Effects of Elicitation Method on Willingness-to-Pay: Evidence from the Field

Jared G. Carlberg and Eve J. Froehlich

This paper compares willingness-to-pay (WTP) estimates elicited using three separate methods: in-store experimental auctions, a mailed survey with a cheap talk script included, and the same survey with no script. The products in question were four steaks bearing hypothetical brands representing various brandable attributes. It is found that WTP elicited using experimental auctions was the lowest, followed by WTP elicited by the mail survey with a cheap talk script, then the mail survey with no cheap talk script. Tobit and double-hurdle econometric models are used to identify factors influencing respondents' WTP; model results are largely consistent with previous findings.

A considerable number of empirical studies have been undertaken to determine willingness-to-pay (WTP) for a variety of products. Methods for eliciting WTP have evolved considerably since the earliest work in the field, resulting in a rich variety of experimental auction and contingent valuation methods now being available to researchers. Differences in cost and complexity among elicitation methods are considerable, and as a result, a wide variety of methods are commonly used in both the field of experimental economics and for practical purposes in market research.

The body of literature exploring the effects of elicitation method upon WTP is vast. Extensive work has been undertaken to develop techniques to ensure precise estimates of WTP. However, some issues relating to effects of elicitation method upon WTP estimates remain unexplored; for example opportunities remain to compare results from field experiments with those from a mail survey using a common product. Additionally, the literature comparing econometric estimates based on data from different elicitation methods is relatively sparse.

There are three principal objectives of the research reported in this paper. The first is to determine the effect of elicitation method upon stated WTP for brand-name steaks in Canada. No researchers have previously attempted to compare WTP estimates from experimental auctions with those from mail surveys. The second objective is to quantify the effects of "cheap talk" (i.e. instructions regarding the presence and reasons for overstatement of bids in a hypothetical buying situation) in reducing hypothetical bias. If the form of experimental auction chosen is incentive compatible, then a comparison of bids from these auctions with results from a survey incorporating cheap talk should illustrate the extent to which cheap talk mitigates hypothetical bias. The third and final objective is to model the factors affecting WTP for the fictional steak brands developed for the research. The finding that econometric models based upon experimental auction vs. survey results have substantially different results has important implications from a research perspective.

Five sections comprise the remainder of this paper. First, the fictional steak brands developed for this research are described. An overview of the survey and experimental auction methodologies is then provided, followed by a section outlining the theory behind and econometric procedure used to estimate the WTP models. Results of the experimental auctions and surveys are then presented and discussed, along with estimates from the econometric models. The paper's concluding section summarizes findings, acknowledges limitations, and makes suggestions for future work.

Carlberg is Associate Professor in the department of Agribusiness & Agricultural Economics at the University of Manitoba. Froehlich is a Policy Analyst for the Canadian Grain Commission.

The authors gratefully acknowledge support from the National Beef Industry Development Fund of the Canadian Cattlemen's Association, along with the helpful comments of three reviewers and the Journal editor.

Steak Brands

Froehlich, Carlberg and Ward (2009) observe there are almost no fresh brand-name beef offerings in Canada, a marked difference from the

U.S. case. Accordingly, it was necessary to develop fictional brands to assess WTP. Four steak brands were developed for this research: a local/Canadian brand, a guaranteed tender brand, a natural beef brand, and an Angus brand. Froehlich, Carlberg and Ward (2009) provide further details on the process used to develop the brands and provide complete descriptions for each. Logos for the brands are given in Figure 1. WTP for Canada AAA beef, the second-highest grade of beef available (only 2% of beef grades above AAA) was also elicited from auction participants and survey respondents.

(a) Prairie Prime logo

(b) Tender Grill logo

(c) Nature's Diamond logo

(d) Original Angus logo

Figure 1. Steak Brand Logos

Prairie Prime was the fictional local Canadian brand developed for the research. Steaks bearing this brand were described as being born and raised in the Canadian Prairies and were certified to grade AAA (equivalent to USDA choice) or higher. Additionally, Prairie Prime beef was guaranteed to be aged a minimum of 14 days. The guaranteed tender brand used to measure consumer WTP was called *Tender Grill*. Participants were informed this brand of beef was tested for tenderness using the Warner-Bratzler shear force test, which allowed the steaks to be certified as tender. The steaks were also described as having been aged a minimum of 21 days, but no grade guarantee was given.

