions reveals that only about 10% of *graded* items actually grade *below* near mint.

WTPC2:   Subjects are endowed with the average WTA value from treatment WTA2 and asked to state their maximum willingness to pay for a "near mint" ticket stub. Subjects are reminded that the stub is graded Mint and that typical distributions of grading reveal that only about 2% of *graded* items actually grade *above* mint.

Treatment U subjects:

WTPU1:   Subjects are asked to state their maximum willingness to pay for a "near mint" ticket stub. Subjects are reminded that the stub is graded Near Mint and that an examination of typical grading distributions reveals that only about 10% of *graded* items actually grade *below* near mint.

WTU2:   Subjects are asked to state their maximum willingness to pay for a "mint" ticket stub. Subjects are reminded that the stub is graded Mint and that typical distributions of grading reveal that only about 2% of *graded* items actually grade *above* mint.

By crossing these 6 treatments with two distinct elicitation institutions (more fully discussed below), I have an experimental design comprised of 12 treatments. These treatments are summarized in Table 1. In Table 1, I also denote the number of subjects included in each treatment—30 subjects—for a total of 360 subjects.

Prior to discussing the experimental procedures, it is worthwhile to discuss the various treatments and hypotheses related to these treatments. As aforementioned, comparing values from WTA1 (WTA2) with WTPU1 (WTPU2) represents the typical experimental examination in the extant literature. Yet, a comparison of these WTA and WTP measures of value will not yield a clear-cut delineation between major theories: both neoclassical and alternative theories predict that value divergences may occur across these measures. Many commentators have conceded this point, but have argued the observed disparities are too large to be explained by standard neoclassical arguments. Whether the size of the divergence makes intuitive sense is up for debate, however. As Bateman et al. (1997; p. 482) quip: "In the absence of a direct measure of the price flexibility of income, or of the elasticity of substitution, it is difficult to decide whether an observed divergence between WTA and WTP is too large to be compatible with Hicksian theory".

This represents the first point of departure—gathering WTPC1 and WTPC2 values to delineate between the various theoretical explanations of the value disparity. A comparison of WTPC1 and WTA1 (WTPC2 and WTA2) measures of value provides a relatively clean test of the neoclassical model. As outlined above, by appropriately compensating subjects who are randomly allocated into the WTP treatments, one can effectively place them on the same indifference curve that WTA subjects occupy. While neoclassical theory allows value divergences along different indifference curves, movements along (up and down) identical segments of the same indifference curve should yield equivalent value estimates. Hence, I endow subjects in treatments WTPC1 and WTPC2 with money before they reveal their WTP values.

The amount of money is determined by the average realized values in the WTA1 and WTA2 treatments—WTPC1 (WTPC2) subjects are endowed with the average WTA value in treatment WTA1 (WTA2). In this case, since income and substitution effects are controlled, neoclassical theory predicts that WTPC1 = WTA1 and WTPC2 = WTA2. The endowment effect predicts that pivoting of the indifference curve will render WTPC1 < WTA1 and WTPC2 < WTA2. These disparate predictions form the basis of this test of neoclassical versus reference dependent theory.[10]

This leads to our second test of the value disparity. Since subjects in the Mint and Near Mint treatments are stating values for identical goods except for perceived scarcity, the number of available substitutes is substantially different across treatments WTA1 versus WTA2 and WTPC1 versus WTPC2. It is commonly known that high-quality sports memorabilia is very difficult to find, and subjects are reminded of this in their treatments—see above. Thus, comparison of data across these treatments represents a rough attempt at manipulating the substitutability of the good in question. Economic intuition implies that a movement to Mint stubs (WTA2 and WTPC2) from Near Mint stubs (WTA1 and WTPC2) will shift the relevant indifference curve out (income effect) and cause it to become more convex to the origin (substitution effect). In this case, neoclassical theory predicts that WTA2/WTPC2 = WTA1/WTPC1. Yet, Morrison (1997) shows that the endowment effect is positively correlated with indifference curve convexity.[11] Thus, endowment theory predicts that WTA2/WTPC2 > WTA1/WTPC1.

---

[10] Note that I am not the first to compare WTA with compensated WTP statements of value (see, e.g., Knetsch, 1989; Kahneman et al., 1990). Yet, to my knowledge this is the first paper that has systematically examined compensated and uncompensated value statements in an actual marketplace.

[11] In addition, there is some evidence that suggests the level of endowment and the good's value are positively correlated (see, e.g., Shogren et al, 1994, who find that there may be a threshold where the importance of the endowment effect may drive behavior).

An experimental finding corresponding to this value relationship is evidence in line with the endowment effect.[12]

Experimental Procedure

The experiment was run on the floor of various sportscard shows in Tucson, Arizona. Given that individuals commonly engage in buying, selling, and trading in this marketplace, if I decided to partake in a game of golf instead of running an experiment at the sportscard shows, subjects would have been engaging in similar activities. This highlights the naturalness of this particular setting, and the added realism associated with field experiments in general.

Each participant's experience typically followed four steps: (1) completing a survey, (2) learning the auction rules; 3) considering the bid (offer), and (4) conclusion of the transaction and exit interview. In Step 1, the administrator approached potential subjects entering the trading card show and inquired about their interest in baseball collectibles. If the subject was interested in baseball memorabilia, then the monitor asked the potential subject if he/she would like to fill out a survey that would take about five minutes.[13] If the individual agreed, the monitor briefly explained that in return for completing the survey the subject would receive their endowment, if applicable (for Treatment A subjects this is the good in question; for subjects in Treatment C, this is the

---

[12] A final, much more qualified and less revealing test of the substitutability effect revolves around a comparison of WTA and uncompensated WTP values. Since lower elasticity of substitution values map into larger "price flexibility of income" values, WTA2/WTPU2 *may* be larger than WTA1/WTPU1. This hypothesis must be qualified because, as discussed above, under neoclassical arguments the disparity depends on both the ordinary income elasticity of demand for the good and the Allen-Uzawa elasticity of substitution. Yet, if the ordinary income elasticity of demand is roughly equivalent for the Mint and Near Mint stubs, then neoclassical theory predicts WTA2/WTPU2 > WTA1/WTPU1. Since theoretically substitutability is related to indifference curve convexity, endowment theory makes a similar prediction.
[13] The survey is contained in Appendix A.

dollar value equivalent of the good given to Treatment A subjects). After physically giving the subject the appropriate endowment (when applicable) the subject proceeded to fill out the survey. The monitor worked one-on-one with the participant and no time limit was imposed.

In Step 2, the administrator informed subjects in Treatment types U and C (A) that they now had the opportunity to bid (offer) in an auction for the good on the table (their good). After inspecting the good, the monitor thoroughly explained the rules of the applicable allocation institution: either a random nth price auction or a Becker-DeGroot-Marschak (BDM) discrete-choice auction. As described in List and Shogren (1998), the random nth-price auction can be explained via 4 simple steps: (1) each bidder submits a bid (offer); (2) each bid (offer) is rank-ordered from lowest to highest; (3) the monitor selects a random number (n) uniformly-distributed between 2 and Z (Z bidders); and (4) in the WTP case, the monitor sells one unit of the good to each of the (n-1) highest bidders at the nth-price; in the WTA case, the monitor buys one unit each from the (n-1) lowest bidders and pays the nth-lowest bid. Appendix B contains the experimental instructions for the WTP random nth price auction. [14]

As described in Shogren et al. (2001), the BDM random pricing mechanism works as follows. If the agent is in the WTP (WTA) treatments, she is deemed a buyer (seller) and asked to determine her maximum WTP (minimum WTA) by marking an "X" on a recording sheet listing a discrete price schedule:

	I will buy (sell)	I will not buy (sell)
If the price is $ 50.00	_____	_____
.	.	.
.	.	.
If the price is $ 20.00	_____	_____
If the price is $ 19.50	_____	_____

[14] The BDM instructions are identical except for the necessary rule changes (see below).

