Resolving Differences in Willingness to Pay and Willingness To Accept

By JASON F. SHOGREN, SEUNG Y. SHIN, DERMOT J. HAYES, AND JAMES B. KLIEBENSTEIN*

This paper tests the conjecture that the divergence of willingness to pay (WTP) and willingness to accept (WTA) for identical goods is driven by the degree of substitution between goods. In contrast to well-known results for market goods with close substitutes (i.e., candy bars and coffee mugs), our results indicate a convergence of WTP and WTA measures of value. However, for a nonmarket good with imperfect substitutes (i.e., reduced health risk), the divergence of WTP and WTA value measures is persistent, even with repeated market participation and full information on the nature of the good. (*JEL* D10, C91)

Over the past decade, a consistent and frustrating pattern of empirical evidence has accumulated suggesting a significant divergence of willingness-to-pay (WTP) measures of value, where individuals buy an object, and willingness-to-accept (WTA) measures of value, where individuals sell the same object. Field-contingent valuation studies first uncovered the pattern, and laboratory markets have confirmed that the divergence is persistent (see e.g., Judd Hammack and Gardner M. Brown, Jr., 1974; Robert D. Rowe et al., 1980; Jack L. Knetsch and John A. Sinden, 1984; David S. Brookshire and Don L. Coursey, 1987). The divergence is troubling in that the interpretation of standard theory predicts that with small income effects WTP and WTA should be equivalent, or at least within a tight bound (see Robert Willig, 1976; Alan Randall and John R. Stoll, 1980). Moreover, since valuation measures are used for the study of many public-policy questions, these results raise questions about which procedure to use in practice. The evidence that WTA measures significantly exceed WTP measures suggests a need to reexamine the analytical foundations of value measures.

In response, Michael W. Hanemann (1991) has offered a straightforward explanation of why the value divergence occurs and by how much. By recognizing that substitution effects have a more important role than previously realized, Hanemann demonstrated that the divergence can range from zero to infinity, depending on the degree of substitution between goods and given a positive income elasticity. Hanemann showed that one should only expect convergence of WTP and WTA value measures when the good in question has a very close substitute. When the good has an imperfect substitute, a value divergence will exist and will expand as the degree of substitution decreases.

This paper tests Hanemann's proposition in a nonhypothetical experimental auction market. Our results provide some support for his argument. We find that for two private market goods with a relatively close substitute (a candy bar and a coffee mug)

*Shogren: Department of Economics, Iowa State University, Ames, IA 50011, and School of Forestry and Environmental Studies, Yale University, New Haven, CT 06511; Shin, Hayes, and Kliebenstein: Department of Economics, Iowa State University, Ames, IA 50011. Tom Crocker, John Braden, and Helen Jensen provided helpful comments on experimental design. Young Park and Todd Holt supplied valuable research assistance. We gratefully acknowledge the financial support of the Food Safety Consortium. Two anonymous reviewers added comments that have improved the paper. All errors remain our own. This is Journal Paper No. J-14801 of the Iowa Agriculture and Home Economics Experiment Station, Ames, IA (Project No. 2994).

the divergence of WTP and WTA value measures disappears with repeated exposure to the market. In contrast, for a private nonmarket good with no close substitute (reduction of human health risk) the divergence is robust and persistent, even given repeated market participation and full information on the characteristics of the good.

This paper proceeds as follows. Section I describes our general experimental design in terms of Hanemann's model. Sections II and III outline the experimental procedures and results. Section IV describes a second experiment that explores the relationship between the substitution effect and the endowment effect. Our conclusions are offered in Section V.

I. Substitution Effects and the General Experimental Design

Hanemann (1991) reconsidered the work of Randall and Stoll (1980) on value measures given changed quantities of the good. He argued that the widely interpreted view of Randall and Stoll's result, which implies approximate equality between WTP and WTA is "misconceived." By recognizing that the difference in value measures depends on both income and substitution effects, Hanemann demonstrated that the fewer available substitutes the greater is the divergence between WTP and WTA. Intuitively, this conclusion makes a great deal of sense. One's willingness to accept compensation for a decrease in the level of a unique good, either private or public, need not equal one's willingness to pay for a good where one is constrained by income.

To illustrate, assume that an individual derives utility, $u = u(x, q)$, from consuming a numeraire good (x) and the good under consideration (q; e.g., candy bar, coffee mug, or health). In general, WTA will exceed WTP unless there is no income effect, in which case WTA equals WTP (see e.g., P. R. G. Layard and A. A. Walters, 1978 pp. 150–53). The standard argument is that income effects are small for the changes measured in most valuation studies so this cannot account for the observed WTP–WTA divergence. Randall and Stoll (1980) take this a step further, arguing that when goods are sold in competitive markets with zero transactions costs (i.e., candy bars and coffee mugs), the goods possess the characteristics of money, so there will be perfect substitutability. Perfect substitution implies a linear indifference curve for x and q (i.e., a frictionless intermediate monetary exchange of commodity holdings).

Consider Figure 1. The top panel shows the Randall and Stoll (1980) scenario where intermediate monetary adjustments create perfect substitution between the two goods. Here the WTA measure is the quantity of the numeraire good required to compensate the individual for forgoing a change in q from q_0 to q_1. This is the amount AD which puts the individual on the higher indifference curve, but which maintains the original q_0 consumption level. The WTP measure is the quantity of the numeraire good that one can take from an individual after the change to q_1, while still leaving him or her as well off as before. This amount is BC. Given perfect substitution, BC equals AD. This is Hanemann's (1991 p. 637) proposition 1. Note that Randall and Stoll (1980) also point out that, not only should WTP equal WTA, both should equal the average market selling price of the good.

Now consider the case of goods such as health, where markets are incomplete or goods are lumpy or indivisible. Hanemann's insight was that certain goods like health cannot be perfectly exchanged for money. There is friction in the market, thereby violating Randall and Stoll's case of intermediate monetary adjustments in commodity holdings. This leads to indifference curves of the standard shape, strictly convex to the origin.[1] Hanemann demonstrated that in the quantity-change case WTA will exceed WTP; the divergence depends on the income elasticity divided by the Allen-Uzawa elasticity of substitution. Therefore an increase in income elasticity or a decrease in

[1]Randall and Stoll (1980) do recognize this possibility (note their brief discussion on p. 452).