The beef brand marketed as being derived from cattle never given growth hormones or antibiotics was called *Nature's Diamond*. This "natural" beef brand also claimed animals were pasture-fed for 15 months prior to 120 days of grain finishing, were raised using environmentally friendly production methods, and received only chemical-free, natural feed and clean water. No guaranteed minimums for grade or dry aging time were provided. *Original Angus* was the fictional brand used in this research to represent the breed-specific characteristic that has enjoyed widespread demand for many years in the U.S. Steaks bearing this brand were positioned as being derived from beef that was grain fed, dry aged a minimum of 14 days, and verifiably Angus in origin. Emphasis was placed upon careful selection of animals for inclusion in the branding program as well as upon involvement of both the national Angus breed association and independent federal inspectors in monitoring and inspection along the supply chain.

Experiment and Survey Methodology

The use of experimental auctions to elicit WTP has risen as the selection of contemporary auction mechanisms has grown. An experimental auction can be used to elicit a participant's WTP in a manner that is designed to reflect their true valuation of a product. In an experimental auction, measurement of WTP involves the use of actual money; this distinguishes the method from the hypothetical situation that exists when a survey method is used (Lusk et al 2001).

An experimental auction is *incentive compatible* if it elicits values that reflect participants' true WTP. There are numerous different experimental auction designs available to researchers, and incentive compatibility is often of paramount importance in choice of design. For this condition to be met, the participant must have an incentive not to over or understate their bid: if a participant understates their bid, they risk not purchasing a product that is valuable to them, whereas overstatement may result in a situation where the respondent is forced to purchase a product for more than it is worth to them (Feldkamp, Schroeder and Lusk 2005; Umberger and Feuz 2004).

Various types of experimental auctions can be used by researchers to elicit WTP. One of the most popular in recent years has been the Becker-DeGroot-Marshack (BDM) (1964) method. A number of agricultural economists have used this technique; examples include Feldkamp, Schroeder and Lusk (2005); Lusk et al (2001); and Lusk and Fox (2003). In a BDM auction, participants do not bid against one another; rather they evaluate a particular good then are asked to submit a bid. If their bid is greater than a randomly selected price, the participant is obliged to pay *the randomly selected price* (i.e. not the amount they bid) for the item. BDM auctions are incentive compatible.

The BDM method was chosen for this research for a number of reasons. First, the approach is easy to explain to participants and it is easy for them to understand relative to other auction designs (Lusk et al 2001). The BDM auction does not take repeated practice rounds for participants to learn how the auction works. Second, BDM auctions tend to have fewer non-responses and thus less non-response bias than other auction mechanisms and certainly less than contingent valuation (Lusk et al 2001). The BDM design has fewer non-responses because of ease of participation. Participants do not have to go out of their way on second day and drive to a location where another type of experimental auction would be conducted in a group setting (Feldkamp, Schroeder and Lusk 2005). In other words, there is less opportunity cost for the participants to partake in the study than in other experimental auction procedures. Response rates are also generally higher than when contingent valuation is used and a mail survey is simply sent out.

The major alternative to experimental auctions is contingent valuation, a popular method used to elicit willingness-to-pay values from consumers. In agribusiness applications, typically a novel product is described in detail and the participant is asked to state hypothetically in monetary terms how much they would be willing to pay for the good in question or are asked whether they are willing-to-pay a stated amount for the good. When consumers make decisions about what goods to purchase, they evaluate the utility of the attributes of each of the goods and maximize their expected utility by choosing a good with the optimal combination of attributes (Ness and Gerhardy 1994). Sometimes the consumer must make trade offs to achieve the most important attributes they desire in a good. For example, if a consumer's primary concern is a 'natural' beef product, the consumer must be willing to trade off a low price attribute to obtain the 'natural' beef product. Contingent valuation is used to determine a consumer's most preferred attributes and ultimately most preferred goods. This type of method has been used by agricultural economists such as Neill et al (1994), Brown et al (1996) and Loomis, Gonzalez-Caban and Gregory (1996), among others.

The use of a "cheap talk" script involves instructing participants to respond as if they were making a real-world rather than hypothetical purchasing decision. Much of the early work behind cheap talk originated with Loomis, Gonzalez-Caban and Gregory (1996), though they neither coined the term nor actually developed cheap talk itself. Using an open-ended survey instrument, they requested the subjects refrain from bidding what they thought to be the fair market value of the good and instead bid as though they were in a real market and actually had the opportunity to buy the good. Additionally, participants were asked to take their budget constraints into consideration when formulating their bid. Despite these efforts, the authors were unable to demonstrate that that these reminders were effective in eliminating hypothetical bias. Later work by Cummings and Taylor (1999) introduced cheap talk the way most researchers

use it today, and List (2001, 2003) extended the applicability of Cummings and Taylor's cheap talk method to a real functioning market as opposed to a classroom setting.