If the price is $ 19.00   _____   _____

           .   .   .

           .

If the price is $ 0.00   _____   _____

The price range for the goods was $0 to $50.00 in steps of $0.50. This range was chosen because it encompassed all of the offers (bids) in the WTA (WTP) random nth price auctions for both good types, and it included the highest price I had seen the good bring ($40).

In the WTP random nth price auction treatments, after completing the survey, learning the auction rules, and examining the good, each participant privately wrote a bid on the bidding sheet and placed it in an opaque box. The monitor informed the participant that his or her bid would not be opened until after the show and that all bids would be destroyed when the research project was completed. In the WTA random nth price auction treatments, after physically giving the subject the good, the subject filled out the survey. He or she proceeded to learn the auction rules and then completed the recording sheet by stating his or her minimum WTA to sell the stub. Each subject worked one-on-one with the monitor and no time limit was imposed on his or her inspection of the stub. In Step 4, the administrator explained that if the participant won the auction, she would be contacted by email or telephone within three days. Upon receipt of payment, I would send her the stub (or money). I paid all postage.

In the BDM treatments, everything was identical except the timing of the transaction. After collecting the bidding sheet, one price from the sheet was selected randomly via a draw out of a box that included the 101 slips of potential prices. If a buyer was willing to pay at least the random price, she immediately purchased the stub.

If the seller was willing to accept less than or equal to the random price, she immediately sold the stub.

Given that learning may influence whether agents conform to neoclassical theoretic predictions, an added advantage of experimenting within the sportscard marketplace is that individuals bring varying levels of experience to the experiment. For example, while some agents are intense consumers, engaging in numerous trades and transactions in a typical month, others are more casual consumers who execute a transaction quite infrequently. Still others are new to the market (both because of age and agents that enter the marker in their later years), and thus provide a sense of how the behavior profile looks for these particular consumer types. To assure that sub-populations are similar across treatments, I randomize participants into *one* of the 12 treatment types. For example, in general at the top of each hour I changed treatments, and continued to change treatments until each cell had 30 observations (note that I completed the WTA and the random nth price auction treatments first because I needed these values for completion of the WTP compensated and BDM treatments).

A few noteworthy aspects of the experimental design merit further consideration. First, no subjects participated in more than one treatment. Second, note that subjects received the stub as payment for completing the survey, and had the stub in their possession while filling out the survey. These two attributes have been found to strengthen significantly the predictive accuracy of reference dependent theory. Third, since I am not testing the incentive compatibility of the two institutions, and I want to avoid any excess noise, I inform the subjects that it is in their best interest to bid (offer)

their true value in the auctions. I reinforce this notion via several examples that illustrate the optimal strategy of truthtelling in the two mechanisms.

III. Results

To put my findings into perspective, I first review the results of List (2001b), who controlled for Hicksian income and substitution effects by asking subjects to choose between bundles—in one treatment the subject is endowed with good A and has the option to trade it for good B; in a second treatment, a different subject is endowed with good B and has the option to trade it for good A. Since subjects were allocated to one of the two treatments randomly, fewer than 50% of the subjects should swap their good if an endowment effect exists.   Alternatively, if an endowment effect does not exist, approximately 50% of the subjects should trade their good. Pooling his data across more than 300 subjects yields Figure 5, which makes the trade probability a function of previous trading experience (trades in a typical month—see question 2 in the survey contained in Appendix A). In support of the received literature, List's (2001b) field evidence suggests an inefficiently low number of trades occur for naïve traders, consistent with reference-dependent preferences. In contrast, this anomaly is not evident for consumers that have obtained significant trading experience. Figure 5 reveals that individual behavior converges to the neoclassical prediction as trading experience increases. Below I also focus on trading experience and its association with the disparity. Since I have found that *overall* transactions are highly correlated with trading rates, hereafter I will interchange trading experience and "market intensity".

In Table 2, I present descriptive statistics for my new WTA/WTP data. The top panel of Table 2 provides data from the random *n*th price auction treatments, whereas the

bottom panel summarizes the BDM data. Columns 1-6 display the 6 unique treatments used for each auction type. Central tendencies of the variables reported in Table 2 suggest that the subgroups are relatively similar in important demographic characteristics. For example, while there is substantial variation between certain cells, such as average trading experience across WTPU1 (3.8) and WTPC2 (6.0) and average years of market experience across treatments WTA2 (6.9) and WTPU2 (14.3) in the random nth price auctions, in general the data suggest that the randomization procedure was reasonably effective in allocating the 360 subjects randomly into the various treatment cells.

Concerning valuation estimates, a first finding is that consistent with expectations, sellers have higher asking prices for the Mint ticket stub compared to the Near Mint stub: in the random nth price auctions the difference is $3.24 (23.9%), while the BDM auctions reveal a difference of $1.49 (12.4%). Data from the WTP treatments exhibit similar patterns—differences range from 13.8% to over 40%, with larger differences being associated with the compensated WTP treatments. Yet, I should note that all of the data are not perfectly organized. For example, economic intuition suggests that WTP values should be larger in the compensated treatments than the uncompensated treatments.[15] This expectation is realized in the Mint stub auctions, but it is not true in the Near Mint treatments, where average values in WTPU1 across both auction types are larger than average values in WTPC1. While not too much should be made of statistically insignificant differences, the astute reader will notice some similar data anomalies below.

---

[15] Note that subjects randomly allocated into the WTP compensated treatments received the average offer from the appropriate random nth price auction: each WTPC1 subject received $13.50 and each WTPC2 subject received $16.75. For consistency, I endowed subjects with these same dollar values whether they were in the BDM or random nth auction treatments.

More importantly for our purposes, overall there is a large discrepancy between WTA and both WTP measures of value. Whereas WTA values are around $13, WTP values tend to be in the $5 range. Even when subjects are compensated, the differences are large. And, all of these differences are statistically significant at the $p < .05$ level using a large-sample t-test. In each case, this finding is substantiated when one considers results from a Mann-Whitney rank-sum test of treatment differences. The rank-sum test is a standard nonparametric test that has a null hypothesis of no treatment effect, or that the two samples are derived from identical populations.

The data can be pushed a bit further by using the insights gained from equation (5): $\partial WTP/\partial Y = 1 - WTP/WTA$. Using results from the smallest observed WTA/WTP discrepancy (BDM WTA2 versus WTPC2), I find that $\partial WTP/\partial Y = 0.422$, which suggests that if an agent was instantly endowed with $100, he would spend about $42.20 on Ripken ticket stubs. Even if this comparative static represented increased spending on sports memorabilia in general, it seems much too large for the average consumer. In sum, therefore, the data indicate that there is a large value disparity in the field, and standard neoclassical arguments cannot explain them.