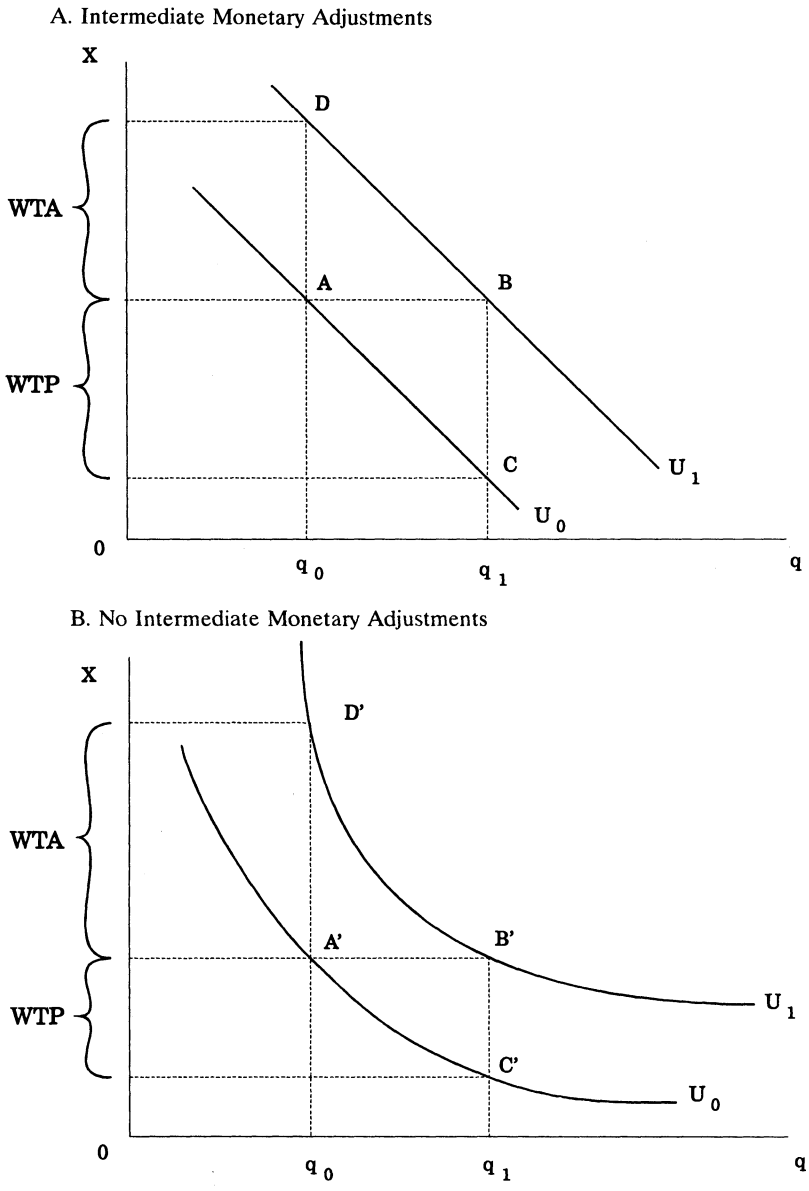

FIGURE 1. THE SIMPLE ANALYTICS OF THE WTA–WTP DIVERGENCE

the degree of substitutability will increase the WTP–WTA divergence.

Figure 1B illustrates. The assumption of health and wealth as imperfect substitutes is reflected by the curvature of the indifference curves. The individual's WTP to secure the new level of health, q_1, keeping him at his original utility level, u_0, is B′C′. In contrast, the compensation (WTA) required to reach the new level of utility, u_1, while remaining at the original health level, q_0, is A′D′. Note that A′D′ exceeds B′C′, or WTA > WTP. As the degree of substitutability decreases, the trade-off between

health and wealth becomes less desirable, implying a greater divergence between WTP and WTA.

We constructed the following general experimental design to determine whether the degree of substitution significantly affects the divergence of value measures. For the case of perfect substitution, we used a regular-size brand-name candy bar. Because the candy bar is readily available in many marketing outlets including stores and vending machines and because there are minimal transactions costs, the substitution possibilities are essentially limitless. Therefore, if the degree of substitution is indeed critical, the value measures should converge with repeated market participation. In addition, both value measures should be close to the average market price.

For the case of low substitution, we auctioned a nonmarket good represented by reduced risk from food-borne illness. We hypothesize that individuals should have a relatively low, or zero, degree of substitution between health and all other commodities as represented by wealth. If Hanemann's conjecture is correct, we expect the WTA measures of value to be significantly greater than the WTP measures. The following proposition summarizes our test.

PROPOSITION 1 (Convergence Proposition): *Given positive income elasticity and repeated market participation, the WTP and WTA measures of value will converge for the market good with close substitutes (candy bar) but will not converge for the nonmarket good with imperfect substitutes (reduced health risk from food-borne pathogens). In addition, the value measures for the market good should be related to the average selling price in the market.*

If we can reject the convergence proposition, then we cannot support Hanemann's argument. In this case, other explanations such as the endowment effect or loss-aversion become more attractive (see e.g., Daniel Kahneman et al., 1990). If we cannot reject the proposition, however, then we can offer support to the conjecture that the degree of substitution is a key to understanding the disparity between WTP and WTA measures of value.

II. Experimental Procedures

The experiment was divided into two stages. Stage 1 was the market-good auction. Stage 2 was the nonmarket-good auction. Subjects participated in both stages either for the WTP or WTA experiment. See the Appendix for instructions for the WTP experiment. The WTA experiment was identical to the WTP experiment in all aspects except for the value measure and initial ownership of the good.

In stage 1, each subject was provided an initial income of $3 and a small piece of candy. To facilitate learning and value formation, the auction was repeated over five trials. The number of trials was selected after pretesting to determining how quickly individual value measures stabilized. Note that to control wealth effects, we made the subjects fully aware that only one of the five trials was binding. The binding trial was selected at random by a Monte Carlo number generator on a personal computer. In an attempt to elicit preference accurately we used a Vickrey second-price sealed-bid auction (see William Vickrey, 1961). The Vickrey auction has successfully elicited values in various experimental settings (see Vikki Coppinger et al., 1980; Don L. Coursey, 1987; Shogren, 1990).

The market good was a regular-size brand-name candy bar. Each subject was asked the maximum he or she would be willing to pay to upgrade the small piece of candy to the brand-name candy bar. For each trial, each subject recorded a bid on a recording card that was collected by the monitor. The highest bidder's identification number and the reigning price (the second-highest bid) were posted as public information on a blackboard. Each subject was given a $3 endowment in the candy-bar stage.