Data for this study were obtained using 274 BDM experimental auctions carried out in seven grocery stores from two major chains in and around the city of Winnipeg, Manitoba. Auctions were conducted at various times of the day during normal store hours on both weekdays and weekends and about 75% of the persons invited to participate agreed to do so. Each auction, conducted near the meat counter in participating grocery stores, took between five and seven minutes to complete after participants had read a two-page information sheet detailing the steak brands and their attributes. Additional information on the auction methodology and information provided to auction participants can be found in Froehlich, Carlberg and Ward (2009).

A mailed survey of 5,100 recipients was also conducted on a random sample of Canadian consumers excluding Quebec (due to its large francophone population) and the three Canadian territories (mostly only frozen beef is available in the northern territories). The survey was designed to be as similar to the BDM auction as possible. Random addresses were purchased from a reputable market research firm. A Canadian one-dollar coin was included with the cover letter of each survey to provide an incentive to complete and return the questionnaire. A reminder postcard was mailed to recipients three weeks following the original survey mailing.

Survey recipients also received a "Steak Fact Sheet" which described the hypothetical brands; this was the same information that was provided to experimental auction participants. Recipients were requested to complete a two page questionnaire and mail it back in the prepaid postage envelope provided. Two separate treatments of the survey were used: the first contained, in addition to the survey package (cover letter, business reply envelope, steak fact sheet and survey instrument), an information sheet discussing how people tend to overstate their willingness-to-pay for products and services in a hypothetical setting (this is referred to as the "cheap talk script"). This cheap talk script was identical to the one used in Lusk (2003) who made small modifications to the original devised by Cummings and Taylor (1999). The script provides an overview of the problem of hypothetical bias, discusses why it may occur and requests that the respondent avoid hypothetical bias when completing the survey. The other survey treatment simply received no cheap talk script.

Theory and Econometric Procedure

Lancaster (1966) observed that a major weakness of consumer demand theory at the time was that it omitted consideration of a good's intrinsic properties. He noted that it is a good's characteristics, rather than the good itself, that determine its value. This concept was extended by Ladd and Martin (1976), who modeled demand for a good as a function of a product's characteristics, along with prices and income. A substantial number of subsequent empirical price analyses have employed this type of framework.

Hedonic price functions are commonly used to empirically estimate the relationship between a good's price and its attributes. Goodman (1998) observes that although earliest use of this type of function is often attributed to Griliches (1958), in fact Court's (1939) work on automobile price indices was the first use of such functions. A simple yet accurate description of hedonic price functions is provided by Nesheim (2006), who observes they "…describe the equilibrium relationship between the economically relevant characteristics of a product or service (or bundle of products) and its price." This definition seems in line with the Lancastrian framework outlined above, and so a hedonic function is employed in this paper to model WTP. Hedonic models are sometimes criticized for a variety of reasons, including the so-called "adding up" problem whereby the value of all attributes, when summed, does not equal a product's price. The reader is thus urged to interpret results with caution.

The model used to describe the relationship between stated WTP and the attributes of the branded steak in question as well as the demographic characteristics of participants is

(1) $WTP_{ij} = \alpha + \beta_1 \times beef\ eaten + \beta_2 \times like\ name + \beta_3 \times confidence + \beta_4 \times gender + \beta_5 \times age + \beta_6 \times income + \beta_7 \times education + e,$

where WTP_{ij} gives the willingness-to-pay for the j^{th} steak stated by the i^{th} respondent, *beef eaten* is the number of times per week the respondent eats beef, *like brand* is the respondent's rating of how much they like the fictional brand on a 7-point Likert scale, *confidence* is the respondent's self-assessment of their ability to select beef on a 7-point Likert scale, the remainder of variables are demographic characteristics of respondents, and *e* is the error term. All of age, income and education were measured as categorical variables.

Because auction participants and survey respondents were not allowed to state negative amounts for premiums, the WTP data are left-censored (censored from below) (Lusk and Shogren 2007). As such, any econometric procedure used to estimate equation (1) must take this characteristic of the data into account; failure to do so could result in biased estimates (Amemiya 1973). The tobit model (Tobin 1958) can be used when left-censored data are encountered. This method explains the relationship between a non-negative latent dependent variable and one or more independent variables, and unlike ordinary least squares, takes explicit account of the limited nature of the dependent variable, yielding unbiased parameter estimates.