Another interesting trend in the data is that WTA/WTP discrepancies are somewhat smaller in the BDM auctions compared to the random *n*th price auctions. This result is unexpected, as previous laboratory data reported in Shogren et al. (2001) suggest that the disparity is larger in BDM auctions. Even though there are minor differences and they are consistent across the value elicitation cells, statistically the BDM data and the random *n*th price auction data are not significantly different, thus below I pool the data.

Table 3 presents WTA and WTP measures of value across 5 distinct subsamples of the data. Figures in the first row are derived from the pooled sample, and highlight the value disparity that is evident in the data. In the auctions for the Near Mint stub, the WTA/WTP ratios are 2.5 (WTA1 vs. WTPU1) and 2.6 (WTA1 vs. WTPC1). The ratios are comparable across auctions for the Mint stub: 2.6 (WTA2 vs. WTPU2) and 2.1 (WTA2 vs. WTPC2).[16] While these ratios are significantly different from one using parametric and non-parametric statistical tests, suggesting again that there is a value disparity evident in the field data, there is no evidence that WTA2/WTPC2 > WTA1/WTPC1. As we will see, this null result will continue to arise for finer classifications of the data, and suggests that the value disparity is not larger for goods that have fewer available substitutes, as predicted by endowment theory (Morrison, 1997).

The remaining rows in Table 3 split the data by consumer type. I define "light" consumers as those subjects that stated they trade zero times in a typical month; "casual" consumers execute fewer than 5 trades in a typical month; "intense" consumers execute 5 or more trades per month; and "super-intense" consumers trade at a rate of 9 or more times per month. While these thresholds may appear *ad hoc*, I have attempted to delineate by first and second moments of the trading variable. Because the average number of trades in a typical month is 4.7, I label subjects below 5 trades per month as "casual" and those above 5 trades per month as "intense". And, given that the standard deviation of the number of monthly trades is 4.2, I label those subjects that execute 9 or

---

[16] Recall (from footnote 12) that both neoclassical and endowment theory predict WTA2/WTPU2 > WTA1/WTPU1. While that relationship is observed here and in spots below, it is not consistently significant at conventional levels.

more trades ($\approx$ mean + one standard deviation) as "super-intense". The general findings do not rely on these thresholds.

Even though the subsamples of "light" consumers are small, the data strongly suggest that WTP is much less than compensation demanded. Results of both a small sample t-test and a Mann-Whitney rank-sum test of treatment differences suggest that WTA>WTP for every applicable treatment comparison at the $p < .05$ level. In these cases, the ratio of WTA/WTP ranges from 4.4 to 8, which is consistent with previous experimental studies, but at odds with conventional economic theory, which assumes that indifference curves are completely reversible. Using equation (5), these ratios suggest that if an agent was instantly endowed with $100, he would spend somewhere between $77.27 and $87.50 on Ripken ticket stubs. I suspect that even the most ardent of neoclassical supporters would find this figure implausibly high. Yet, note again that there is no evidence that WTA2/WTPC2 > WTA1/WTPC1. Overall, these results are mirrored in the "casual" consumer subsample. While data from the more experienced casual consumers indicate a lesser value disparity across the various treatment cells, in each case the discrepancies are significantly different from zero at conventional levels.

Data contained in the lower rows of Table 3 paint a much different story, however. For example, the data clearly do not exhibit the large value disparity that is evident amongst inexperienced consumers. In fact, quite the contrary result emerges. In the intense consumer group, the comparison between WTA2 and WTPU2 as well as WTA1 and WTPU1 provides statistically significant differences at conventional levels using a one-sided alternative. Comparing WTA with compensated WTP values yields value ratios ranging from 1.16 to 1.30, but they are never significantly different from one;

although at lesser significance levels they almost achieve statistical significance using a one-sided alternative. These findings are in accord with neoclassical theory. The first set of results—WTA versus uncompensated WTP—may well be due to a comparison between value estimates drawn from different indifference curves, thus WTA exceeding WTP is predicted by neoclassical theory. The latter set of null results, which arise from a comparison of WTA and compensated WTP, suggests that indifference curves are reversible. This prediction, while at odds with endowment theory, is consonant with neoclassical theory, which predicts that no disparity should exist in such a setting. Yet, if one takes the observed point estimates literally, WTA/WTP ratios of approximately 1.16 and 1.30 seemingly provide implausible $\partial WTP/\partial Y$ values (0.14 and 0.23).

Even though these results begin to suggest that neoclassical theory holds promise in properly organizing data for experienced subjects, empirical results for the super-intense consumers make an even stronger case. Here, I find that neither parametric nor non-parametric statistical tests can reject the equality null in any of the treatment comparisons; therefore I also cannot reject the WTA2/WTPC2 = WTA1/WTPC1 null. This result is again at odds with endowment theory and consistent with neoclassical explanations. Furthermore, computed WTA/WTP ratios of approximately 1.09 provide a more plausible $\partial WTP/\partial Y$ value of 0.08, suggesting that $8 of a $100 increase in income will be used for Ripken ticket stubs (or perhaps more generally sports memorabilia).

But how plausible is this estimated income effect? Running the risk of making too much of a few point estimates rather than relying on inference gained from the statistical tests, I will conduct some back of the envelope calculations to shed light on this issue. Average income levels in the sample are roughly $30,000 ($\approx$ 3.5 on the survey)

and average WTP values for Near Mint and Mint stubs are about $5 and $6. Assume now that the average person's income increases by $100 to $30,100. If WTP increased by the same proportion to income, it would increase by about $0.02 to $5.02 or $6.02. To get the expected increase in WTP ($8), one would need to assume that WTP increases with extremely large powers of income. This illustrates that when the WTP is only a small fraction of income, neoclassical theory predicts very small divergences, even smaller than what the point estimates herein suggest.

To provide a sense of the relationship between the value disparity and market intensity I provide Figure 6. Akin to Figure 5, Figure 6 reveals that individual behavior converges to the neoclassical prediction as trading experience increases. These empirical results supplement List's (2001b) trading results summarized in Figure 5 and provide further evidence that experienced consumers' utility functions may be void of an important attribute that tends to induce a kinked indifference curve for everyday consumable goods. Yet, with these new auction data one can more meaningfully understand what is behind the declivity observed in Figure 6.

The statistical summary presented in Table 3 implies that more experienced subjects state much lower WTA figures than inexperienced subjects, which induces the value divergence to lessen—super-intense consumers' WTA is approximately $17 less than light consumers WTA, while their WTP is generally only $2-$4 greater. One straightforward economic explanation of this result is that undergraduate students, who are typically used as subjects in laboratory experiments, are not likely to have been arbitraged in real markets like the experienced consumers in the sportscard and memorabilia market. A complementary psychology-based explanation is that

inexperienced subjects have dealt with selling and trading multiple objects in the past and

have learned to treat a good leaving their endowment as an opportunity cost rather than a

loss.[17]

Although analysis of the raw data provides evidence that supports the main

conjecture of the study, there has been no attempt to control for other factors that may

affect the bid/offer. These other subject-specific factors, which include years of trading

experience, gender, income, education, and age can be adequately accounted for in a well-

specified econometric model. To condition on these factors, I estimate the following

bid/offer model:

$$(6) \qquad\qquad Y = g(\alpha + \beta`X),$$

where Y is subject i's bid or offer; X includes subject-specific variables that may affect

the bid or offer. Variables in X are listed in Table 2 and include market intensity, years of

trading experience, yearly income, age, gender, education, and two dichotomous

treatment variables: (i) denoting whether the subject was bidding/offering for the Mint

stub and (ii) whether the subject was in the compensated treatments. Equation (1)

therefore provides a test of whether market intensity matters upon conditioning on several

potentially important variables.

