

Endowment Effects and Drinking Water Quality

Pamela L. Booth, Todd Guilfoos, and Emi Uchida

We conduct a laboratory experiment to test for the existence of the endowment effect—a gap between willingness to accept and willingness to pay—for improved drinking water quality using a within-subject design. We find a statistically significant and positive gap. Willingness to accept is 62 to 125 cents higher than willingness to pay on average, indicating the presence of endowment effects. This gap is robust to information about the quality of water being consumed. We also identify some heterogeneity in the size of the gaps that is associated with differences in subjects' knowledge of drinking water quality and disparities in their incomes.

Key Words: drinking water quality, endowment effect, experimental economics

The quality of drinking water in the United States is high thanks to stringent federal standards and public demand. Municipalities must decide how to provide high-quality drinking water that meets public expectations, which may exceed standards set by the Safe Drinking Water Act. At the same time, there will be an estimated countrywide investment shortfall in drinking water infrastructure of \$72 billion over the next twenty years (Environmental Protection Agency (EPA) 2013). City planners are exploring alternative investment strategies that will allow cities to preserve the quality of their water supplies by protecting watersheds and conserving source water through projects that offer large ecological and social benefits (Pires 2004, Postel and Thompson 2005). Pivotal to evaluations of alternative projects and investments is how improvements to the quality of drinking

Pamela Booth is a doctoral student, Todd Guilfoos is an assistant professor, and Emi Uchida is an associate professor in the Department of Environmental and Natural Resource Economics at University of Rhode Island. *Correspondence:* Todd Guilfoos ▪ Department of Environmental and Natural Resource Economics ▪ Kingston Coastal Institute ▪ 1 Greenhouse Road ▪ University of Rhode Island ▪ Kingston, RI 02881 ▪ Phone 1.401.874.2471 ▪ Email guilfoos@uri.edu.

This material is based on work supported in part by the College of the Environment and Life Sciences for graduate research assistantship and by the National Science Foundation EPSCoR Track-2 Cooperative Agreement IIA-1330406, Collaborative Research: North East Water Resources Network. The authors are grateful to two anonymous reviewers for valuable input and comments. We also thank Maik Kecinski, Jacob Fooks, Carrie Gill, and the participants of the Northeastern Agricultural and Resource Economics Association Water Quality Economics Workshop for helpful comments and Haoran Miao and Christopher Brozyna at University of Rhode Island for coding assistance. The views expressed are the authors' and do not necessarily represent the policies or views of any sponsoring agencies.

Agricultural and Resource Economics Review 45/2 (August 2016) 338–366

© The Author(s) 2016. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

water are valued, especially when considering improvements that will generate water that will exceed the basic safety standards.

We conduct a laboratory experiment in which we test for the existence of the endowment effect—additional value attached to a good because of possession of it, which translates to a gap between willingness to accept (WTA) and willingness to pay (WTP)—for improved drinking water quality. Many studies have examined the valuation of drinking water quality or increased reliability of service (Brox, Kumar, and Stollery 2003, Jordan and Elnagheeb 1993, McConnell and Rosado 2000, Whittington et al. 1990). However, those studies typically elicited WTP for improvements and largely ignored the role of endowments in valuations of drinking water quality.

There is reason to believe that the endowment effect applies not only to goods but also to the degree of quality of goods and that a single commodity provided at different levels of quality can be viewed essentially as separate goods (Corrigan and Rousu 2006, Lusk, Feldkamp, and Schroeder 2004, MacDonald and Bowker 1994). Whenever an endowment effect exists, there can be gaps between WTP for a good and WTA an offer to give up that good, in this case high-quality drinking water. A significant gap between WTA and WTP for drinking water quality would suggest, for example, that municipalities should aim to provide drinking water that exceeds the quality standards set by regulatory authorities and invest accordingly while projects that would improve water quality further are not highly valued by consumers.

Numerous studies have examined gaps between WTA and WTP, potential drivers of those gaps, and the potential policy implications of this behavioral anomaly (e.g., Brown and Gregory 1999, Del Saz-Salazar, Hernandez-Sancho, and Sala-Garrido 2009, Hjorth and Fosgerau 2011, MacDonald, Morrison, and Barnes 2010, Rowe, D'Arge, and Brookshire 1980). The gap has been identified in both experimental and nonexperimental studies for many goods (Horowitz and McConnell 2002, Tunçel and Hammitt 2014), and in general, the magnitudes of the gaps appear to depend on the goods' characteristics with nonmarket goods exhibiting the largest gaps (Amigues et al. 2002, Brown and Gregory 1999, Horowitz and McConnell 2002). In many cases, the gaps can be explained by large discrete changes in the goods (Randall and Stoll 1980) or by a lack of market substitutes (Hanemann 1991). A number of empirical studies (e.g., Morrison 1998, Shogren et al. 1994) and meta-analyses (Horowitz and McConnell 2002, Sayman and Öncüler 2005, Tunçel and Hammitt 2014) have identified large WTA-WTP gaps for goods with incomplete substitutes. Another source of gaps is uncertainty, including lack of familiarity with the good and time preferences (Kling, List, and Zhao 2013, Sayman and Öncüler 2005, Tsur 2008, Zhao and Kling 2001).

Experiment designs and elicitation techniques can also influence the size of the gaps measured in a study (e.g., Plott and Zeiler 2005, 2007), and their persistence can be explained by loss aversion and a behavioral propensity to place a larger value on losses (WTA) than on gains (WTP) (Tversky and Kahneman 1991, Knetsch and Sinden 1984). This study contributes to the

literature by testing for the presence of an endowment effect in drinking water quality using a laboratory experiment that elicits valuations of various levels of water quality.

In the experiment, we construct a competitive market for drinking water quality using filtered and unfiltered tap water and use information treatments about the waters' quality. The experiment allows subjects to trade tap water of different qualities by endowing buyers with unfiltered tap water and sellers with filtered tap water. Though drinking water quality is sometimes regarded as a public environmental good, we treat it as an attribute or characteristic of a private good.

We find significant differences between WTA and WTP for drinking water quality. The ratio of average WTA to average WTP ranges from 2.72 to 10.77, which is consistent with results from studies of many other types of goods (e.g., Tuncel and Hammitt 2014). We also find that the gap persists but is narrower after subjects are provided with information about the quality of tap water and benefits of filtered water. These findings suggest that drinking water quality is influenced by endowment effects.

A Model of Endowment Effects with Environmental Quality

We use a reference-dependent model in which the reference point is defined as in Tversky and Kahneman (1991) as

$$(1) \quad U_r(x) = U(R(x))$$

where R is a reference function for a good, x , that is associated with a reference point, r . We extend the reference-dependent utility framework so that the reference function depends on a composite good, x , and its quality, q . Alternatively, we could interpret the reference point to encompass quality characteristics that extend to multiple dimensions. We specify the reference point in our model as an attribute of a good with a kink in utility occurring where $q = r$:

$$(2) \quad U_r(x, q) = U(R(x, q))$$

This model follows Tversky and Kahneman (1991); the reference point is specified in terms of the attribute of the good rather than the good itself. An increase in a good's quality increases consumers' utility, which leads to greater WTP for the good. We do not claim that this is a true extension of the original model (which did not exclude this interpretation). Rather, this model provides a brief introduction to how to think about the endowment effect for attributes of goods.

Hypotheses

We use a within-subject design to elicit WTP and WTA from each subject. We define WTP as the amount that a subject who is endowed with low-quality drinking water is willing to pay to receive higher-quality water and WTA as the amount that a subject who is endowed with high-quality drinking water is willing to accept to receive lower-quality water. This design provides data we can use to test the following hypotheses by comparing individuals' WTA and WTP.

Hypothesis 1: Valuations of drinking water quality are not influenced by the endowment effect.

$$H_0: WTA = WTP$$

This first hypothesis is tested by determining whether there is a gap between a subject's WTA and WTP. Since the endowment effect is well established for a number of market and nonmarket goods (Tunçel and Hammitt 2014), we expect that the null hypothesis will be rejected and average WTA will exceed average WTP.

Hypothesis 2: Information does not affect the endowment effect for drinking water quality.

$$H_0: WTA_A - WTP_A = WTA_B - WTP_B$$

The second hypothesis examines the robustness of the results of the first hypothesis by testing whether WTA-WTP gaps are caused, at least in part, by lack of familiarity with the good being valued and can be mitigated by providing information that improves the consumers' familiarity with the good (List 2003). We construct an information treatment that is presented to all subjects to improve their understanding of water quality and posit that the information will decrease the gap. We expect to reject the null hypothesis and to find, instead, that the pre-information gap in valuations ($WTA_A - WTP_A$) is larger than the post-information gap in valuations ($WTA_B - WTP_B$).