Stage 2 was the nonmarket-good auction. The procedures were similar to those in stage 1, with some noted exceptions. Each participant was given an initial income of $15. Two types of food items were then shown to the subjects with a description of

each item. The first type was the test product, which represented food purchased from a local source with a typical chance of being contaminated with a food-borne pathogen from one-time consumption. Five food-borne pathogens were considered in five separate experimental sessions: *Campylobacter*, *Salmonella*, *Staphylococcus aureus*, *Trichinella spiralis*, and *Clostridium perfringens*.[2] All five pathogens occur in the United States. The test product was provided to every participant as a free lunch. The second food type was stringently screened food. The stringently screened food had been tested for food-borne pathogens and had a low probability (one in 100 million) of causing food-borne illness.

Each participant was then asked the maximum he or she would be willing to pay to upgrade the test product to the screened food product. The bidding procedure was the same as that used in stage 1 except that there were 20 trials in stage 2. "Naive" bids were elicited in the first ten trials. The bids were naive in that the subjects were not given any information on the actual probabilities of contracting a food-borne illness from consuming the typical food product. After the tenth trial, the monitor supplied three items of information: (a) the objective probability of becoming ill from eating a year's supply of the typical food product; (b) a description of the severity of the illness; and (c) the symptoms and average medical costs of a mild case of infection. For *Salmonella*, the following information was provided (see John V. Bennett et al., 1987; Roberts, 1989):

> *Description of Salmonellosis:* Symptoms are those of a mild flu-like intestinal disease of short duration with abdominal pains, nausea, vomiting, and diarrhea. The actual individual chance of infection of salmonellosis is 1 in 125 annually. Of those individuals who get sick, 1 individual out of 1,000 will die annually. The average cost for medical expenses and productivity losses from a mild case of salmonellosis is $220.

Given this information, "informed" bids were elicited in trials 11–20.

The computer randomly selected one of the 20 trials as binding. The highest bidder paid the displayed second-highest bidding price and ate the stringently screened food. The highest bidder's take-home income was $15 minus the price paid for the screened food product. The other bidders ate the test product and took home $15. Note that the subjects had to eat the food item to leave the experiment with the take-home income.

Table 1 summarizes the experimental design for both the WTP and WTA experiments. One hundred forty-two subjects participated in the experiment. Each experimental session included between 12 and 15 subjects. All were undergraduate and graduate students from Iowa State University (ISU), recruited campus-wide. Note that a subject participated in either the WTA or the WTP experiment, not both. Also, each subject was only confronted with one food-borne pathogen description, not all five, regardless of whether that subject was in the WTA or the WTP experiment. After each subject read the instructions and answered a set of questions to test his or her understanding of the experiment and the monitor answered all relevant questions, the experiment began. All experiments were conducted in the ISU meat-testing laboratory with modern kitchen facilities. The ISU meat lab conducts food-tasting experiments on a regular basis. The lab is actively involved in all aspects of meat processing and handling, thereby providing a unique setting for our experiment.

III. Results and Discussion

Overall, we cannot reject the convergence proposition. Table 2 and Figure 2 illustrate that the WTP and WTA measures of value for the market good were not significantly different, with the exception of the first

[2]We report results for all five pathogens because measures of consumers' WTA and WTP to reduce or eliminate these pathogens are interesting in their own right (see Tanya Roberts and David Smallwood, 1991).

TABLE 1—SUMMARY OF EXPERIMENTAL DESIGN

Procedure	Experiment: Willingness to pay (WTP)	Experiment: Willingness to accept (WTA)
Stage 1 (market good; five trials; Vickrey second-price, sealed-bid auction; one trial binding):		
Initial conditions	$3 income; small piece of candy	$3 income; regular-size brand-name candy bar
Auctioned good	regular-size brand-name candy bar	small-piece of candy
Value measure	WTP to exchange piece of candy for candy bar	WTA to exchange candy bar for small piece of candy
Stage 2 (nonmarket good; 20 trials [10 naive, 10 informed]; one trial binding; Vickrey auction; five food-borne pathogens [*Campylobacter*, *Salmonella*, *Staphylococcus aureus*, *Trichinella spiralis*, *Clostridium perfringens*]):		
Initial conditions	$15 income; typical food product with average health risk from food-borne pathogens	$15 income; stringently screened food
Auctioned good	stringently screened food with one in 100 million chance of health risk from food-borne pathogen	typical food product
Value measure	WTP to exchange typical food product for screened food product	WTA to exchange screened food product for typical food product

TABLE 2—COMPARISON OF MEAN WTP AND WTA IN CANDY-BAR EXPERIMENT

H_0: WTP = WTA
H_1: WTP < WTA

Value measure	Mean ($): Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
WTP	0.40 (0.36)	0.38 (0.20)	0.40 (0.23)	0.40 (0.19)	0.39 (0.20)
WTA	0.51 (0.35)	0.44 (0.34)	0.39 (0.35)	0.37 (0.36)	0.37 (0.35)
t:[a]	−1.81**	−1.19	0.22	0.55	0.57
U:[b]	4,047.5**	4,607	5,185*	5,342**	5,332.5**

Notes: The sample size for the WTP experiments was $n = 68$; the sample size for the WTA experiment was $n = 74$. Sample standard deviations are in parentheses.

[a] One-tailed t test.

[b] Mann-Whitney U test.

*H_0 rejected at the 5-percent significance level.

**H_0 rejected at the 1-percent significance level.

trial. The mean values and standard deviations by trial are presented in the table. Repeated participation in the auction market caused the values to converge (also see Coursey et al. [1987]). Trial 1 represents the inexperienced bid analogous to the contingent-valuation method. The average WTP–WTA difference in the initial bid equaled 11 cents, and the null hypothesis that WTP and WTA were equal is rejected at the 5-percent significance level. The value disparity converged, however, to a differ-

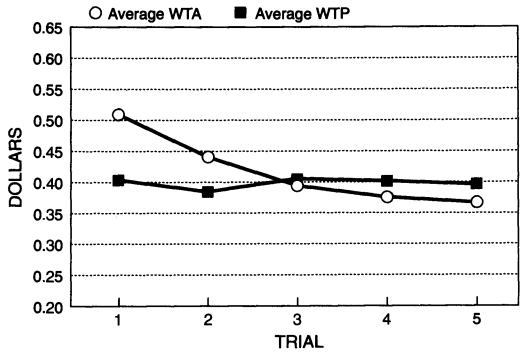

FIGURE 2. WTP AND WTA COMPARISON: CANDY BARS

ence of 6 cents in trial 2, which is not statistically significant. By trials 3, 4, and 5, the average WTP and WTA values converged to differences between 1 cent and 3 cents. We cannot reject the hypothesis of equality of the WTP and WTA measures. In addition, the results in Table 2 support Randall and Stoll's (1980) prediction that the value for goods with limitless substitution will also converge to the average market price. Because we are exchanging a small piece of candy worth about 10 cents for a candy bar worth about 50 cents, the observed average value near 40 cents is striking. With perfect substitution, value measures converged, and they converged on the average market price.