Summary estimates of equation (6) are presented in Table 4. Since there were

signs of nonspherical disturbances, I use White's robust standard errors; and because

there is evidence that learning may be nonlinear, I include models that have market

intensity entering both monotonically and non-monotonically. Columns 1 and 2 in Table

---

[17] While these and a number of other explanations are undoubtedly plausible, it is beyond the scope of this particular study to parse out the various explanations of these findings. I will leave these important issues for future research, but return to them briefly in the concluding section.

4 provide insights into whether market intensity is correlated with WTA statements of value. Empirical results strongly support the unconditional findings in Table 3. For example, in the linear learning model, the coefficient estimate of intensity is significantly different from zero at the $p < .05$ level and implies that a one standard deviation increase in intensity ($\approx$ 4.2) is correlated with a $4.28 decrease in WTA. Coefficients in the nonlinear learning specification also suggest that market intensity is correlated with compensation demanded, yet here the coefficients suggest that there are diminishing returns to experience. The majority of other coefficient estimates are not significantly different from zero at conventional levels, except for the dummy variable *Mint*. In this case, both coefficient estimates are significant at the $p < .05$ level and indicate that subjects demanded about $4 more for Mint stubs than they demanded for Near Mint stubs.

Columns 3-6 contain empirical estimates from the WTP specifications, where in the Tobit models I present marginal effects computed at the sample means. While market intensity again enters in a statistically significant manner, its marginal influence is much lower than comparable estimates in the WTA specifications. For example, in the OLS linear model presented in column 3, an estimate of $0.21 suggests that WTP values are $0.84 larger for each one standard deviation increase in market intensity. This estimate represents about 20% of the marginal effect estimate in the WTA specifications. This result corresponds well with findings in Table 3, and further suggests that the convergence of values tends to occur for experienced subjects because of lower WTA values rather than significant increases in WTP. Consistent with the WTA specifications, there is some evidence that learning dissipates for intense consumers. In this case, the

learning curve reaches its peak at approximately 20 trades in a typical month, which is on the boundary of the sample (only one WTP subject stated he executed more than 20 trades in a typical month).[18]

Other WTP coefficient estimates are generally insignificant, but there are spots of significance. Interestingly, experience, as measured by the number of years in the sportscard market, gains statistical significance at the $p < .10$ level in each specification, but appears to be attenuating the bidding level. This effect also appears to have economic significance: one extra year of experience maps into roughly a $0.10 decrease in WTP, and a one standard deviation increase in years of market experience yields a $1.00 decrease in WTP. There is also evidence that men bid higher than women. While the OLS estimates are relatively noisy, in the Tobit specifications these effects are significant at conventional levels and imply that men bid about $1.75 more than women. This result, coupled with the fact that men provide lower WTA values, suggests that men have lower WTA/WTP disparities than women. This finding is potentially of interest, but due to small samples sizes of women I find the evidence premature to make a strong statistical statement concerning the role of gender on the WTA/WTP disparity.

IV. Concluding Remarks

Defining preferences over changes in consumption is a violation of the basic independence assumption, which is used in most theoretical and applied economic models to assess the operation of markets. Yet, substantial experimental evidence has mounted that indicates a nontrivial gap exists between WTA and WTP statements of value. These findings have induced even the most ardent supporters of neoclassical

---

[18] While the results herein do not discern between treatment and selection effects, List (2001b) presents evidence that suggests both play a role, but that treatment effects dominate.

theory to doubt the validity of certain neoclassical modeling assumptions. Now that experimental economists have obtained several unique insights on the value disparity, however, it is important to examine whether these experimental findings, which are generally from the laboratory, generalize to the environment that theorists' models purport to explain. Field experiments are a natural next step, as they provide a useful intermediary between the sterile lab environment and the vagaries of completely uncontrolled field data.

In this study, I examined data gathered from observing behavior in a naturally occurring marketplace. Even though economists have prided themselves on relevancy, few other studies have perturbed real markets for the sake of testing important economic theories.[19] Examining data from twelve experimental treatments yields several unique insights: across institutions that are substantially heterogeneous, reference dependent theory has good predictive power for many consumer types. Yet, for those consumers that have had a considerable amount of exchange opportunity, neoclassical theory predicts reasonably well, as I find sharp evidence that marketlike experience matters: across all consumer types intensity of market experience and the WTA/WTP disparity are negatively related. Accordingly, this paper argues that we may have prematurely discarded the neoclassical model, as the results herein suggest that the observed discrepancy in WTA and WTP values may merely be a mistake made by inexperienced consumers. Further, super-experienced consumers' behavior is in broad accord with predictions of the neoclassical model.

---

[19] But, see for example the recent work of List (2001a) and List and Lucking-Reiley (2001) who use field experiments to examine other issues.

A few clear next steps in this research agenda are to bolster sample sizes of the experienced consumer subset and examine whether these results spill-over to the domain of public goods. This latter extension is necessary since the disparity has such profound impacts on applied welfare issues such as nonmarket valuation and more broadly cost/benefit analysis. Besides these obvious extensions, it makes sense to investigate the underlying causal mechanism at work. For example, why do experienced subjects display no value disparity? Is it simply that the market has arbitraged them, and via learning they become rational utility maximizing agents? If so, what have they learned? Is it to treat the loss as an opportunity cost? Or are they more certain of their preferences because they are experienced consumers? Could lower transactions costs play a role?

While I find answers to these provocative questions quite important, I defer a more extensive discussion until another occasion. Yet, I would be remiss not to briefly discuss a few of the possible explanations. In terms of neoclassical explanations, preference uncertainty seems to be the logical place to start. Indeed, the overall pattern of observed results herein is consistent with Kolstad and Guzman's (1999) theoretic bidding model in which subjects have private values, but know them imperfectly. In their model, an increase in the level of individual information decreases the gap between WTP and WTA. In addition, as the cost of information increases, the expected gap between WTP and WTA increases. Coupling the intuition of these two insights yields their Proposition 7 (p. 77), which aptly describes my experimental findings: "In a private value, costly information rational expectations equilibrium, with information costs such that there are informed and uninformed bidders, a mean-increasing shift in the distribution of information costs results in a decrease in the number of potential trades."

Relatedly, Kling et al. (2002) have recently obtained promising experimental results suggesting that a neoclassical options based theory may have some predictive power. Using a good that is commonly traded in the sportscard market, they find that even for those consumers that plan to purchase the good for resale, WTP (WTA) increases (decreases) in the difficulty of delay and decreases (increases) in the difficulty of reversing the transaction. With certain reasonable assumptions, their comparative static findings are consistent with the results herein. Besides more fully exploring these neoclassical-based explanations, I believe the psychology literature can also lend insights into the effect of experience on the value disparity. For example, George Loewenstein and associates have recently found that experience may lessen the endowment effect by lowering the "pain" of the transaction. In this sense, agents are more likely to treat the good leaving their endowment as an opportunity cost rather than a loss.

References

Bateman, Ian, Munro, Alistair, Rhodes, Bruce, Starmer, Chris, and Sugden, Robert. "A Theory of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 1997, 112, pp. 479-505.