Hypothesis 3: The initial endowment of drinking water quality does not influence subsequent valuations.

$$H_0: WTA_C = WTA_D; H_0: WTP_C = WTP_D$$

This hypothesis tests the market role of the first round in the experiment on subsequent bids to buy and asks to sell. Subjects may use their initial market role as a seller or a buyer as a reference point and make choices based on that reference point throughout the remainder of the experiment (Kahneman, Knetsch, and Thaler 1991). We posit that subjects who were initially endowed with a high quality of water (sellers, designated with subscript C) will make choices that maintain the quality of their water to a greater degree

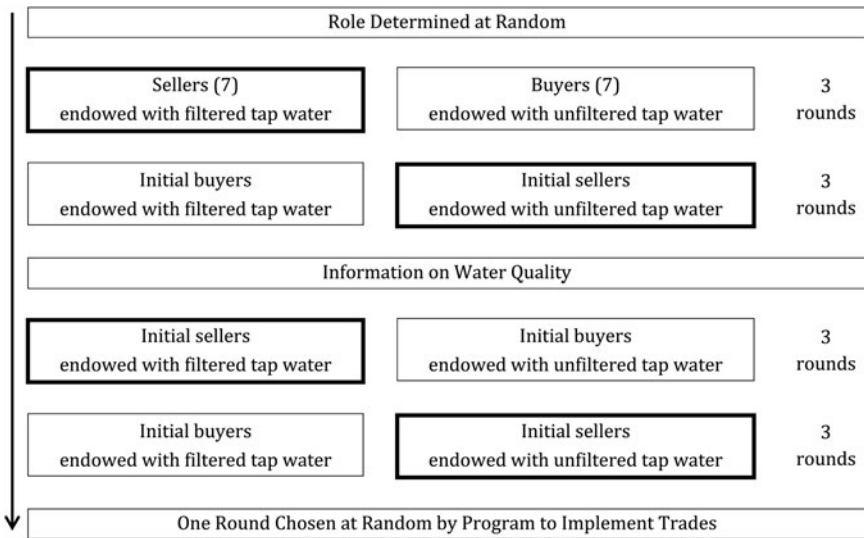


Figure 1. Water Round Flowchart

Note: Market roles were assigned randomly; sellers and buyers played as their starting roles for three rounds and then switched roles and played for three additional rounds. Every subject then received information stating that the United States had both high standards for drinking water and high-quality drinking water. The program chose one round at random from the twelve rounds played to implement as trades. Bolded boxes follow the path through the market of subjects initially endowed with high-quality water.

than subjects who were initially endowed with low-quality water (buyers, designated with subscript D) regardless of their roles in the later rounds.

Experiment Design

In the experimental market for trading high-quality and low-quality drinking water, we use a two-to-the-two factorial in a competitive market framework to identify endowment effects. The market consists of an equal number of buyers and sellers with subjects randomly assigned to each role (see Figure 1). Buyers are endowed with the low-quality good and sellers are endowed with the high-quality good.

The subjects had an opportunity to become familiar with the market mechanism through several practice rounds involving a known good, candy, before beginning the formal experiment. During the experiment, each subject maintained the role initially assigned for the first three rounds and then switched roles for the next three rounds. This pattern—buyer-seller or seller-buyer—was the first treatment and allowed us to conduct a within-subject comparison of WTA and WTP. After the first six rounds were completed, the administrator provided the information treatment, which described standards

for drinking water quality in the United States and the effects of the technology used to filter the tap water in the experiment. The buyer-seller/seller-buyer treatment was then repeated for an additional six rounds.¹

The quality market is structured like a double auction in which buyers are endowed with a low-quality good and sellers are endowed with a high-quality good. In each round, the subject first chooses whether to participate in the market. Buyers who choose to participate submit one bid (their WTP) and sellers who choose to participate submit one ask (their WTA). The market-clearing price is calculated from the intersection of the demand and supply schedules using all bids and asks submitted. Trades and transactions occur at this market-clearing price and are dependent on meeting two conditions: (i) a buyer's bid is equal to or greater than the market price or a seller's ask is equal to or less than the market price and (ii) there are an equal number of buyers and sellers who meet the first condition. If both conditions are met, an individual can trade in that round. If only the first is met, only the highest bidding buyer(s) and lowest bidding seller(s) will actually trade in that round; individuals who do not meet the first condition cannot trade. When the demand and supply curves do not cross at any price point, no one trades when supply (WTA) exceeds demand (WTP) and everyone trades when demand exceeds supply.

Experimental Procedure

Five sessions, each involving fourteen subjects, were conducted at an experimental computer lab in spring of 2015. The subject pool consisted of 70 undergraduate (82.9 percent) and graduate (17.1 percent) students, 55.7 percent of whom were from environment- and science-oriented fields. The subjects were recruited through university-sponsored e-mails, class announcements, and fliers, and they were prescreened to prevent anyone from participating more than once.

The experiment was programmed and conducted in z-Tree software (Fischbacher 2007). Subjects were randomly assigned to each computer, which was randomly designated as a buyer or seller in the market. The computers were equipped with privacy shields, and subjects could view the item they were valuing from a seated position. The bids and asks were submitted anonymously through the software interface via a code identifying each subject. Subjects were provided information regarding the market-clearing price and their eligibility to trade for each round but were not informed of the decisions or bids/asks made by others. They each received an account balance of \$10 to use during the experiment and were told that their final earnings would consist of the balance remaining in the account

¹ Within-subject designs and repeat experiences have been used in a handful of WTP-WTA studies (e.g., Harless 1989, Schmidt and Traub 2009).

plus \$5 for participating in the experiment. Sessions lasted approximately one hour and subjects earned an average of \$10.76 plus the goods (candy and water) and the \$5 participation payment.²

Each session involved fifteen rounds—three practice rounds using candy and twelve rounds using drinking water. In each round, subjects made two decisions: whether to participate in the market and how much to bid/ask. Those who chose to participate in a round were told that they had to remain within the budget (the account balance) when submitting bids/asks. The market price was calculated at the end of each round using all bids/asks submitted and was independent of the results of previous rounds. In addition, subjects were told that a market price of \$0 (i.e., all asks exceeded all bids) meant that no transactions would occur in that round.

Once the subjects were assigned to a computer, they were given information about the market structure and then participated in the three practice rounds.³ In the practice rounds, buyers were endowed with one piece of a generic milk chocolate valued at less than 10 cents, sellers were endowed with one Hershey's milk chocolate kiss valued at less than 10 cents, and both groups received an account balance of \$10. Bags of the candy were opened in front of the subjects and remained in view at the front of the room throughout the practice rounds.⁴ The market value of the candy was not disclosed. After the practice rounds were completed, one of the rounds was randomly chosen for implementation of the actual transactions. Subjects were given the pieces of candy won and their accounts were credited according to their trades.

The formal water-quality experiment consisted of twelve rounds. Buyers were endowed with four ounces of unfiltered tap water sourced from a campus well and sellers were endowed with four ounces of the same tap water that had been filtered through a Zero Water filter. At the end of each round, subjects were told the market price for that round and whether their bid/ask would result in a trade if that round was chosen. After the twelfth round, one of the rounds was chosen at random to implement.

² Because of a computer error that occurred when multiple buyers and/or sellers submitted the same price, more buyers/sellers traded in 28 percent of the rounds than should have. In those instances, all buyers/sellers who submitted the same value traded regardless of our bilateral trade rule with the "house" playing the other side of the trades. This error increased our average payout but did not affect other participants' payouts or trades. Rerunning the analysis when excluding the participants who should not have traded increased our average WTA-WTP gap from 62 cents in the original sample ($N = 48$) to 74 cents in the subsample ($N = 32$). This gap was still significant at the 1 percent, level, but was not statistically different from the WTA-WTP gap in the original 48-pair sample.

³ The market mechanism was described using a generic good—a coffee mug—and two arbitrary quality levels, high and low. Since mugs served no purpose in the experiment other than to provide an example of how the market worked, we did not provide additional information about the quality differences.