Tables 3 and 4 and Figures 3–7 illustrate that the majority of the WTA measures for

TABLE 3—COMPARISON OF WTP AND WTA FOR FIVE PATHOGENS WITHOUT ELIMINATION

Pathogen (probability of illness)	Value measure	H_0: WTP = WTA; H_1: WTP < WTA — Inexperienced (1st trial) Mean ($)	t	U	Naive (trials 7–10) Mean ($)	t	U	Informed (trials 17–20) Mean ($)	t	U
Campylobacter (1/125,143)	WTP	0.60 (0.50)	−3.65**	141**	0.71 (0.43)	−1.57*	201	0.86 (0.38)	−1.84**	228
	WTA	5.06 (4.55)			2.36 (3.89)			3.03 (4.39)		
Salmonella (1/137,000)	WTP	0.61 (0.53)	−1.20	136.5**	0.44 (0.23)	−1.15	120**	0.55 (0.25)	−2.04**	156**
	WTA	8,029 (25,957)			8.01 (25.46)			1.62 (2.00)		
Staphylococcus aureus (1/173,694)	WTP	0.97 (0.39)	−2.25**	140**	0.92 (0.32)	−1.40**	187	0.84 (0.33)	−1.04	170.5
	WTA	5.55 (7.86)			3.89 (8.19)			56.2 (205.87)		
Trichinella spiralis (1/2,628,000)	WTP	0.48 (0.42)	−1.93**	115**	0.69 (0.46)	−1.50**	155**	0.81 (0.55)	−1.31	172.5
	WTA	12.8 (24.80)			10.51 (25.36)			18.0 (50.82)		
Clostridium perfringens (1/26,280,000)	WTP	0.64 (0.63)	−2.26**	111**	0.58 (0.41)	−3.77**	109.5**	0.42 (0.33)	−4.00**	91**
	WTA	30.2 (50.56)			1.98 (1.37)			2.21 (1.70)		

Notes: Sample sizes are as follows: *Camplyobacter* (WTP = 15, WTA = 14), *Salmonella* (WTP = 15, WTA = 15), *Staphylococcus aureus* (WTP = 12, WTA = 15), *Trichinella spiralis* (WTP = 13, WTA = 15), *Clostridium perfringens* (WTP = 13, WTA = 15). Columns for t and U report results of one-tailed t tests and Mann-Whitney U tests, respectively. Sample standard deviations are in parentheses.

*H_0 rejected at the 5-percent significance level.

**H_0 rejected at the 1-percent significance level.

TABLE 4—COMPARISON OF WTP AND WTA FOR FIVE PATHOGENS WITH ELIMINATION

Pathogen (probability of illness)	Value measure	H_0: WTP = WTA; H_1: WTP < WTA — Inexperienced (1st trial) Mean ($)	t	U	Naive (trials 7–10) Mean ($)	t	U	Informed (trials 17–20) Mean ($)	t	U
Campylobacter (1/125,143)	WTP	0.53 (0.31)	−3.87**	100**	0.71 (0.36)	−1.58*	150	0.88 (0.32)	−1.61*	177
	WTA	4.63 (3.65)			1.50 (1.70)			2.29 (3.02)		
Salmonella (1/137,000)	WTP	0.55 (0.38)	−1.02	96**	0.44 (0.20)	−4.00**	91**	0.56 (0.22)	−1.91**	114**
	WTA	1,572 (5.537)			1.49 (0.92)			1.23 (1.25)		
Staphylococcus aureus (1/173,694)	WTP	1.02 (0.26)	−2.63**	100	0.97 (0.21)	−0.99	143*	0.91 (0.23)	−1.19	127.5
	WTA	4.08 (4.19)			3.12 (7.81)			3.33 (7.37)		
Trichinella spiralis (1/2,628,000)	WTP	0.44 (0.31)	−4.03**	78**	0.69 (0.44)	−2.32**	110*	0.82 (0.51)	−2.26**	128.5
	WTA	7.08 (5.93)			4.43 (5.79)			5.42 (7.33)		
Clostridium perfringens (1/26,280,000)	WTP	0.57 (0.49)	−3.47**	78**	0.60 (0.38)	−3.67**	76.5**	0.43 (0.32)	−4.11**	66**
	WTA	19.4 (19.51)			1.83 (1.14)			2.00 (1.34)		

Notes: Sample sizes are as follows: *Camplyobacter* (WTP = 13, WTA = 12), *Salmonella* (WTP = 13, WTA = 13), *Staphylococcus aureus* (WTP = 10, WTA = 13), *Trichinella spiralis* (WTP = 11, WTA = 13), *Clostridium perfringens* (WTP = 11, WTA = 13). Columns for t and U report results of one-tailed t tests and Mann-Whitney U tests, respectively. Sample standard deviations are in parentheses.

*Denotes rejection of H_0 at the 5-percent significance level.

**Denotes rejection of H_0 at the 1-percent significance level.

the nonmarket good significantly exceed the WTP measures. This holds for both the naive bids (trials 7–10) and the informed bids (trials 17–20). Note that the WTP and WTA measures for each pathogen are examined with two mean values: without elimination of the highest and lowest bids and with elimination. Again mean values and standard deviations are presented in the tables. We consider elimination to explore Robin Gregory and Lita Furby's (1987) argument that values are extremely sensitive to one or two outliers (also see Robert C. Mitchell and Richard T. Carson [1989]). This work reexamined Coursey et al.'s (1987) sucrose octa-acetate (SOA) experiment with elimination of outliers and found that the results of value convergence depend on inclusion of an outlier. To illustrate the robustness of our results, we consider values with and without the elimination of outliers.