Becker, G., DeGroot, M., and Marschak, J. "Measuring Utility by a Single Response Sequential Method." *Behavioral Science*, July 1964, 9, pp. 226-36.

Brookshire, David and Coursey, Don. "Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures," *American Economic Review,* September 1987, 77(4), pp. 554-566.

Brookshire, David, Randall, Alan, and Stoll, John R. "Value Increments and Decrements in Natural Resource Service Flows," *American Journal of Agricultural Economics* 1980, 62(3), pp. 185-193.

Coursey, Don, Hovis, John, and Schulze, William. "The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value." *Quarterly Journal of Economics*, 1987, 102(3), pp. 679-90.

Horowitz, John K. and McConnell, K.E. "A Review of WTA/WTP Studies," *Journal of Environmental Economics and Management*, forthcoming, 2002a.

Horowitz, John K. and McConnell, K.E. "Willingness to Accept, Willingness to Pay, and the Income Effect," *Journal of Economic Behavior and Organization*, forthcoming, 2002b.

Kahneman, Daniel, Knetsch, Jack L., and Thaler, Richard H. "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy,* December 1990, 98(6), pp. 1325-48.

Kling, Catherine, List, John A., and Zhao, Jinhua. "Commitment Costs and the Basic Independence Assumption: Evidence from the Field," Working paper, Iowa State University, 2002.

Knetsch, Jack L. "The Endowment Effect and Evidence of Nonreversible Indifference Curves," *American Economic Review,* December 1989, 79(5), pp. 1277-1284.

——"Environmental-Policy Implications of Disparities Between Willingness To Pay and Compensation Demanded Measures of Values," *Journal of Environmental Economics and Management*, 1990, 18 (3): 227-237.

Knetsch, Jack L. and J.A. Sinden. "The Persistence of Evaluation Disparities," *Quarterly Journal of Economics*, 1987, 102(3), pp. 691-695.

Knetsch, Jack L., Fang-Fang Tang, and Thaler, Richard H. "The Endowment Effect and repeated Market Trials: Is the Vickrey Auction Demand Revealing?" *Experimental Economics*, December 2001, 4(3), pp. 257-269.

Knez, Peter, Smith, Vernon L., and Williams, Arlington. "Individual Rationality, Market Rationality, and Value Estimation," *American Economic Review*, May 1985, 75(2), pp. 397-402.

Kolstad, Charles and Guzman, Rolando. "Information and the Divergence between Willingness to Accept and Willingness to Pay," *Journal of Environmental Economics and Management*," July 1999, 38(1), pp. 66-80.

List, John A. "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards**,"** *American Economic Review,* December, 91[5], 2001a, 1498-1507.

——"The Endowment Effect: True Preferences or Mutable Anomaly? Evidence from the Field," Working paper, University of Maryland, 2001b.

List, John A. and Lucking-Reiley, David. "Demand Reduction in Multi-Unit Auctions: Evidence from a Sportscard Field Experiment," *American Economic Review,* September 2000, 90(4), pp. 961-972.

List, John A. and Shogren, Jason F. "The Deadweight Loss of Christmas: Comment," *American Economic Review*, December 1998, 88(5), pp. 1350-1355.

Morrison, G.C. "Resolving Differences in Willingness to Pay and Willingness to Accept: Comment," *American Economic Review,* March 1997, 87(1), pp. 236-40.

Plott, Charles. "Laboratory Experiments in Economics: The Implication of Posted-Price Institutions." *Science*, May 1986, 232, pp. 732-38.

Randall, Alan and Stoll, John R. "Consumer's Surplus in Commodity Space." *American Economic Review,* June 1980, 71(3), pp. 449-57.

Shogren, Jason F., Shin, Seung Y., Hayes, Dermot J., and Kliebenstein, James B. "Resolving Differences in Willingness to Pay and Willingness to Accept." *American Economic Review,* March 1994, 84(1), pp. 255-70.

Shogren, J. Cho, S., Koo, C. List, J., Park, C., Polo, P. Wilhelmi, R. "Auction Institutions and the Measurement of WTP and WTA," *Resource and Energy Economics* April, 23(2) pp. 97-109 (2001).

Smith, Vernon L. "Microeconomic Systems as an Experimental Science." *American Economic Review*, December 1982, 72(5), pp. 589-97.

Sugden, Robert. "Alternatives to the Neoclassical Theory of Choice," in Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries (Ian Bateman and K.G. Willis, eds.), Oxford University Press, 1999.

Zhao, Jinhua and Kling, Catherine. "Real Options and the WTA/WTP Disparity," Working paper, Iowa State University, 1999.

Table 1 Experimental Design

Stub Quality	WTA	WTP (uncompensated)	WTP (compensated)
Near Mint	Denoted WTA1 Random *n*th: n =30 BDM: n =30	Denoted WTPU1 Random *n*th: n =30 BDM: n =30	Denoted WTPC1 Random *n*th: n =30 BDM: n =30
Mint	Denoted WTA2 Random *n*th: n =30 BDM: n =30	Denoted WTPU2 Random *n*th: n =30 BDM: n =30	Denoted WTPC2 Random *n*th: n =30 BDM: n =30

Notes: Each cell represents two unique treatments. For example, "Random *n*th: n =30" in row 1, column 1 denotes that one treatment had 30 offerers competing in a random *n*th price auction to sell their endowed good—a Near Mint Ripken ticket stub. The treatment is labeled WTA1, for WTA for the Near Mint stub. "BDM" denotes Becker-DeGroot-Marschak (1964) discrete auction mechanism. No subject participated in more than one treatment.

Table 2 Sample Statistics

	WTA1 Mean (Std. Dev.)	WTA2 Mean (Std. Dev.)	WTPU1 Mean (Std. Dev.)	WTPU2 Mean (Std. Dev.)	WTPC1 Mean (Std. Dev.)	WTPC2 Mean (Std. Dev.)
Random nth						
Bid or Offer	13.53	16.78	4.32	4.90	4.25	5.98
	(9.2)	(12.0)	(4.85)	(4.57)	(4.61)	(6.12)
Trading Experience	4.3	5.9	3.8	4.9	4.3	6.0
	(3.4)	(3.9)	(3.7)	(3.2)	(4.6)	(6.1)
Years of Market Experience	6.9	6.9	8.7	14.3	10.1	12.6
	(9.2)	(6.4)	(8.3)	(14.3)	(13.5)	(13.1)
Income	3.3	3.4	3.4	4.6	4.0	3.7
	(1.9)	(1.9)	(1.9)	(1.9)	(1.9)	(2.2)
Age	31.9	30.9	31.1	35.6	35.7	28.8
	(14.9)	(10.6)	(12.7)	(12.7)	(15.4)	(11.6)
Gender (%male)	0.89	0.90	0.88	0.86	0.86	0.72
	(0.32)	(0.31)	(0.32)	(0.35)	(0.35)	(0.45)
Education	2.8	3.4	3.2	3.9	3.5	3.5
	(1.7)	(1.7)	(1.5)	(1.4)	(1.8)	(1.7)
BDM						
WTA	11.98	13.47	5.85	6.67	5.53	7.78
	(11.6)	(12.1)	(5.1)	(8.7)	(5.9)	(6.7)
Trading Experience	4.4	4.8	5.6	5.2	3.5	4.1
	(4.5)	(5.3)	(4.8)	(5.1)	(3.3)	(3.7)
Years of Market Experience	10.7	9.5	8.4	9.5	8.5	7.9
	(11.1)	(10.5)	(6.3)	(12.7)	(8.6)	(9.2)
Income	3.9	3.5	3.8	3.5	3.5	3.4
	(2.4)	(2.1)	(2.3)	(2.1)	(2.0)	(1.9)
Age	33.8	30.4	32.1	30.3	29.9	34.0
	(14.1)	(13.1)	(14.6)	(14.0)	(11.8)	(14.9)
Gender (%male)	0.86	0.78	0.79	0.88	0.86	0.87
	(0.35)	(0.42)	(0.42)	(0.33)	(0.35)	(0.35)
Education	3.8	3.4	3.7	3.1	2.6	3.0
	(1.8)	(1.7)	(1.7)	(1.6)	(1.5)	(1.5)

Notes:
1.   Thirty subjects are included in each of the twelve treatments, for a total of 360 subjects.