⁴ The logistics of the available space did not allow us to put the endowed goods directly in front of participants. The good corresponding to each market role was explicitly noted, clearly labeled, and within view of every participant.

The water filtration process was conducted in front of subjects and the resulting water samples remained in view and clearly labeled throughout the experiment. All subjects were informed in writing and verbally at the start of the formal experiment that they would be asked to drink a glass of water and their decisions during the experiment would determine the quality of the water they would drink. We also emphasized that their bids/asks during the experiment affected the amount of their final payout and that all the rounds were independent.

The twelve rounds were divided into three-round sections, two conducted prior to the information treatment and two conducted after the information treatment. The seven initial sellers and seven initial buyers played three rounds and then switched roles and endowments and played another three rounds. Upon completing those six rounds, the administrator provided the information treatment. Subjects then completed the last six rounds, repeating their initial market roles for rounds seven through nine and switching roles for rounds ten through twelve. After completing the experiment, the subjects filled out a questionnaire that collected information on their demographic characteristics and water quality preferences.

In the final step of the sessions, the subjects were informed of their final payouts and whether the chosen round (selected at random by the software) resulted in their completing a trade that determined the quality of the water they would drink. Subjects who traded were asked to drink water of the quality for which they traded. Subjects who did not trade in the chosen round were asked to drink water of the quality endowed to them in that round. The subjects were brought individually to a separate room, given their final payment, and asked to drink the water in the presence of the experimenter.

Pairing Methodology and Data

We use a within-subject design to measure the significance of endowment in our experiment. We define an observation for each subject as a set consisting of the subject's average WTP and average WTA. Thus, each observation is in fact a pair of observations and that is how we refer to them in the analysis.⁵ The advantage of this method is that it statistically controls for unobservables related to individuals and tests the means of the distribution of differences between WTA and WTP. The type-I error probability for the paired t-test is likely greater than for the standard t-test when data are paired (Zimmerman 1997).

Of the 70 subjects who participated in the experiment, 49 submitted both bids (WTP) and asks (WTA) and 21 submitted asks only. To make full use of the data, we conduct two parallel analyses that account for the nonbidding participants. In the first analysis, we treat nonbidding buyers as having bid \$0 (referred to as

⁵ This terminology should not be confused with other econometric methods that pair or match observations based on characteristics. We do not use any of those matching methods in our analysis.

the inclusive analysis); in the second analysis, we exclude nonbidding buyers (referred to as the exclusive analysis). We do not include nonparticipating sellers' asks (28 percent of seller rounds) in our analysis since interpretation of their potential WTA is not obvious. Inclusion of \$0 bids results in 69 pairs (one anomalous pair was excluded from the analysis⁶) and excluding them results in 48 pairs to use to test our hypotheses. We report the model that excludes nonparticipating buyers primarily because their true bids are not observable and may not be zero.

In these parallel analyses, hypothesis 1 tested the endowment effect using all pairs, 69 pairs in the inclusive analysis and 48 pairs in the exclusive analysis. We find that the WTA-WTP gaps in the inclusive analysis are much larger than the gaps in the exclusive analysis. Average WTP decreases 62 percent (dropping from \$1.23 to \$0.47) and the average gap grows from \$0.62 to \$1.25. Therefore, we view the results of the analysis that excludes nonparticipating sellers as more conservative estimates.

The parallel analyses of hypothesis 2, which relates to the effect of information on the average WTA-WTP gap, uses pairs for which bids and asks were made both before and after the information treatment—66 pairs in the inclusive analysis and 29 pairs in the exclusive analysis. We also test the effect of the information treatment on the extensive margin by modeling the determinants of market participation using a probit model.

The parallel analyses of hypothesis 3, which examines the effect of the initial endowment of either high-quality or low-quality water on subsequent rounds, uses 34 low-quality-endowment and 35 high-quality-endowment pairs in the inclusive analysis and 22 low-quality-endowment and 26 high-quality-endowment pairs in the exclusive analysis.

We use a nonparametric Wilcoxon signed-rank test and a Wilcoxon (Mann-Whitney) rank-sum test (Mann and Whitney 1947, Wilcox 1945) as the preferred statistical tests. We also report Welch's (1938) t-test, which incorporates different population sizes and variances for comparison.

Demographic Characteristics

We collected demographic information on subjects' gender, education, and car ownership as a proxy for income. The results are reported in [Table 1a](#). The

⁶ A review of the pairs for outliers resulted in exclusion of one pair. When reviewing the raw data for influential observations, we found that all of the submitted bids (WTP) for that person were in the 95–99th percentile of bids by all participants. In addition, that individual's average WTA-WTP gap was in the 99th percentile of the sample used for analysis. The p-value comparing average WTA-WTP gaps with and without the outlier was 0.86. Qualitatively, the average gap was \$0.55 with the outlier and \$0.62 without the outlier. The difference in average WTA between initial sellers and initial buyers was \$0.55 with the outlier and \$0.59 without the outlier, and the difference in average WTP between initial sellers and initial buyers was \$0.13 with and \$0.18 without the outlier. Significance did not change in either test.

Table 1a. Demographic Characteristics of the Subjects

| Variable | Subset of Subjects | | | | | | | | |
|--------------------------------------|--------------------|--------------------|--------|-----------------------|--------|-------------------|--------|---------------------|--------|
| | All Subjects | All WTA /WTP Pairs | t-Stat | Information Treatment | t-Stat | Buyer First Round | t-Stat | Seller First Round | t-Stat |
| Female | 0.657 (0.057) | 0.688 (0.068) | 0.76 | 0.655 (0.092) | 0.03 | 0.727 (0.097) | 0.85 | 0.654 (0.095) | 0.04 |
| Income proxies | | | | | | | | | |
| Own car | 0.600 (0.059) | 0.583 (0.072) | 0.42 | 0.621 (0.094) | 0.29 | 0.545 (0.109) | 0.61 | 0.615 (0.097) | 0.20 |
| Lease car | 0.186 (0.047) | 0.146 (0.052) | 1.15 | 0.172 (0.077) | 0.24 | 0.182 (0.084) | 0.06 | 0.115 (0.064) | 1.24 |
| Borrow car | 0.086 (0.034) | 0.104 (0.045) | 0.92 | 0.103 (0.063) | 0.43 | 0.182 (0.084) | 1.57 | 0.038 (0.039) | 1.22 |
| Do not drive | 0.129 (0.040) | 0.167* (0.054) | 1.71 | 0.103 (0.052) | 0.53 | 0.091 (0.063) | 0.68 | 0.231* (0.084) | 1.76 |
| New car | 0.186 (0.047) | 0.125* (0.048) | 1.72 | 0.172 (0.071) | 0.24 | 0.227 (0.091) | 0.57 | 0.038*** (0.039) | 3.00 |
| Used car | 0.686 (0.056) | 0.708 (0.066) | 0.58 | 0.724 (0.082) | 0.58 | 0.682 (0.102) | 0.05 | 0.731 (0.089) | 0.63 |
| Environment and Life Science Degrees | | | | | | | | | |
| Undergraduate | 0.414 (0.059) | 0.438 (0.072) | 0.58 | 0.482 (0.094) | 0.96 | 0.409 (0.107) | 0.06 | 0.462 (0.099) | 0.60 |

Continued

Table 1a. Continued

| Variable | Subset of Subjects | | | | | | | | |
|--|--------------------|--------------------|--------|-----------------------|--------|-------------------|--------|--------------------|--------|
| | All Subjects | All WTA /WTP Pairs | t-Stat | Information Treatment | t-Stat | Buyer First Round | t-Stat | Seller First Round | t-Stat |
| Graduate | 0.143 (0.042) | 0.146 (0.052) | 0.10 | 0.103 (0.038) | 0.81 | 0.136 (0.075) | 0.10 | 0.154 (0.072) | 0.20 |
| Not Environment and Life Science Degrees | | | | | | | | | |
| Undergraduate | 0.414 (0.059) | 0.396 (0.071) | 0.45 | 0.379 (0.089) | 0.49 | 0.409 (0.107) | 0.06 | 0.385 (0.097) | 0.38 |
| Graduate | 0.029 (0.020) | 0.021 (0.021) | 0.49 | 0.034 (0.038) | 0.24 | 0.045 (0.046) | 0.49 | 0.0 (0) | 1.43 |
| <i>N</i> | 70 | 48 | | 29 | | 22 | | 26 | |

Note: We restrict our full sample of 70 subjects to the paired subset of 48 pairs testing hypothesis 1. For testing the information treatment in hypothesis 2, we restrict our sample to 29 pairs. To test the effect of the first-round role, we divide our 48 pairs into Buyers First Round and Sellers First Round. All variables are equal to 1 if yes. All t-tests compare the subsample to the full sample. The car ownership/use demographic characteristics are used as income proxies since the subjects are students. Standard errors are shown in parentheses and the asterisks represent significance levels of a Welch's t-test with ***p < 0.01, **p < 0.05, and *p < 0.10.

average subject was a female (66 percent) undergraduate (82 percent) who owned a car (60 percent). The profile of the 48-pair sample was similar: 69 percent were female, 83 percent were undergraduates, and 58 percent owned a car. Generally, the differences between the full sample and the subsamples were fairly small. Those who did not drive and those who had a new car were weakly significantly different from the full sample. Among the information-treatment pairs, 66 percent were female, 86 percent were undergraduates, and 62 percent owned a car.