Means of the WTP experiment without elimination of outliers closely coincided with the means with elimination. In the WTA experiment, outliers change the majority of the mean values, especially for *Salmonella*, *Staphylococcus aureus*, and *Trichinella spiralis*. For the initial bid in trial 1, we observed extremely high WTA values. For *Salmonella*, the mean WTA is more than 13,000 times greater than the mean WTP without elimination and is still 3,000 times greater with elimination. WTA for *Clostridium perfringens* is 47 times greater than WTP without elimination. WTA divergence for the other pathogens ranges from four to

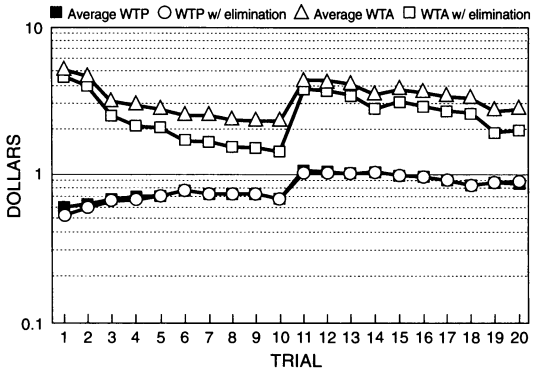

FIGURE 3. WTP AND WTA COMPARISON: *Camplyobacter*

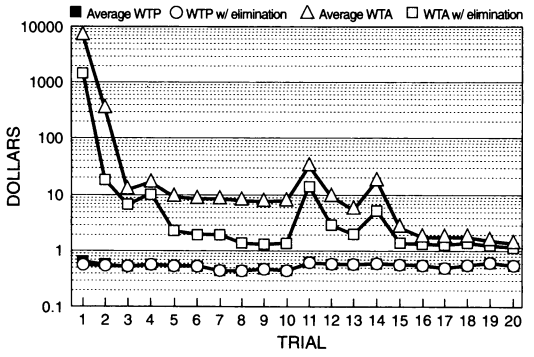

FIGURE 4. WTP AND WTA COMPARISON: *Salmonella*

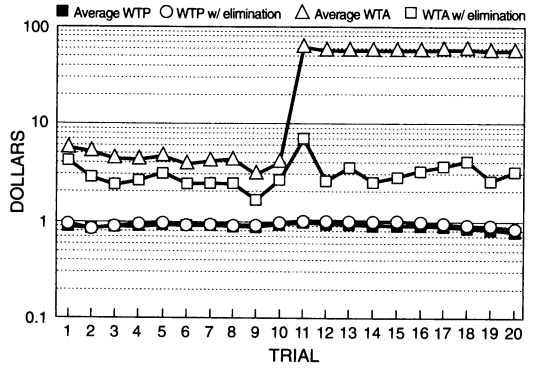

FIGURE 5. WTP AND WTA COMPARISON: *Staphylococcus aureus*

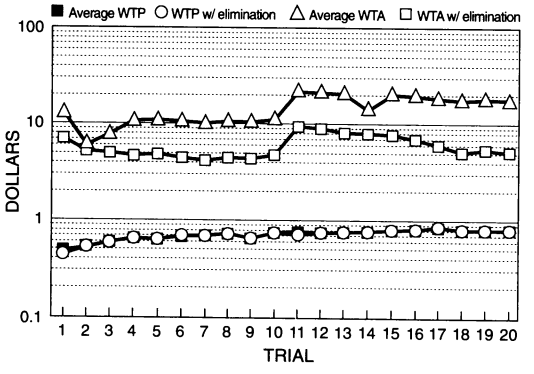

FIGURE 6. WTP AND WTA COMPARISON: *Trichinella spiralis*

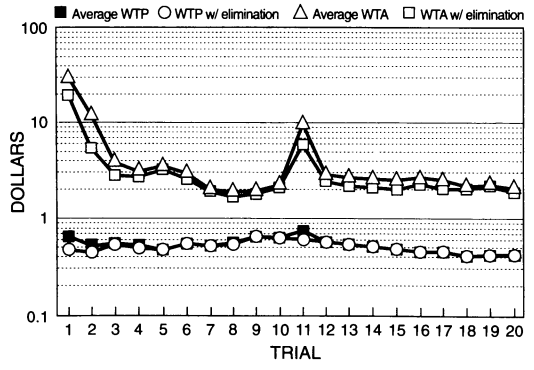

FIGURE 7. WTP AND WTA COMPARISON: *Clostridium perfringens*

34 times greater than that for WTP in trial 1. For the initial bid, we performed a one-tailed t test and the Mann-Whitney rank-sum U test to test the significance of the divergence of WTP and WTA. According to the rank-sum test, the null hypothesis of all pathogens, that WTP and WTA values are from the same parental population, is rejected at the 5-percent significance level.

For most of the naive bids (trials 7–10), the average bidding prices stay relatively constant in both the WTP and the WTA bids. This result is consistent with Coursey's (1987) observation that Vickrey auctions usually stabilize by the sixth or seventh trial. The mean WTA for trials 7–10 ranges from approximately three times greater than that

of the mean WTP for *Campylobacter* to approximately 18 times greater than that of the mean WTP for *Salmonella* without elimination. With elimination of outliers, the results indicate that the mean WTA is from two to six times greater than the mean WTP. The disparities between WTP and WTA for each pathogen are tested by performing a multivariate analysis[3] and a *U* test. Although the WTP and WTA experiments are statistically independent, we used multivariate analysis to account for the between-trial correlation among bids from the same subjects. For most pathogens, the difference between WTP and WTA in the naive bids is statistically significant, based on *t* tests, both with and without elimination of outliers. The WTP–WTA difference for the *Salmonella* experiment is statistically insignificant according to a test without elimination, but significant with elimination.

For the informed bids (trials 17–20), we observed that bids initially increased from the information shock. The WTP experiments have a smaller increase relative to the WTA experiments. Again, after six trials with information, the mean WTP bid stabilizes. Mean WTA bids converge to lower values, with some variation in the last two or three trials. For trials 17–20, the differences between WTP and WTA range from threefold to fivefold for *Salmonella*, *Campylobacter*, and *Clostridium perfringens*. The WTP and WTA bids for these three pathogens are statistically significant with and without elimination of outliers. *Trichinella spiralis* bids were significantly different with elimination, but not different without. *Staphylococcus aureus* bids were not significantly different, either with or without elimination.

In sum, we cannot reject the convergence proposition. For the market good with close substitutes, WTP and WTA measures of value are not statistically different with repeated market exposure. In contrast, for the nonmarket good with imperfect substitutes, WTP and WTA measures are significantly different, even after repeated market participation and with full information about the probability and severity of the health risk. Our results support Hanemann's (1991) proposition that the degree of substitution drives the divergence between value measures.