2. *Trading experience* represents the number of trades made in a typical month.
3. *Years of market experience* denotes years that the subject has been active in the market.
4. *Income* denotes categorical variable (1-8): 1) Less than $10,000, 2) $10,000 to $19,999, 3) $20,000 to $29,999, 4) $30,000 to $39,999, 5) $40,000 to $49,999, 6) $50,000 to $74,999, 7) $75,000 to $99,999 8) $100,000 or over.
5. *Age* denotes actual age in years.
6. *Gender* denotes categorical variable: 0 if female, 1 if male.
7. *Education* denotes categorical variable 1) Eighth grade or less, 2) High School 3) 2-Year College, 4) Other Post-High School, 5) 4-Year College, 6) Graduate School Education.

Table 3 Bid/Offer Sample Statistics Across Consumer Types

	WTA1 Mean (Std. Dev.)	WTA2 Mean (Std. Dev.)	WTPU1 Mean (Std. Dev.)	WTPU2 Mean (Std. Dev.)	WTPC1 Mean (Std. Dev.)	WTPC2 Mean (Std. Dev.)
Pooled						
Bid or Offer	12.76	15.12	5.08	5.78	4.89	6.88
	(10.4)	(12.0)	(4.9)	(7.0)	(5.3)	(6.4)
	(n=60)	(n=60)	(n=60)	(n=60)	(n=60)	(n=60)
"Light" Consumers						
Bid or Offer	23.40	22.21	3.34	2.75	4.75	5.06
	(14.7)	(13.7)	(5.6)	(3.8)	(7.2)	(9.9)
	(n=10)	(n=7)	(n=13)	(n=6)	(n=12)	(n=12)
"Casual" Consumers						
Bid or Offer	16.31	18.11	4.48	6.21	3.50	4.51
	(11.6)	(14.0)	(5.7)	(9.7)	(4.7)	(6.6)
	(n=35)	(n=28)	(n=30)	(n=24)	(n=34)	(n=32)
"Intense" Consumers						
Bid or Offer	7.78	12.51	5.68	5.50	6.71	9.59
	(5.7)	(9.4)	(3.5)	(4.4)	(5.6)	(5.2)
	(n=25)	(n=32)	(n=30)	(n=36)	(n=26)	(n=28)
"Super-Intense" Consumers						
Bid or Offer	6.71	8.33	5.67	7.17	6.17	7.65
	(5.7)	(8.1)	(4.1)	(5.3)	(4.1)	(4.9)
	(n=12)	(n=15)	(n=12)	(n=12)	(n=9)	(n=10)

Notes: Data are pooled across BDM and random nth price auctions in each major treatment cell. "Light" consumers are those that trade 0 times in a typical month; "Casual" consumers are those that trade less than 5 times in a typical month; "Intense" consumers are those that trade 5 or more times in a typical month; "Super-Intense" consumers are those that trade 9 or more times per month. In choosing these thresholds, I considered both the mean number of trades in a typical month (4.7) and its standard deviation (4.2). Thus, "Casual" consumers are those subjects below the average trading rate; "Intense" consumers are those subjects above the average trading rate, and "Super-Intense" consumers are those consumers at least one standard deviation above the average trading rate.

Table 4 Regression Results

	WTA (OLS)	WTA (OLS)	WTP (OLS)	WTP (Tobit)	WTP (OLS)	WTP (Tobit)
Intensity	-1.07**	-2.04**	0.21**	0.17**	0.39**	0.49**
	(0.28)	(0.65)	(0.08)	(0.07)	(0.20)	(0.17)
Intensity^2	---	0.07*	---	---	-0.01	-0.01**
		(0.04)			(0.09)	(0.007)
Mint	4.1**	3.9**	1.08	0.59	1.09	0.75
	(2.0)	(1.9)	(0.83)	(0.76)	(0.83)	(0.75)
Compensated	---	---	0.24	0.50	0.25	0.37
			(0.81)	(0.76)	(0.31)	(0.73)
Years of Market Experience	0.12	0.10	-0.07*	-0.08**	-0.07*	-0.09**
	(0.09)	(0.8)	(0.04)	(0.03)	(0.04)	(0.03)
Income	-0.33	-0.35	-0.23	-0.03	-0.24	-0.03
	(0.50)	(0.49)	(0.21)	(0.22)	(0.22)	(0.21)
Gender (%male)	-0.67	-0.52	1.43	1.62*	1.44	1.83**
	(3.38)	(3.47)	(1.02)	(0.87)	(1.0)	(0.81)
Age	0.05	0.08	0.03	0.01	0.03	0.01
	(0.09)	(0.09)	(0.04)	(0.03)	(0.04)	(0.03)
Constant	12.2**	15.2**	2.32	1.21	1.96	-0.25
	(5.0)	(4.6)	(2.02)	(1.83)	(2.01)	(1.8)
N	112	112	215	215	215	215
R^2	0.17	0.19	0.04	---	0.05	---