In the sample for analysis of the effect of the initial endowment, the proportion of students who had new cars and who initially were sellers was statistically significantly different from the full sample. In the full sample, thirteen subjects had new cars and six of those were part of the 48-pair sample; however, only one initially was a seller. Thus, the proportion of subjects who had new cars in our 48-pair sample was small to begin with relative to the proportion who had used cars. For initial buyers in this analysis, the demographic profile was similar to that of the full sample.

Preferences for Water Quality

We collected information on subjects' preferences for drinking water at home and on campus, knowledge of their at-home drinking water quality, and subjective beliefs regarding water quality in general. The results are reported in [Table 1b](#). For the full sample, the average subject preferred bottled water (53 percent) and filtered tap water (61 percent) at home, preferred bottled water while on campus (57 percent), and believed that water quality was somewhat important but was not familiar with the quality of his/her home drinking water (41 percent). The profiles for all of the subsamples were largely similar to the full-sample profile. There was a slight difference between the 29-pair (exclusive) information sample and the full sample in preference for bottled water at school. The few other differences were weakly significant and could have been caused by the limited sample size.

In general, the average subject was a female undergraduate who owned a used car, preferred filtered and bottled water at home, preferred bottled water on campus, considered water quality somewhat important, and was not familiar with the quality of her home drinking water. We found that the subject profiles were consistent in terms of demographics and water preferences across all subsamples and thus believe there is limited, if any, selection bias in our sampling.

Market Summary Statistics

Participation in the experiment market was skewed; sellers participated more often than buyers and submitted higher asks on average. The average participation rate for sellers in the full sample was 71.4 percent before the information treatment and 71.9 percent after while the average participation

Table 1b. Subjects' Preferences for Water

| Variable | Subset of Subjects | | | | | | | | |
|---------------------------------|--------------------|--------------------|--------|-----------------------|--------|-------------------|--------|--------------------|--------|
| | All Subjects | All WTA /WTP Pairs | t-Stat | Information Treatment | t-Stat | Buyer First Round | t-Stat | Seller First Round | t-Stat |
| Importance of water quality | 2.371 (0.079) | 2.396 (0.083) | 0.39 | 2.448 (0.105) | 0.85 | 2.364 (0.124) | 0.07 | 2.423 (0.113) | 0.53 |
| I know my water quality at home | 0.414 (0.059) | 0.396 (0.071) | 0.45 | 0.448 (0.092) | 0.48 | 0.318 (0.102) | 1.12 | 0.462 (0.1) | 0.60 |
| Water Types Consumed at Home | | | | | | | | | |
| Mineral | 0.157 (0.044) | 0.146 (0.052) | 0.36 | 0.172 (0.063) | 0.29 | 0.136 (0.075) | 0.33 | 0.154 (0.072) | 0.06 |
| Bottled | 0.529 (0.060) | 0.563 (0.072) | 0.83 | 0.586 (0.094) | 0.81 | 0.636 (0.105) | 1.23 | 0.500 (0.1) | 0.36 |
| Filtered | 0.614 (0.059) | 0.646 (0.07) | 0.78 | 0.621 (0.094) | 0.09 | 0.591 (0.107) | 0.27 | 0.692 (0.092) | 1.04 |
| Unfiltered | 0.457 (0.06) | 0.396 (0.071) | 1.51 | 0.448 (0.095) | 0.12 | 0.500 (0.109) | 0.48 | 0.308* (0.092) | 1.99 |
| Water Types Consumed at School | | | | | | | | | |
| Mineral | 0.114 (0.038) | 0.146 (0.052) | 1.46 | 0.172 (0.063) | 1.20 | 0.136 (0.075) | 0.37 | 0.154 (0.072) | 0.75 |

| | | | | | | | | | |
|-------------------|------------------|------------------|------|--------------------|------|------------------|------|------------------|------|
| Bottled | 0.571 (0.06) | 0.646* (0.07) | 1.85 | 0.724** (0.089) | 2.26 | 0.682 (0.102) | 1.29 | 0.615 (0.097) | 0.57 |
| Filtered | 0.400 (0.059) | 0.354 (0.07) | 1.13 | 0.276* (0.089) | 1.83 | 0.273 (0.097) | 1.53 | 0.423 (0.099) | 0.30 |
| Unfiltered | 0.414 (0.059) | 0.417 (0.072) | 0.06 | 0.414 (0.095) | 0.01 | 0.364 (0.105) | 0.58 | 0.462 (0.1) | 0.60 |
| Other | 0.014 (0.014) | 0.021 (0.021) | 1.00 | 0.0 (0) | 1.00 | 0.0 (0) | 1.00 | 0.038 (0.039) | 1.00 |
| Do not consume | 0.029 (0.020) | 0.021 (0.021) | 0.49 | 0.0 (0) | 1.43 | 0.045 (0.046) | 0.49 | 0.0 (0) | 1.43 |
| <i>N</i> | 70 | 48 | | 29 | | 22 | | 26 | |

Note: We restrict our full sample of 70 subjects to the paired subset of 48 pairs testing hypothesis 1. For testing the information treatment in hypothesis 2, we restrict our sample to 29 pairs. To test the effect of the first-round role, we divide our 48 pairs into Buyers First Round and Sellers First Round. Knowledge of home water quality and water types consumed equal 1 if yes. Subjects were asked their opinion on the importance of water quality to them using a Likert scale. The omitted category is Not important at all. All variables equal 1 if yes. All t-tests compare the subsample to the full sample. Standard errors are shown in parentheses and the asterisks represent significance levels of Welch's t-tests with *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

rate for buyers in the full sample was 35.7 percent before the information treatment and 37.1 percent after. Summary statistics for the bids and asks are presented in [Table 2](#).

For the observations from the 48-pair exclusive analysis, the average (median) ask for all rounds was \$1.85 (\$1.68) and was slightly higher before the information treatment than after. The average (median) bid for all rounds was \$1.23 (\$0.88) and again was slightly higher before the information treatment than after. The asks in the 29-pair exclusive analysis that was used to test the information treatment were lower than the asks from the 19 pairs that were excluded. However, the average (median) bid for the 29 pairs, \$1.20 (\$0.86), was not significantly different from the bids from the excluded observations. We find no statistically significant differences in the demographic characteristics of those two groups. The average ask in the 69-pair inclusive analysis was similar to the average ask in the 70-participant exclusive analysis, but the average bid was significantly less in the paired inclusive analysis. As expected, the average bids in the inclusive analyses used to test the information treatment were lower than the average and median bids in the 48-pair exclusive analysis; the average bid for all rounds was \$0.52 compared to \$1.30 in the exclusive observations.

In general, the bids and asks in the experiment were relatively large. A potential explanation for the high valuations is the subjects' stated preferences for bottle water, which suggest that they were willing to pay a premium for higher-quality drinking water; 65 percent of the full sample preferred to drink bottled water while on campus and 56 percent preferred to drink bottled water while at home. Only 42 percent preferred to drink unfiltered water and 35 percent preferred to drink filtered water while on campus. Another potential reason for the unexpectedly large WTA and WTP values is that subjects inflated their valuations because they were playing with house money. Since our primary objective is to identify differences between WTA and WTP, any inflation of values is a secondary concern.

Results

We find evidence of the endowment effect in valuations of drinking water quality that is robust to the water-quality information treatment and several subsamples of our data. We find no evidence that the initial quality endowment influences subsequent valuations, though this result is weaker because our subsample populations are small. There is also some variation in the results depending on whether nonbidders were included or excluded.