IV. Substitution versus Endowment Effects

An alternative explanation of the WTP–WTA divergence has been put forth by Kahneman et al. (1990). They argue that there may be a fundamental "endowment effect" underlying the theory of choice (also see Knetsch [1989]). The endowment effect exists when an individual becomes attached to the good because he or she is often rewarded for doing so in many contexts. This attachment induces the individual to demand a higher level of compensation than he or she was originally willing to pay.

The inability to substitute goods may be the underlying motivation behind Kahneman et al.'s (1990) observations of an endowment effect. They recognize this possibility, stating that "... endowment effects will almost certainly occur when owners are faced with an opportunity to sell an item purchased for use that is not easily replaceable" (p. 1344). If the endowment effect was not driven by substitutability, we should have observed a divergence in value measures for the candy bar, which we did not.

[3]Let X_{ijk} be the subject's kth bid in the jth trial of the ith group, with $i=1,2$ ($i=1$, WTP experiment; $i=2$, WTA experiment), $j=7,8,9,10$ (trial), and $k=1,2,\ldots,n_i$ (number of subjects in experiment). Because X_{ijk} and $X_{ij'k}$ ($j\neq j'$) are not independent (measured repeatedly), multivariable analysis or split plot design can be applied. Suppose vector $\mathbf{X}_i \equiv (X_{i,7}, X_{i,8}, X_{i,9}, X_{i,10})'$ ($i=1,2$) $\sim$ MVN($\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i$) where $\boldsymbol{\mu}_i = (\mu_{i,7},\ldots,\mu_{i,10})'$ and where $\boldsymbol{\Sigma}_i$ is the corresponding variance-covariance matrix. Consider $Y_i \equiv \mathbf{a}'\mathbf{X}_i$ where $\mathbf{a}' = \frac{1}{4}(1,1,1,1)'$ ($i=1,2$). Then $Y_1 = \frac{1}{4}(X_{1,7} + X_{1,8} + X_{1,9} + X_{1,10})$ and $Y_2 = \frac{1}{4}(X_{2,7} + X_{2,8} + X_{2,9} + X_{2,10})$ are normally distributed with mean $\mathbf{a}'\boldsymbol{\mu}_1$ and $\mathbf{a}'\boldsymbol{\mu}_2$ and variance $\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}$ and $\mathbf{a}'\boldsymbol{\Sigma}_2\mathbf{a}$, respectively. Because Y_1 and Y_2 are independent, $(Y_1 - Y_2)$ is normally distributed with mean $(\mathbf{a}'\boldsymbol{\mu}_1 - \mathbf{a}'\boldsymbol{\mu}_2)$ and variance $(\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a} + \mathbf{a}'\boldsymbol{\Sigma}_2\mathbf{a})$. There are n_1 and n_2 samples from the WTP and the WTA experiments, respectively (i.e., using $y_{1,1},\ldots y_{1,n_1}$ and $y_{2,1},\ldots y_{2,n_2}$).

To test the null hypothesis that there is a difference between the WTP and WTA experiments, we can use the *t* test for the difference of the mean between the WTP and WTA experiments (see Richard A. Johnson and Dean W. Wichern, 1988).

At the request of a reviewer, we ran an additional experiment to explore more rigorously the relationship between the substitution effect and the endowment effect. Kahneman et al.'s work suggests that perhaps an endowment effect can exist for a good with available substitutes if the subjects have less information on its value and cost. For a Cornell coffee mug available at the campus store, Kahneman et al. (1990) observed sellers demanding twice the compensation that buyers were willing to pay, a pattern that remained unchanged with market experience (see their tables 3 and 4, pp. 1334–35). Because we did not observe an endowment effect for our candy-bar treatments, this suggests a potential threshold where the importance or size of the endowment drives behavior.

To test for a possible endowment effect we repeated our experiment, but the auctioned good was now an Iowa State coffee mug available from the campus bookstore for \$5.20. All design aspects followed as closely as possible to our earlier experiments: a Vickrey auction, 15 subjects per treatment, mugs placed directly in the subject's hands, a \$15 initial endowment, and repeated market participation (ten trials for the ISU mug stage). In addition, we replicated the initial candy-bar stage (\$3 initial endowment and five trials) to maintain consistency with the food-borne pathogen treatments. The baseline treatment was a WTP auction that allowed the highest bidder to upgrade his or her plain plastic mug (worth \$1.60) for an ISU mug. We compared two WTA treatments to the WTP baseline. The first treatment was a WTA auction where the low bidder would receive compensation (second-lowest bid) for exchanging his or her ISU mug for a plain mug. The second treatment was identical to the first, except now subjects were told they could purchase (at the market price) an unlimited number of ISU mugs right outside the door immediately after the experiment. This eliminated any transaction costs that could create friction in trade. Our hypothesis is summarized below.

PROPOSITION 2 (Endowment Proposition): *If the endowment effect exists, then the willingness to accept compensation will exceed the willingness to pay for the market good (Iowa State coffee mug) with or without zero transactions costs.*

If we observe a persistent divergence of value measures with repeated market experience, a fundamental endowment effect may well exist in goods that have available substitutes but have some degree of uncertainty regarding value and cost. However, if we observe a convergence of value measures for both treatments, we can reject the endowment proposition. Our results then cast doubt on the generality of Kahneman et al.'s evidence and provide more support for Hanemann's argument on the importance of the substitution effect.

Overall, we reject the endowment proposition. Table 5 and Figure 8 illustrate the results for all subjects. Although for the initial trial mean WTA exceeds mean WTP by ratios of 2.74 and 2.76 to 1 for the two treatments, by trial 4 the ratio nearly equals unity: 1.03 and 1.06. Clearly, the value measures converged, a pattern generally consistent in the remaining trials. Considering the last three experienced trials (trials 8–10), the WTA/WTP ratio equals 1.08 and 1.05 for the two treatments. We cannot reject the null hypothesis of equality of WTP and WTA either without or with elimination of outliers at the 5-percent level (without elimination, $t = -0.2412$ and $t = -0.1509$ for treatments 1 and 2; with elimination, $t = 0.1768$ and $t = 0.3813$ for treatments 1 and 2). This holds for the Wilcoxon rank-sum test as well ($Z = 0.996$, $Z = 0.975$, $Z = 1.103$, and $Z = 1.206$ for treatments 1 and 2 without and with elimination).[4]

We do not observe a fundamental endowment effect under our experimental procedures, contrary to the findings reported by Kahneman et al. (1990). There are numerous differences between our experimental

[4]Note that the stage-1 candy-bar experiments followed an identical pattern to the earlier runs. In trial 1, the average WTA was \$0.73 while the average WTP was \$0.38. But by trial 3, average WTA and WTP equaled \$0.37 and \$0.36. Again, values converged, and they converged to the differential market price.