Notes:
1. Dependent variable is subjects offer or bid.
2. Mint equals 1 if subject was in the treatment with Mint stubs, 0 otherwise. Compensated equals 1 if subject was in a compensated treatment, 0 otherwise. Gender = 1 if male, 0 otherwise.
3. Tobit coefficients are marginal effects computed at the sample means. Standard errors are in parentheses beside coefficient estimates.
4. Sample sizes may not match with sample sizes in Table 1 due to some respondents not filling in the entire questionnaire (e.g., not including income responses). Reported results omit these observations. If means are used to fill in the missing observations, results are not quantitatively or qualitatively different from the results reported.
5. **(*)Denotes coefficient estimate is significant at the p < .05 (.10) level.
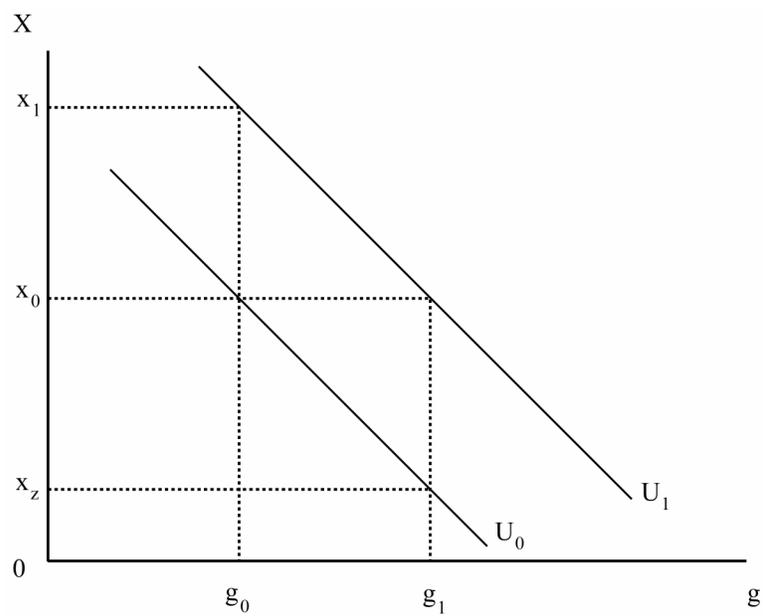
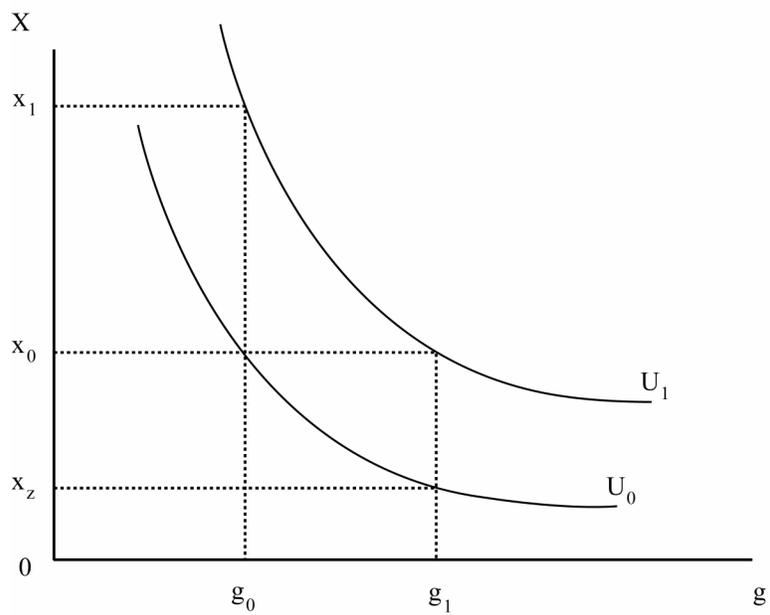Figure 1.   Linear Indifference Curves and WTA/WTP

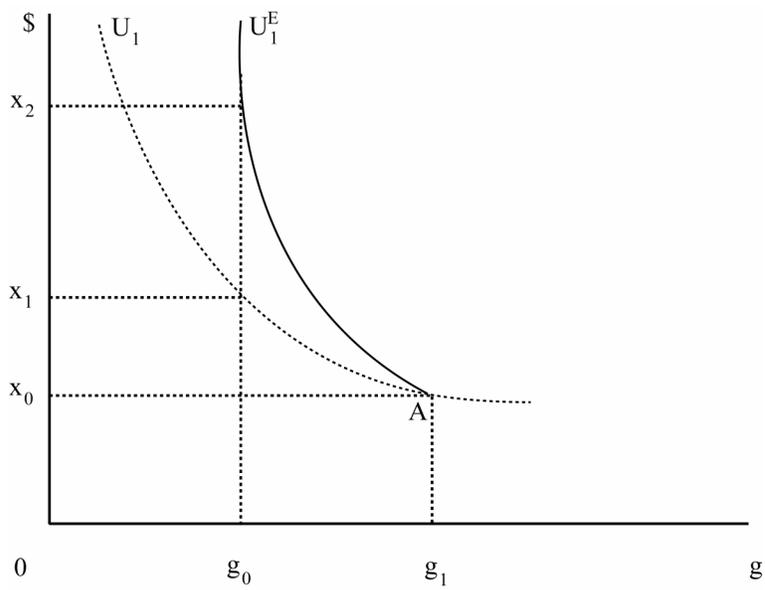Figure 2.   Convex to the Origin Indifference Curves and WTA/WTP

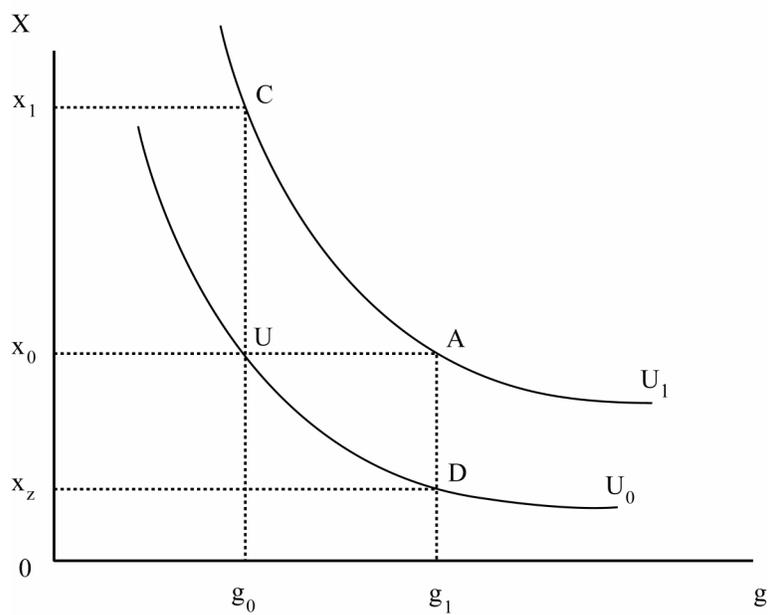Figure 3.   Endowment Effects and the WTA/WTP Disparity