Hypothesis 1: Valuations of drinking water quality are not influenced by the endowment effect.

As shown in [Table 3](#), we find a positive and significant gap between WTA and WTP for all rounds collectively, for the rounds before the information treatment,

Table 2. Summary Statistics for Average Bids and Asks Before and After the Information Treatment

| Bid/Ask in U.S. Dollars | Mean | Median | Min. | Max. | Standard Deviation | No. of Subjects |
|---|------|--------|------|------|--------------------|-----------------|
| All Participants Excluding No-bids (<i>N</i> = 70) | | | | | | |
| Seller Ask | 1.72 | 1.49 | 0.02 | 5.23 | 1.20 | 70 |
| Before information | 1.98 | 1.75 | 0.03 | 5.50 | 1.33 | 66 |
| After information | 1.48 | 1.19 | 0.01 | 5.00 | 1.27 | 70 |
| Buyer Bid | 1.30 | 0.88 | 0.01 | 5.17 | 1.25 | 49 |
| Before information | 1.60 | 1.00 | 0.07 | 7.00 | 1.48 | 41 |
| After information | 1.09 | 0.50 | 0.01 | 4.33 | 1.22 | 40 |
| All Pairs Excluding No-bids (<i>N</i> = 48) | | | | | | |
| Seller Ask | 1.85 | 1.68 | 0.02 | 5.23 | 1.27 | 48 |
| Before information | 2.10 | 2.00 | 0.03 | 5.50 | 1.39 | 46 |
| After information | 1.62 | 1.63 | 0.01 | 5.00 | 1.36 | 48 |
| Buyer Bid | 1.23 | 0.88 | 0.01 | 5.17 | 1.15 | 48 |
| Before information | 1.47 | 1.00 | 0.07 | 6.00 | 1.22 | 40 |
| After information | 1.03 | 0.50 | 0.01 | 4.33 | 1.18 | 39 |
| Pairs Used to Test the Information Treatment Excluding No-bids (<i>N</i> = 29) | | | | | | |
| Seller Ask | 1.63 | 1.38 | 0.19 | 5.23 | 1.18 | 29 |
| Before information | 1.89 | 1.58 | 0.30 | 5.46 | 1.30 | 29 |
| After information | 1.40 | 1.50 | 0.08 | 5.00 | 1.22 | 29 |
| Buyer Bid | 1.20 | 0.86 | 0.07 | 5.17 | 1.15 | 29 |
| Before information | 1.38 | 1.00 | 0.07 | 6.00 | 1.28 | 29 |
| After information | 1.05 | 0.62 | 0.08 | 4.33 | 1.13 | 29 |
| All Pairs Including No-bids (<i>N</i> = 69) | | | | | | |
| Seller Ask | 1.72 | 1.48 | 0.02 | 5.23 | 1.20 | 69 |
| Before information | 1.98 | 1.75 | 0.03 | 5.50 | 1.33 | 66 |
| After information | 1.47 | 1.12 | 0.01 | 5.00 | 1.28 | 69 |
| Buyer Bid | 0.47 | 0.17 | 0.00 | 5.17 | 0.82 | 69 |
| Before information | 0.58 | 0.20 | 0.00 | 6.00 | 1.05 | 69 |
| After information | 0.36 | 0.06 | 0.00 | 4.33 | 0.67 | 69 |

Continued

Table 2. Continued

| Bid/Ask in U.S. Dollars | Mean | Median | Min. | Max. | Standard Deviation | No. of Subjects |
|---|------|--------|------|------|--------------------|-----------------|
| Pairs Used to Test the Information Treatment Including No-bids ($N = 66$) | | | | | | |
| Seller Ask | 1.69 | 1.47 | 0.02 | 5.23 | 1.19 | 66 |
| Before information | 1.98 | 1.75 | 0.03 | 5.50 | 1.33 | 66 |
| After information | 1.43 | 1.01 | 0.01 | 5.00 | 1.25 | 66 |
| Buyer Bid | 0.44 | 0.17 | 0.00 | 5.17 | 0.77 | 66 |
| Before information | 0.54 | 0.18 | 0.00 | 6.00 | 0.97 | 66 |
| After information | 0.34 | 0.06 | 0.00 | 4.33 | 0.65 | 66 |

Note: Statistics are for N subjects' average bid or ask. The statistics for panels 1, 2, and 3 do not include nonparticipating buyers (bid = \$0) or sellers (offer greater than \$10). The statistics for panels 4 and 5 include nonparticipating buyers as bid = \$0 but do not include nonparticipating sellers (offer greater than \$10). In all panels, before information refers to water rounds 1–6 before the information treatment and after information refers to water rounds 7–12 after the information treatment.

Table 3. Mean WTA, WTP, and WTA-WTP Gap for Paired Observations

| | N | WTA | WTP | WTA-WTP Difference | Wilcoxon z-Score | Welch's t-Statistic |
|--------------------|-----|-----------------|-----------------|--------------------|------------------|---------------------|
| Including No-bids | | | | | | |
| All rounds | 69 | 1.72 (0.145) | 0.47 (0.099) | 1.25*** (0.196) | 7.14 | 6.37 |
| Before information | 66 | 1.98 (0.164) | 0.54 (0.12) | 1.44*** (0.237) | 6.72 | 6.06 |
| After information | 69 | 1.47 (0.154) | 0.36 (0.08) | 1.11*** (0.187) | 6.91 | 5.97 |
| Excluding No-bids | | | | | | |
| All rounds | 48 | 1.85 (0.183) | 1.23 (0.166) | 0.62*** (0.156) | 4.49 | 3.98 |
| Before information | 38 | 2.08 (0.218) | 1.41 (0.189) | 0.67*** (0.193) | 3.67 | 3.47 |
| After information | 39 | 1.57 (0.218) | 1.03 (0.189) | 0.54*** (0.179) | 3.42 | 3.02 |

Note: The difference is WTA less WTP. The Welch's t-statistic and Wilcoxon signed-rank z-score test the difference between WTA and WTP. Standard errors are shown in parentheses and the asterisks represent the significance level of the Wilcoxon tests with *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Using Welch's t-tests, we find similar significance.

Table 4. Mean WTA-WTP Gap Before and After Information Treatment

| | N | Information | | | Wilcoxon z-Score | Welch's t-Statistic |
|-------------------|----|-----------------|-----------------|--------------------|------------------|---------------------|
| | | Before | After | Difference After | | |
| Including no-bids | 66 | 1.44 (0.237) | 1.09 (0.190) | 0.35*** (0.115) | -3.20 | 3.03 |
| Excluding no-bids | 29 | 0.51 (0.192) | 0.35 (0.171) | 0.16 (0.169) | -1.30 | 0.97 |

Note: The information difference is calculated as difference before information minus difference after information. The Welch's t-statistic and Wilcoxon signed-rank z-score test the difference before information minus the difference after information. Standard errors are shown in parentheses and the asterisks represent the significance level of the Wilcoxon tests with ***p < 0.01, **p < 0.05, and *p < 0.10. Using Welch's t-tests, we find similar significance.

and for the rounds after the information treatment in both the inclusive and the exclusive analysis, rejecting this hypothesis. The average WTA-WTP gap for all rounds is \$0.62 to \$1.25 represented by WTA-WTP ratios of 2.72 to 10.77. For the rounds prior to the information treatment, the average gap is \$0.67 to \$1.44, and for the rounds after the information treatment, the average gap is \$0.54 to \$1.11. All of the differences are statistically significant at the 1 percent level using both the Wilcoxon signed-rank test and the Welch's paired t-test.

The average gap in the inclusive analysis is statistically different from the gap in the exclusive analysis (p-value of 0.00 using the Wilcoxon rank-sum test). Inclusion of the zero bids adds 21 pairs to the sample and reduces average WTP for 44 of the original 48 observations by an average of \$0.61. Inclusion affected average WTP more than average WTA, as expected, resulting in a significant expansion of the average WTA-WTP gap. These results suggest that there is a statistically significant endowment effect associated with drinking water quality and that the effect could be economically significant. If individuals value a loss of water quality at 2.72 to 10.77 times more than they value a gain in water quality, their current endowment will have a strong influence on how they value a change in that water quality.

Hypothesis 2: Information does not change the endowment effect for drinking water quality.