TABLE 5—COMPARISON OF MEAN WTP AND WTA BIDS (IN DOLLARS) IN MUG TREATMENTS

Treatment	n	Trial 1	2	3	4	5	6	7	8	9	10
WTP, baseline without mug sale	15	2.37 (1.484)	2.40 (1.275)	2.89 (1.111)	2.90 (1.072)	2.70 (1.047)	2.86 (1.139)	2.74 (1.022)	2.62 (1.291)	2.68 (1.187)	3.09 (1.357)
WTA, without ISU mug sale	15	6.55 (4.894)	4.17 (3.261)	3.19 (2.990)	3.06 (2.401)	2.88 (1.903)	4.14 (5.125)	3.53 (3.411)	3.25 (3.632)	3.12 (2.522)	2.30 (1.970)
WTA, with ISU mug sale	15	6.50 (4.452)	4.74 (4.147)	3.67 (3.202)	2.98 (3.104)	3.27 (3.262)	3.75 (4.331)	2.65 (2.746)	3.13 (3.010)	3.29 (3.135)	2.49 (2.175)

Note: Sample standard deviations are in parentheses.

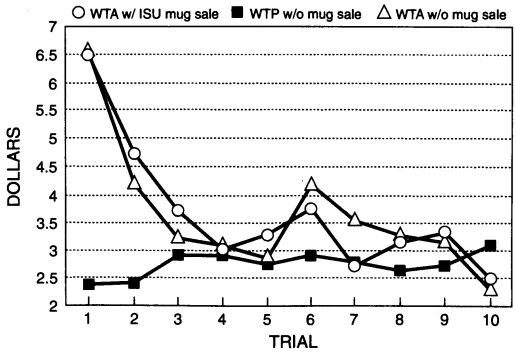

FIGURE 8. WTP AND WTA COMPARISON: ISU MUGS (AVERAGE VALUES)

design and that of Kahneman et al., and our use of the Vickrey auction may play a role (see e.g., John H. Kagel et al. [1987] on overbidding). Regardless of why our results differ from Kahneman et al., we observe the convergence of WTP and WTA values for the candy-bar and coffee-mug treatments but a persistent divergence in the values for the food-borne pathogen treatments. Hanemann's (1991) substitution effect appears to organize observed valuation behavior in a predictable fashion. Future research should continue to compare how alternative elicitation mechanisms affect revealed values in experimental markets, holding constant monetary incentives, opportunity for market experience, subject pools, and the auctioned or traded good.

V. Conclusion

The divergence in WTP and WTA measures of value has troubled economists for the past decade. The divergence led Ronald G. Cummings et al. (1986) to recommend in their "reference operating conditions" (pp. 102–9) for contingent valuation of environmental goods that only WTP measures be elicited in the attempt to value nonmarket goods. Hanemann (1991) has offered an explanation grounded in economic theory that may calm the fears that the divergence is some form of cognitive mistake. Our experimental results support his argument that the degree of substitutability between goods may drive the difference between WTA and WTP measures of value. For market goods with close substitutes which are readily available in commercial outlets with minimal transaction costs (i.e., candy bars and coffee mugs), we find that WTP and WTA value measures converge. In contrast, for a nonmarket good with no close substitutes (i.e., reduced health risk), the value measures diverge and persist, even with repeated market participation and full information on the nature of the good.

Further research on discrepancies between willingness-to-pay and willingness-to-accept measures of valuation is clearly needed. First, researchers should replicate our experiment to test the robustness of our findings. Second, researchers should improve the understanding of substitutability when conducting field studies in nonmarket valuation. As shown by Brookshire and Coursey (1987), wild discrepancies in measures of value may well be only an artifact of survey methods that do not employ repeated market experience and real payments. However, with truly unique private or public goods that have limited substi-

tutes, Hanemann predicts a wide divergence of value. Defining an individual-specific index of substitution for the nonmarket good in question could improve the correspondence between economic theory and observed phenomena.

APPENDIX
Experimental Instructions [Exact Transcript]

GENERAL INSTRUCTIONS

You are about to participate in an experiment about decision making. Please follow the instructions carefully. The United States Department of Agriculture has provided funds for this research.

SPECIFIC INSTRUCTIONS

In this experiment, you will be asked to decide how much you would be willing to pay for safer food [to decide the minimum amount you would be willing to accept for taking the test product food, instead of keeping your safer food]. The experiment has two stages.

Your starting income will be \$3 in stage 1. Your income will be \$15 for stage 2. Your take-home income will consist of your initial income (\$3 + \$15) minus [plus] the values of goods purchased.

You will submit your bidding price on a recording card. Note only one of the trials in stage 1 will be binding and only one of the twenty trials in stage 2 will be binding. A number will be randomly selected to identify these binding trials.

You cannot reveal your bids to any other participant. Any communication between bidders during a trial will result in an automatic penalty of \$3.

STAGE 1

Step 1: You *own* the candy [candy bar] free in front of you. Your initial income is \$3.

Step 2: Let's say you are willing to pay \$X for the piece of candy and \$Y for a candy bar. The *difference* (\$Y − \$X) is what you are willing to pay to upgrade your piece of candy for a candy bar. Please indicate your willingness to pay to trade the piece of candy for a candy bar. Do *not* state what you would pay for an entire candy bar. Only state the *difference* (\$Y − \$X) you are willing to pay.

Step 3: Please write your bid (difference) for the *one* candy bar on the recording card. The monitor will announce the highest bidder and display the price of the candy bar (*second-highest bidding price*) on the blackboard.

Note: For example, if the highest bid was $\$\alpha$ and the second-highest was $\$\beta$, the highest bidder would receive the candy bar and must pay $\$\beta$.

Step 4: There will be five trials.

Step 5: Only one trial will be binding. After the five trials, a number will be randomly selected to determine which trial is binding. The highest bidder of that trial will exchange the piece of candy for the candy bar and must pay the displayed price (i.e., *the second-highest bid*).

Note: In the event that there is a tie for the highest bid, those participants will be asked to bid again.

Questions

Please answer the following questions, which are designed to help you understand stage 1. Do not hesitate to ask the researchers if you have questions.