Figure 4.   Experimental Design

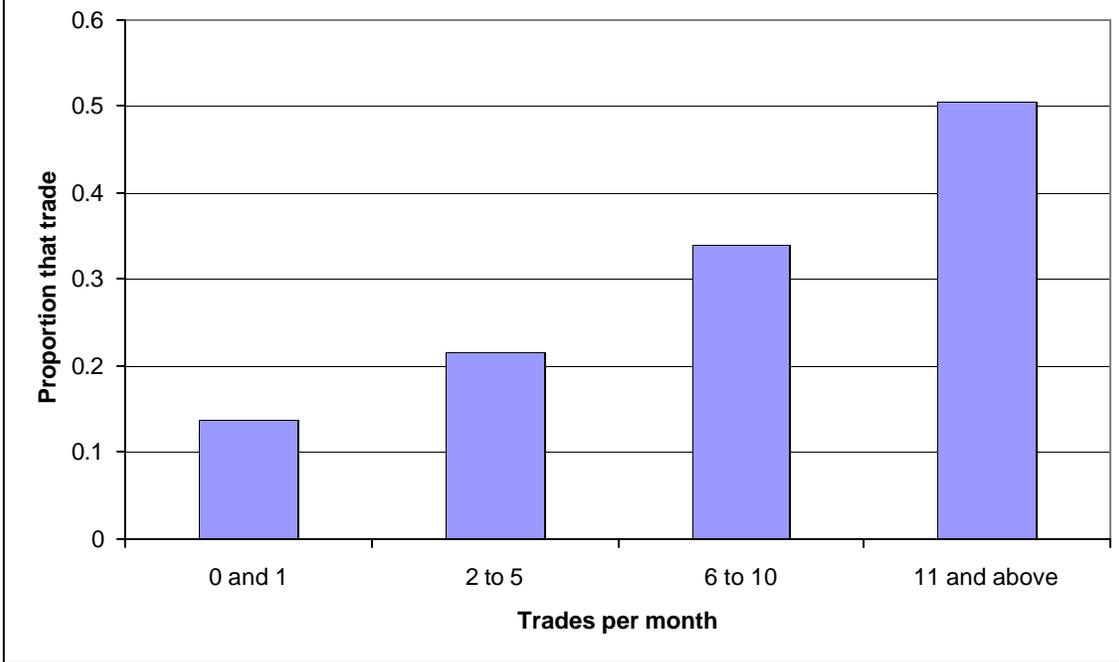**Figure 5. Summary of List's (2001) Trading Results**
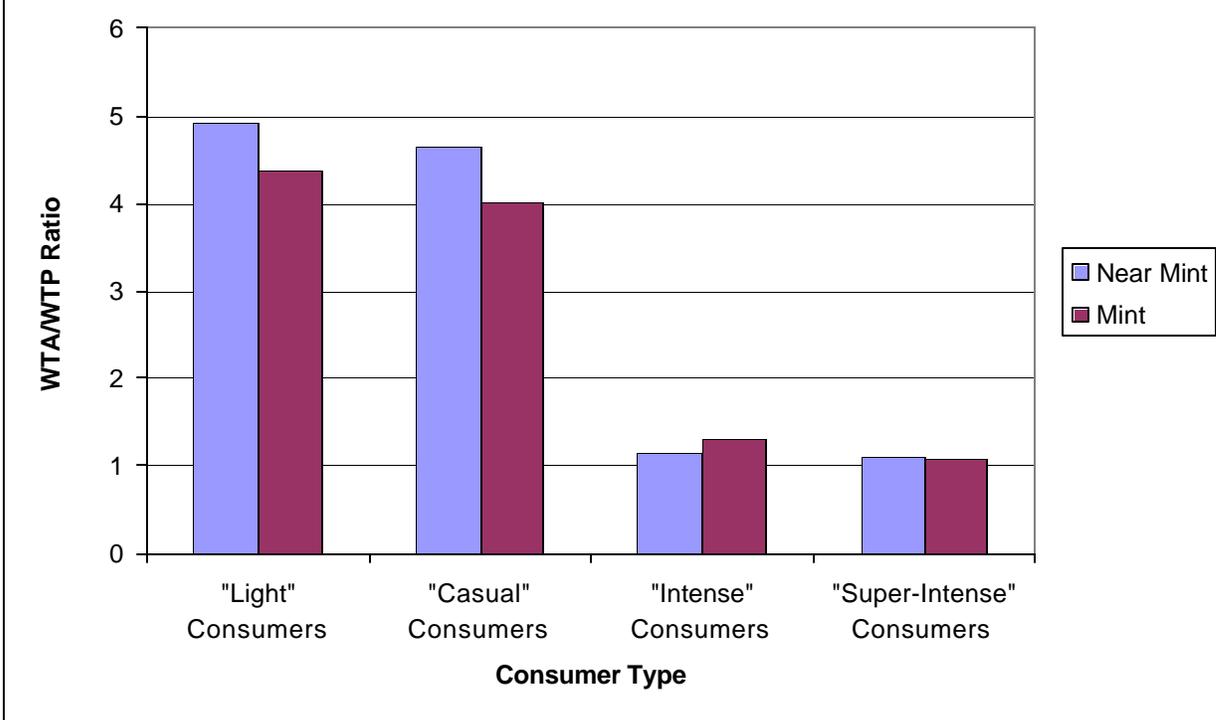

Figure 6. Summary of WTA/WTP Ratios

Appendix A Survey

These questions will be used for statistical purposes only. THIS INFORMATION WILL BE KEPT STRICTLY CONFIDENTIAL AND WILL BE DESTROYED UPON COMPLETION OF THE STUDY.

1. How long have you been active in the sportscards and memorabilia market?
   _____yrs

2. Approximately how many trades (cards or memorabilia) do you make in a typical month? _____. Note that trades could include pokemon cards, sportscards, other trading cards, and sports memorabilia.

3. Are you a sportscard or sports memorabilia professional dealer?_____

4. Gender: 1) Male   2) Female

5. Age _____   Date of Birth _____

6. What is the highest grade of education that you have completed. (Circle one)
   1) Eighth grade   3) 2-Year College   5) 4-Year College
   2) High School   4) Other Post-High School   6) Graduate School Education

7. What is your approximate yearly income from all sources, before taxes?
   1) Less than $10,000   5) $40,000 to $49,999
   2) $10,000 to $19,999   6) $50,000 to $74,999
   3) $20,000 to $29,999   7) $75,000 to $99,999
   4) $30,000 to $39,999   8) $100,000 or over

Appendix B Subject instructions for WTP random *n*th price auction.

Welcome to Lister's Auctions. You have the opportunity to bid in an auction for the good on the table. There are only 30 bidders in this auction; the other 29 bidders will be randomly chosen from other participants at today's card show.

Auction Rules:

You are asked to submit one bid. Since there are 29 other bidders, there will be a total of 30 bids submitted. The monitor will rank these 30 bids from highest to lowest and the winning bidder(s) will be determined in a random fashion. Here is how it works: if the monitor randomly selected the bid ranked #20 (the 20th highest bid), then each of the 19 bidders who bid more than this bid would win in the auction and receive the good after they sent me the value of the 20th highest bid. There is an equal chance that the selected bid will be the 2nd, 3rd, 4th, 5th, 6th, …. or 30th highest bid. Lets go through a few examples so you understand the auction rules.

Assume that the randomly determined bid is the 8th highest. I will rank the bids from highest to lowest to determine the winners.

> **Example 1:** if the 10 highest bids are ranked highest to lowest as follows:
>
> | | |
> |-------|-----------------------------|
> | $C | High bidder |
> | $D | 2nd highest bidder |
> | $A | 3 |
> | $B | 4 |
> | $F | 5 |
> | $G | 6 |
> | $L | 7 |
> | $K | 8 |
> | $V | 9 |
> | $Z | 10th highest bidder |

Since the 8th highest bid was chosen, the top 7 bidders win and pay the value of the 8th highest bid. In this case, that is $K.

Would you like to go through some more examples?

As you can see, in this type of auction you should bid your true value for the good (e.g., what it is worth to you). If you bid too much, then you stand a risk of winning but paying more than what it is worth to you. If you bid below your true value, then you stand a risk of not winning the good when you could have won and paid less than your true value. This is true because your bid never affects the price that you pay in the auction, just whether you win or lose. In this way, it is similar to sealed bid 2nd price auctions that you see sometimes at card shows or in trade magazines.

Would you like to go through some examples that illustrate why bidding your true value makes sense?

Final Transaction

At 6PM Sunday night I will determine the winners of the auction. After the winners pay me (cash or check) for the good, the good will be awarded to the winners. Note, regardless of price, the goods will be awarded to the winners. In case you cannot attend the "determination of winners" session at 6pm, please provide your name, mailing address, and phone number below:

Name_____

Address_____

Phone#_____

If you are unable to attend at 6PM, I will contact you by phone. Upon receipt of your check or cash, I will send you the good that you won. All postage will be paid by Lister's Auctions for goods mailed to winners.

I guarantee to sell the goods to the winners no matter what the final auction prices turn out to be. Your bid represents a binding commitment to buy the good you win at the prices specified by the auction outcomes.

Good luck—please write your bids on the sheets provided.

Thanks for participating.