The information treatment failed to reduce the average WTA-WTP gap in the exclusive analysis but had a positive and significant effect in the inclusive analysis. For the 66 subjects who submitted bids and asks both before and after the information treatment, the average gap in the inclusive analysis was \$1.44 before information and \$1.09 after. The average difference-in-differences of \$0.35 is statistically significant at the 1 percent level using the Wilcoxon signed-rank test and Welch's t-test. This result rejects the second

hypothesis. For the 29 subjects in the exclusive analysis, the average gap was \$0.51 before information and \$0.35 after, and the average difference-in-differences of \$0.16 is statistically insignificant at the 10 percent level using the Wilcoxon signed-rank test and Welch's paired t-test. This result fails to reject the second hypothesis.

When we explored the reasons behind the drastic change in magnitude and significance of the gaps in response to the information treatment between the exclusive and inclusive analyses, we found that the change was driven by 20 subjects (of 66) who never submitted a bid. When we excluded those pairs in the inclusive analysis, we found an average difference-in-differences of \$0.25, which is significant at the 10 percent level using the Wilcoxon signed-rank test and Welch's paired t-test.

These findings provide a mixed interpretation of the effect of increasing U.S. consumers' familiarity with drinking water quality. Using the more-conservative sample of 29 pairs in the exclusive analysis, we fail to reject the hypothesis that familiarity with water quality does not reduce the endowment effect. However, when we assume that nonparticipation is actually a WTP of zero, the hypothesis is rejected.

The positive information-treatment effect in the inclusive analysis suggests that the information provided may be affecting the extensive margin—that is, the decision to participate in the market by submitting a bid or ask. Therefore, we conducted an additional analysis that modeled the determinants of market participation and tested whether the information treatment affected the extensive margin and, if so, whether the effect persisted when subjects played the role of seller more often than the role of buyer.⁷

We find that the subject's role in the market strongly affects market participation (see Table 5). The likelihood of participation is 34.5 percent greater when subjects play the role of seller than when they play the role of buyer. The information treatment also increases market participation—by 24.3 percent on average—and the effect is even larger when the subjects also act as sellers. Combined, these findings suggest that the information treatment increases market participation, especially by the seller, and reduces the WTA-WTP gap.

Hypothesis 3: The initial endowment of drinking water quality does not influence subsequent valuations.

As shown in Table 6, we find a positive and significant effect from the initial endowment of water in the inclusive analysis. Those who were initially

⁷ Ideally, we would take an additional step and jointly estimate the extensive and intensive margins in one analysis using a selection model such as Heckman. However, we chose not to pursue that step because of the small sample size and lack of a variable in our data to satisfy the exclusion restriction.

Table 5. Probit Results for the Effect of the Information Treatment on Market Participation

| | Dependent Variable: Market Participation | |
|--------------------------------|--|---------------------|
| | (1) | (2) |
| Seller | 0.961*** (0.098) | 1.701*** (0.138) |
| Information treatment | 0.696*** (0.110) | 1.543*** (0.156) |
| Seller × Information treatment | — | 1.566*** (0.156) |
| Observations | 770 | 770 |
| Pseudo R-squared | 0.156 | 0.222 |

Note: Market participation is 1 if the subject participated as a seller or buyer. The seller and information treatments equal 1 if yes. Standard errors are shown in parentheses and the asterisks represent the significance level with *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Other variables included as controls include the round, lag in the market price, gender, student status, car ownership, importance and knowledge of water quality, and types of drinking water at home and at school. All of the controls not shown in this table are insignificant at the 10 percent level.

sellers had an average WTA for all rounds that was \$0.67 greater than those who were initially buyers, a difference that is significant at the 5 percent level for both the Wilcoxon nonparametric test and Welch’s t-test. In the exclusive analysis, however, the average difference in WTA dropped to \$0.59, which is no longer significant at the 10 percent level. Still, the average difference in WTA is economically large at 49 percent and 39 percent of the average buyer’s WTA.

Turning to WTP, we fail to find a significant effect associated with subjects’ first-round role/endowment on subsequent market behavior in either model.⁸ The results of Welch’s t-test are similar but have slightly higher t-statistics than the Wilcoxon z-scores. Though we fail to find a consistently significant effect of subjects’ first-round role/endowment, the test may be relatively weak because of the constraint of the small sample.

Robustness

We checked the results from our analysis of the endowment effect using various subsets of the data and report the results in Table 7. We tested the average

⁸ We find the same insignificant result when testing subsets of buyers and sellers before and after the information treatment.

Table 6. Comparison of Mean WTA and WTP by First-round Role

| | First Round | | Difference between First Round Sellers and Buyers | Wilcoxon z-Score | Welch's t-Statistic |
|-------------------|-----------------|-----------------|--|---------------------|------------------------|
| | Seller | Buyer | | | |
| Including No-bids | | | | | |
| WTA | 2.05 (0.223) | 1.38 (0.168) | 0.67** (0.279) | -2.14 | 2.42 |
| WTP | 0.48 (0.157) | 0.45 (0.121) | 0.03 (0.198) | -0.32 | 0.14 |
| N | 35 | 34 | | | |
| Excluding No-bids | | | | | |
| WTA | 2.12 (1.359) | 1.53 (1.101) | 0.59 (0.355) | -1.40 | 1.65 |
| WTP | 1.31 (1.259) | 1.13 (1.032) | 0.18 (0.331) | -0.34 | 0.53 |
| N | 26 | 22 | | | |

Note: The difference is calculated as the first-round seller minus the first-round buyer. In the first round, first-round sellers were endowed with high-quality drinking water and first-round buyers were endowed with low-quality drinking water. Welch's t-statistic and the Wilcoxon rank-sum z-score test the difference between the first-round groups. Standard deviations are shown in parentheses and the asterisks represent the significance level of the Wilcoxon tests with *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Using Welch's t-tests, we find similar significance.

WTA-WTP gap for first-round-role subsets and when excluding some rounds. All of the robustness checks of the subsamples in the inclusive analysis were significant at the 1 percent level using both the Wilcoxon signed-rank test and Welch's paired t-test, and the average gaps were 31–52 percent higher than the average gaps in the exclusive analysis. Therefore, we present the robustness checks only for the subsamples in the exclusive analysis.

For the subsets of first-round roles (22 subjects initially were buyers and 26 initially were sellers), we find a positive and significant WTA-WTP gap for all rounds (Wilcoxon signed-rank test and Welch's t-test). The average gap for the initial sellers is \$0.81, representing a ratio of 3.45, that is significant at the 1 percent level. The average gap for initial buyers is \$0.40, representing a ratio of 1.87, that is significant at the 5 percent level.

Since our experimental design provided subjects with an opportunity to learn about the market mechanism through repetition of rounds, we also checked our initial findings using subsets of the number of rounds completed (see Table 7). We report significance in this case using the Wilcoxon signed-rank test. When

Table 7. Robustness Check: Comparison of Mean WTA, WTP, and WTA-WTP Gap

| Subset | N | WTA | WTP | WTA-WTP Difference | Wilcoxon z-Score | Welch's t-Statistic |
|------------------------------|----|-----------------|-----------------|--------------------|------------------|---------------------|
| First Round Role | | | | | | |
| Seller | 26 | 2.12 (0.267) | 1.31 (0.247) | 0.81*** (0.248) | 3.84 | 3.26 |
| Buyer | 22 | 1.53 (0.235) | 1.13 (0.220) | 0.40** (0.179) | 2.32 | 2.23 |
| Rounds Included | | | | | | |
| 2-12 | 48 | 1.76 (0.177) | 1.15 (0.159) | 0.60*** (0.146) | 4.87 | 4.14 |
| 2-6 and 8-12 | 47 | 1.71 (0.175) | 1.16 (0.162) | 0.57*** (0.147) | 4.37 | 3.87 |
| 2, 3, 5, 6, 8, 9, 11, and 12 | 43 | 1.68 (0.176) | 1.12 (0.159) | 0.51*** (0.149) | 3.68 | 3.40 |
| 3, 6, 9, and 12 | 31 | 1.49 (0.213) | 1.16 (0.180) | 0.33 (0.187) | 1.52 | 1.76 |
| 8-12 | 36 | 1.48 (0.24) | 1.04 (0.205) | 0.43*** (0.185) | 2.81 | 2.40 |
| 8, 9, 11, and 12 | 29 | 1.39 (0.261) | 1.00 (0.204) | 0.38** (0.189) | 2.44 | 2.04 |