1. Suppose that person A is the highest bidder in the first trial, person B is the highest bidder in third trial, and person C is the highest bidder in fifth trial. If, after five trials are finished, we randomly select the third trial, then who will purchase the candy bar?_____
2. If your $\$\alpha$ bid is the highest in the third trial, and the second-highest bid is $\$\beta$, what price will you pay for the candy bar? $_____
3. If your bid is not the highest in the third trial, which is randomly selected, how much should you pay for the piece of candy? $_____

STAGE 2

Step 1: There are two types of food. The features are each described below.

Test Product	**Stringently Screened**
This food has a typical chance of being contaminated with the food-borne pathogen *Salmonella*; i.e., it is purchased from a local source.	This food has been subjected to stringent screening for *Salmonella*. There is a *1 in 100,000,000* chance of getting salmonellosis from consuming

Step 2: You *own* a test product sandwich free in front of you. Everyone has the same sandwich. You also have initial income, $15.

Step 3: Let's say you are willing to pay $X for the test product sandwich and $Y for the stringently screened sandwich. The *difference* ($Y − $X) is what you are willing to pay to reduce the risk of illness from the food-borne pathogens. Please indicate your willingness to pay to reduce the risk of illness. Do *not* state what you would pay for the entire stringently screened sandwich. Only state the *difference* ($Y − $X) you are willing to pay. The highest bidder will upgrade his or her test product sandwich for the stringently screened sandwich. He or she will pay *the second-highest bidder's price*.

Step 4: There will be twenty trials.

Step 5: After all twenty trials are complete, we will randomly select *one* binding trial to determine who buys the stringently screened food.

Note: The sandwich has to be eaten to leave with the take-home income.

Questions

Please answer the following questions, which are designed to help you understand stage 2. Do not hesitate to ask the researchers if you have questions.

1. There are twenty bidding trials. If person A is the highest bidder in the first trial, person B is the highest bidder in the eighteenth trial, and the eighteenth trial is selected, then who will receive the stringently screened food?_____
2. If your $\$\alpha$ bid is the highest in the eighteenth trial, and the second highest bid is $\$\beta$, what price will you pay for the stringently screened food? $_____

Note: Please answer the questions below.

1. What do you think is the chance of becoming ill from *Salmonella*, given that you eat an average amount of typical food products in the United States *over one year*?

Answer: _____chance out of 1 million people

2. What do you think are the important sources of the food-borne pathogen, *Salmonella*, in the United States?

_____ _____ _____

Please list the type of food items.

Information for Trials 11-20

Test Product

If you eat this food, there is a *1 in 137,000* chance that you will become ill from *Salmonella*.

Stringently Screened

This food has been subjected to stringent screening for *Salmonella*. There is a *1 in 100,000,000* chance of getting salmonellosis from consuming this food.

Description of Salmonellosis:

Symptoms are those of a mild "flu-like" intestinal disease of short duration with abdominal pains, nausea, vomiting, and diarrhea. The actual individual choice of infection of Salmonellosis is 1 in 125 annually. Of those individuals who get sick, 1 individual out of 1,000 will die annually. The average cost for medical expenses and productivity losses from a mild case of Salmonellosis is $220.

REFERENCES

Bennett, John V.; Holmberg, Scott D.; Rogers, Marta F. and Solomon, Steven L. "Infectious and Parasitic Diseases," in R. W. Amler and H. B. Dull, eds., *Closing the gap: The burden of unnecessary illness*. New York: Oxford University Press, 1987, pp. 102–14.

Brookshire, David S. and Coursey, Don L. "Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures." *American Economic Review*, September 1987, *77*(4), pp. 554–66.

Coppinger, Vikki; Smith, Vernon L. and Titus, Jon. "Incentives and Behavior in English, Dutch and Sealed-Bid Auctions." *Economic Inquiry*, January 1980, *18*(1), pp. 1–22.

Coursey, Don L. "Markets and the Measurement of Value." *Public Choice*, October 1987, *55*(3), pp. 291–97.

Coursey, Don L.; Hovis, John J. and Schulze, William D. "The Disparity between Willingness to Accept and Willingness to Pay Measures of Value." *Quarterly Journal of Economics*, August 1987, *102*(3), pp. 679–90.

Cummings, Ronald G.; Brookshire, David S. and Schulze, William D. *Valuing environmental goods: An assessment of the contingent valuation method*. Totowa, NJ: Rowman and Allanheld, 1986.

Gregory, Robin and Furby, Lita. "Auctions, Experiments and Contingent Valuation." *Public Choice*, 1987, *55*(3), pp. 273–89.

Hanemann, W. Michael. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review*, June 1991, *81*(3), pp. 635–47.

Hammack, Judd and Brown, Gardner M., Jr. *Waterfowl and wetlands: Toward bioeconomic analysis*. Baltimore, MD: Johns

Hopkins University Press, 1974.

Johnson, Richard A. and Wichern, Dean W. *Applied multivariate statistical analysis*, 2nd Ed. Englewood Cliffs, NJ: Prentice Hall, 1988.

Kagel, John H.; Harstad, Ronald M. and Levin, Dan. "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study." *Econometrica*, November 1987, *55*(6), pp. 1275–1304.

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.

Kentsch, Jack. "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *American Economic Review*, December 1989, *79*(5), 1277–84.

Knetsch, Jack L. and Sinden, John A. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value." *Quarterly Journal of Economics*, August 1984, *99*(3), pp. 507–21.

Layard, P. R. G. and Walters, A. A. *Microeconomic theory*. New York: McGraw-Hill, 1978.

Mitchell, Robert C. and Carson, Richard T. *Using surveys to value public goods: The contingent valuation method*. Washington, DC: Resources for the Future, 1989.

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

Roberts, Tanya. "Human Illness Costs of Food-Borne Bacteria." *American Journal of Agricultural Economics*, May 1989, *71*(2), pp. 468–74.

Roberts, Tanya and Smallwood, David. "Data Needs to Address Economic Issues in Food Safety." *American Journal of Agricultural Economics*, August 1991, *73*(3), pp. 933–43.

Rowe, Robert D.; d'Arge, Ralph C. and Brookshire, David S. "An Experiment on the Economic Value of Visibility." *Journal of Environmental Economics and Management*, March 1980, *7*(1), pp. 1–19.

Shogren, Jason F. "The Impact of Self-Protection and Self-Insurance on Individual Response to Risk." *Journal of Risk and Uncertainty*, June 1990, *3*(2), pp. 191–204.

Vickrey, William. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance*, March 1961, *16*(1), pp. 8–37.

Willig, Robert. "Consumer's Surplus Without Apology." *American Economic Review*, September 1976, *66*(4), pp. 589–97.