Note: There were twelve water rounds: 1-6 before the information treatment and 7-12 after the information treatment. Panel 1 tests the average WTA-WTP gap across all rounds for the first-round-seller subset and average WTA-WTP gap across all rounds for the first-round-buyer subset. Panel 2 tests the average WTA-WTP gap for different subsets of rounds. Difference is WTA less WTP. Standard errors are shown in parentheses. The Welch's t-statistic and Wilcoxon signed-rank z-score test the difference between WTA and WTP. The asterisks represent the significance level of the Wilcoxon tests with ***p < 0.01, **p < 0.05, and *p < 0.10. Using Welch's paired t-tests, we find similar significance except for the subset that included rounds 3, 6, 9, and 12, which is significant at the 10 percent level using Welch's paired t-test.

we exclude the first round, we find a gap of \$0.60 with a ratio of 2.48, which is significant at the 1 percent level. When we exclude the first round in the experiment and the first round after the information treatment, we find an average gap of \$0.57 and a ratio of 2.50, which is significant at the 1 percent level. When we exclude the first round in each of the three-round experiment sections, we find an average gap of \$0.51 and a ratio of 1.89, which is significant at the 1 percent level. Using only the last round of each three-round section produced an insignificant average gap of \$0.33 and a ratio of 1.53.⁹

We also restrict the data to a subset of paired observations in rounds 8 through 12, which occurred after the information treatment, and find an average gap of \$0.43 and a ratio of 2.44 that is significant at the 1 percent level.¹⁰ When we restrict the sample even further by excluding the first round of each post-information-treatment section, we find an average gap of \$0.38 and a ratio of 1.63, which is significant at the 5 percent level. These checks provide evidence that our results are relatively robust and that the average WTA-WTP ratio is approximately 2.72 for the change in water quality used in this experiment.

We also regress the individual WTA-WTP gaps on various demographic characteristics and water preferences and report the results for four exclusive samples and one inclusive sample in [Table 8](#). The only significant effect is a reduction in the average WTA-WTP gap of \$0.56 to \$0.62 for subjects who knew the quality of their home drinking water, a difference that is significant at the 10 percent level. This information effect is distinct from the effect of our information treatment in that it captures the change in the average WTA-WTP gap. It is not an average treatment effect; instead, it shows that valuations of drinking water quality vary among demographic groups and with preferences for drinking water quality. Knowing the quality of one's home drinking water could account for part of the \$0.62 gap under hypothesis 1 ([Table 3](#)), which would suggest that individuals who have such knowledge have a small or nonexistent gap. For the inclusive sample, the effect of knowing the quality of one's home drinking water is much smaller and statistically insignificant, but the proxy for income (car ownership) is significant. It is not clear why income would affect the size of the endowment effect. It may be related to subjects' economic ability to pay for higher-quality (bottled) water.

Discussion

By using a within-subject experimental design to elicit WTA and WTP for drinking water quality and endowing sellers and buyers with different

⁹ Using Welch's t-test, this gap is significant at the 10 percent.

¹⁰ Using Welch's t-test, this gap is significant at the 5 percent level.

Table 8. Ordinary Least Square Results for the Effect of Subject Characteristics on the Average WTA-WTP Gap

| Characteristic | Dependent Variable: Average WTA-WTP Gap | | | | |
|---------------------------------------|---|--------------------|-------------------|--------------------|--------------------|
| | Exclusive | | | | Inclusive |
| | (1) | (2) | (3) | (4) | (5) |
| Female | -0.344 (0.429) | -0.391 (0.470) | -0.213 (0.349) | -0.123 (0.393) | -0.076 (0.275) |
| Own car | 0.194 (0.411) | -0.005 (0.432) | 0.137 (0.404) | -0.0106 (0.429) | -0.551* (0.291) |
| Used car | -0.282 (0.441) | -0.081 (0.468) | -0.291 (0.447) | -0.170 (0.470) | 0.309 (0.324) |
| I know my home drinking water quality | -0.617 (0.387) | -0.881* (0.430) | -0.56* (0.306) | -0.622* (0.325) | -0.210 (0.266) |
| Major | | | | | |
| Science and environment undergraduate | 0.236 (0.357) | 0.325 (0.371) | 0.158 (0.318) | 0.167 (0.340) | -0.063 (0.274) |
| Not science and environment graduate | 0.825 (1.224) | 0.698 (1.247) | 0.614 (1.069) | 0.551 (1.112) | 0.446 (0.778) |
| Science and environment graduate | 0.190 (0.568) | 0.214 (0.630) | 0.199 (0.457) | 0.018 (0.502) | -0.157 (0.416) |
| Water Types Consumed at Home | | | | | |
| Mineral | 0.097 (0.556) | 0.206 (0.583) | — — | — — | — — |

Continued

Table 8. Continued

| Characteristic | Dependent Variable: Average WTA-WTP Gap | | | | |
|--|---|-------------------|------------------|------------------|-------------------|
| | Exclusive | | | Inclusive | |
| | (1) | (2) | (3) | (4) | (5) |
| Bottled | -0.185 (0.417) | -0.464 (0.473) | — — | — — | — — |
| Unfiltered | -0.164 (0.503) | -0.464 (0.549) | — — | — — | — — |
| Other | -0.626 (1.212) | -1.360 (1.349) | — — | — — | — — |
| Water Types Consumed at School | | | | | |
| Mineral | 0.115 (0.516) | -0.027 (0.563) | — — | — — | — — |
| Filtered | -0.211 (0.333) | -0.055 (0.370) | — — | — — | — — |
| Unfiltered | -0.058 (0.483) | -0.046 (0.502) | — — | — — | — — |
| Other | 0.484 (1.212) | 0.280 (1.269) | — — | — — | — — |
| Water Types Consumed at School, Home, or Both | | | | | |
| Mineral | — — | — — | 0.105 (0.379) | 0.132 (0.438) | -0.274 (0.341) |

Continued

Table 8. Continued

| Characteristic | Dependent Variable: Average WTA-WTP Gap | | | | |
|-----------------------|---|-------|---------|-----------|---------|
| | Exclusive | | | Inclusive | |
| | (1) | (2) | (3) | (4) | (5) |
| Filtered | — | — | 0.058 | 0.113 | 0.148 |
| | — | — | (0.314) | (0.358) | (0.276) |
| Unfiltered | — | — | −0.016 | −0.0653 | −0.176 |
| | — | — | (0.321) | (0.388) | (0.262) |
| Other | — | — | 0.091 | −0.315 | −0.780 |
| | — | — | (1.068) | (1.170) | (0.721) |
| N | 48 | 48 | 48 | 48 | 69 |
| R-squared | 0.170 | 0.252 | 0.135 | 0.190 | 0.353 |
| F-statistic | 0.44 | 0.5 | 0.51 | 0.5 | 1.93 |
| Session fixed effects | N | Y | N | Y | Y |

Note: All of the variables are binary (1 = yes). Not science and environment undergraduate and no car were omitted in all models. Filtered water at home and bottled water on campus were omitted in models 1 and 2. Bottled water was omitted in models 3, 4, and 5. The water-type variables in models 3, 4, and 5 equal 1 if water type is consumed at school, home, or both. Standard errors are shown in parentheses and the asterisks represent significance levels of the coefficient where ***p < 0.01, **p < 0.05, and *p < 0.10.

qualities of drinking water, we sought to determine whether the endowment effect exists for drinking water quality. We find that it does. The analysis identifies a significant and positive gap of \$0.62 to \$1.25 between WTA and WTP that is robust to provision of information, first-round roles, and restricted subsamples. Relative to such gaps for other goods, our range of WTA-WTP ratios of 2.72 to 10.77 is larger than the range found for ordinary private goods (Tunçel and Hammitt 2014). The sizes of the gaps and the ratios are somewhat surprising given that drinking water quality can be thought of as a private good. The 2.72 ratio can be viewed as a lower bound; students tend to have smaller WTA-WTP ratios than the general public (Horowitz and McConnell 2002), within-subject designs tend to reduce WTA-WTP gaps (Sayman and Öncüler 2005), and our change in water quality is relatively small.

Our finding of a significant and economically large gap between WTA and WTP for drinking water quality presents potential policy implications. The most obvious is that evaluations of projects meant to raise the quality of drinking water beyond required standards would not be especially valuable to consumers. While our specific results are not generalizable to water bodies and environmental goods, we believe that identification of this significant gap indicates that the endowment effect should be considered in valuations of water quality projects.

We identified several interesting trends in our demographic and robustness checks. It turns out that consumers' incomes and their being familiar with the quality of their home drinking water are associated with reductions in the size of the endowment effect. We also find that the WTA-WTP gap is generally robust to restricting the number of rounds analyzed.

Provisions of information on water quality and experience with a water market may reduce WTA-WTP gaps but will not eliminate them. Our analysis uses multiple observations from each subject to control for idiosyncratic differences in the subjects via our pairing methodology. Alternative methods present an avenue for future research, but may lead to different results depending on the elicitation method (Lusk, Feldkamp, and Schroeder 2004).

Another potential explanation for the WTA-WTP gap is that buyers underbid their true valuations for moral reasons. According to the concept of moral omission/commission (DeScioli, Bruening, and Kurzban 2011), buyers are likely to offer lower prices to avoid taking away a seller's filtered water, which results in underestimation of the buyers' willingness to pay for filtered water. If moral concerns govern a buyer's choice in the market, the buyer may prefer to opt out of the market to eliminate the risk of affecting other players. Alternatively, sellers could ask for lower prices to avoid excluding buyers and thus would always choose to participate in the market. Future studies could more-explicitly test moral motivations as potential drivers of market behavior and valuations. Additional research is also needed to improve our understanding of motivations for market behavior and links between endowment and valuation of environmental goods related to water quality.

References

- Amigues, J., C. Boulatoff, B.C. Gauthier, and J.E. Keith. 2002. "The Benefits and Costs of Riparian Analysis Habitat Preservation: A Willingness to Accept/Willingness to Pay Contingent Valuation Approach." *Ecological Economics* 43(1): 17–31.
- Brown, T.C., and R. Gregory. 1999. "Why the WTA-WTP Disparity Matters." *Ecological Economics* 28(3): 323–335.
- Brox, J.A., R.C. Kumar, and K.R. Stollery. 2003. "Estimating Willingness to Pay for Improved Water Quality in the Presence of Item Nonresponse Bias." *American Journal of Agricultural Economics* 85(2): 414–428.
- 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.
- Del Saz-Salazar, S., F. Hernandez-Sancho, and R. Sala-Garrido. 2009. "The Social Benefits of Restoring Water Quality in the Context of the Water Framework Directive: A Comparison of Willingness to Pay and Willingness to Accept." *Science of the Total Environment* 407(16): 4574–4583.
- DeScioli, P., R. Bruening, and R. Kurzban. 2011. "The Omission Effect in Moral Cognition: Toward a Functional Explanation." *Evolution and Human Behavior* 32(3): 204–215.
- Environmental Protection Agency. 2013. "2011 Drinking Water Needs Survey and Assessment Fact Sheet." EPA 816-F-13-001, Washington, DC.
- Fischbacher, U. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10(2): 171–178.
- Hanemann, W.M. 1991. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *The American Economic Review* 81(3): 635–647.
- Harless, D.W. 1989. "More Laboratory Evidence on the Disparity between Willingness to Pay and Compensation Demanded." *Journal of Economic Behavior and Organization* 11(3): 359–379.
- Hjorth, K., and M. Fosgerau. 2011. "Loss Aversion and Individual Characteristics." *Environmental Resource Economics* 49(4): 573–596.
- Horowitz, J.K., and K.E. McConnell. 2002. "A Review of WTA/WTP Studies." *Journal of Environmental Economics and Management* 44(3): 426–447.
- Jordan, J.L., and A.H. Elnagheeb. 1993. "Willingness to Pay for Improvements in Drinking Water Quality." *Water Resources Research* 29(2): 237–245.
- Kahneman, D., J.L. Knetsch, and R.H. Thaler. 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *The Journal of Economic Perspectives* 5(1): 193–206.
- Kling, C.L., J.A. List, and J. Zhao. 2013. "A Dynamic Explanation of the Willingness to Pay and Willingness to Accept Disparity." *Economic Inquiry* 51(1): 909–921.
- Knetsch, J.L., and J.A. Sinden. 1984. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measure of Values." *Quarterly Journal of Economics* 99(3): 507–521.
- List, J.A. 2003. "Does Market Experience Eliminate Market Anomalies?" *The Quarterly Journal of Economics* 118(1): 41–71.
- Lusk, J.L., Y. Feldkamp, and T.C. Schroeder. 2004. "Experimental Auction Procedure: Impact on Valuation of Quality Differentiated Goods." *American Journal of Agricultural Economics* 86(2): 389–405.
- MacDonald, D.H., M.D. Morrison, and M.B. Barnes. 2010. "Willingness to Pay and Willingness to Accept Compensation for Changes in Urban Water Customer Service Standards." *Water Resources Management* 24(12): 3145–3158.
- MacDonald, H.F., and J.M. Bowker. 1994. "The Endowment Effect and WTA: A Quasi-experimental Test." *Journal of Agricultural and Applied Economics* 26(2): 545–551.
- Mann, H.B., and D.R. Whitney. 1947. "On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other." *The Annals of Mathematical Statistics* 18(1): 50–60.

- McConnell, K.E., and M.A. Rosado. 2000. "Valuing Discrete Improvements in Drinking Water Quality through Revealed Preferences." *Water Resources Research* 36(6): 1575–1582.
- Morrison, G.C. 1998. "Understanding the Disparity between WTP and WTA: Endowment Effect, Substitutability, or Imprecise Preferences?" *Economics Letters* 59(2): 189–194.
- Pires, M. 2004. "Watershed Protection for a World City: The Case of New York." *Land Use Policy* 21(2): 161–175.
- Plott, C.R., and K. Zeiler. 2005. "The Willingness to Pay – Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Valuations." *American Economic Review* 95(3): 530–545.
- . 2007. "Exchange Asymmetries Incorrectly Interpreted As Evidence of Endowment Effect Theory and Prospect Theory?" *The American Economic Review* 97(4): 15–38.
- Postel, S.L., and B.H. Thompson Jr. 2005. "Watershed Protection: Capturing the Benefits of Nature's Water Supply Services." *Natural Resources Forum* 29(2): 98–108.
- Randall, A., and J.R. Stoll. 1980. "Consumer's Surplus in Commodity Space." *American Economic Review* 70(3): 449–455.
- Rowe, R.D., R.C. D'Arge, and D.S. Brookshire. 1980. "An Experiment on the Economic Value of Visibility." *Journal of Environmental Economics and Management* 7(1): 1–19.
- Sayman, S., and A. Öncüler. 2005. "Effects of Study Design Characteristics on the WTA–WTP Disparity: A Meta Analytical Framework." *Journal of Economic Psychology* 26(2): 289–312.
- Schmidt, U., and S. Traub. 2009. "An Experimental Investigation of the Disparity between WTA and WTP for Lotteries." *Theory and Decision* 66(3): 229–262.
- Shogren, J.F., S.Y. Shin, D.J. Hayes, and J.B. Kliebenstein. 1994. "Resolving Differences in Willingness to Pay and Willingness to Accept." *American Economic Review* 84(1): 255–270.
- Tsur, M. 2008. "The Selectivity Effect of Past Experience on Purchasing Decisions: Implications for the WTA–WTP Disparity." *Journal of Economic Psychology* 29(5): 739–746.
- Tunçel, T., and J.K. Hammitt. 2014. "A New Meta-analysis on the WTP–WTA Disparity." *Journal of Environmental Economics and Management* 68(1): 175–187.
- Tversky, A., and D. Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-dependent Model." *Quarterly Journal of Economics* 106(4): 1039–1061.
- Welch, B.L. 1938. "The Significance of the Difference between Two Means When the Population Variances Are Unequal." *Biometrika* 29(3/4): 350–362.
- Whittington, D., J. Briscoe, X. Mu, and W. Barron. 1990. "Estimating the Willingness to Pay for Water Services in Developing Countries: A Case Study of the Use of Contingent Valuation Surveys in Southern Haiti." *Economic Development and Cultural Change* 38(2): 293–311.
- Wilcoxon, F. 1945. "Individual Comparisons by Ranking Methods." *Biometrics Bulletin* 1(6): 80–83.
- Zhao, J., and C.L. Kling. 2001. "A New Explanation for the WTP/WTA Disparity." *Economics Letters* 73(3): 293–300.
- Zimmerman, D.W. 1997. "Teacher's Corner: A Note on Interpretation of the Paired-sample t-Test." *Journal of Educational and Behavioral Statistics* 22(3): 349–360.