Cragg's (1971) double-hurdle model recognizes that left-censored and uncensored data could be affected in disparate ways by various factors included in a model. For example, a given regressor could exert a positive (negative) influence upon a respondent's stated WTP, but a negative (positive) influence upon the likelihood that a respondent reports a zero bid for the good in question (Lusk and Shogren 2007). In order to use the double-hurdle technique, three separate models must be used: tobit, binomial probit, and truncated tobit. The log-likelihood statistic is captured from each, then used to calculate the following likelihood ratio statistic:

(2) $LR = -2[lnLF_{Tobit} - lnLF_{Binomial\ Probit} - lnLF_{Truncated\ Regression}]$.

This test statistic is then compared to a critical value from the chi-squared distribution, with the degrees of freedom equal to the number of independent variables in the model. If the null hypothesis that the tobit model is the correct specification is rejected, then the double-hurdle model should be used. The interested reader is invited to consult Lusk and Shogren (2007) for additional details on use of the double-hurdle model.

Results

Table 1 shows the mean WTP by elicitation method for each of the fictional steak brands along with Canada AAA beef. Stated WTP elicited via the "conventional" survey is highest for each fictional brand, WTP from the survey containing a cheap talk script is the second-highest, and bids from the experimental auctions are the lowest. In general, WTP was highest for the Original Angus steak, followed by Nature's Diamond, Tender Grill and Prairie Prime. Generic Canada AAA beef generally had the lowest stated WTP. Table 1 also provides results of t-tests for statistical equivalence of average WTP by elicitation method.

Table 1. Mean Consumer Willingness-to-pay in each treatment and t-test results

Steak	Experimental Auction ($/steak)	Cheap Talk Survey ($/steak)	Conventional Survey ($/steak)
Canada AAA	1.116	1.425[a]	1.472[a,d]
Prairie Prime	1.205	1.406[a]	1.481[a,d]
Tender Grill	1.317	1.431[c]	1.567[a,d]
Nature's Diamond	1.312	1.576[a]	1.767[a,e]
Original Angus	1.308	1.641[a]	1.810[a,e]

[a] indicates statistically different from the experimental auction at $\alpha = 0.05$.
[b] indicates statistically different from the experimental auction at $\alpha = 0.10$
[c] indicates not statistically different from the experimental auction
[d] indicates not statistically different from the cheap talk survey
[e] indicates statistically different from the cheap talk survey at $\alpha = 0.10$

The effectiveness of the cheap talk script in mitigating hypothetical bias can be quantified

approximately by calculating the difference in stated WTP for the five steaks across the two survey treatments. It should be noted that although average WTP for the cheap talk survey was lower than for the conventional survey for every beef brand, the WTP were only *statistically* different for two brands; this implies that caution should be exercised in interpreting these results. In percentage terms, the cheap talk script appears to have lowered stated WTP by between 2.8% (for Canada AAA beef) and 12% (for Original Angus). In general, the reduction in bids was smallest for the steaks with the fewest premium quality attributes (Canada AAA and Prairie Prime) and highest for those with several (Original Angus and Nature's Diamond). If WTP estimates elicited by the BDM auction are regarded as incentive compatible, then the difference between auction and cheap talk survey results approximates the amount of hypothetical bias remaining despite the use of a cheap talk script, of course recognizing that there are other potential reasons for the differences in stated WTP. Results from the auctions are lower by amounts ranging from 8.3% (for Tender Grill) to 27.7% (for Canada AAA beef).

Willingness-to-pay estimates elicited via BDM auctions were thus 20% lower on average than those elicited by a mail survey incorporating a cheap talk script. This is a potentially important finding for those engaged in market research—given the ease and relatively low cost of reaching a wide range of respondents via mail survey, the benefits of using this method are clear. However, researchers should interpret results of such surveys with caution, noting the results here imply a significant amount of hypothetical bias exists in this contingent valuation method, even when a cheap talk script is used.

Treatment costs invariably influence selection of elicitation method. For this research, per-survey costs were calculated to be $3.72, including stationary, printing, postage, student assistance and the $1 monetary incentive provided. Costs for the BDM auction amounted to $16.39 per response, including steaks, stationary (for steak information sheets), student assistance, and miscellaneous related costs. Though this cost discrepancy seems considerable, it must be remembered that the response rate for auctions is effectively one hundred percent—almost no costs are incurred for people who choose not to participate. By contrast, survey costs (except business reply postage, if it is used) are incurred for each recipient, regardless of whether they complete and return the survey.

Given this research's survey response rate of 28%, approximately 3.57 surveys were required to generate one usable response, yielding a cost per usable response of $13.29 (3.57 × $3.72). It should be noted that this response rate is unusually high for a "cold" mail survey (i.e. one being administered by an institution with whom the recipient has no prior relationship and pertaining to an issue with which the recipient has no expected prior specialized knowledge); a lower response rate would result in a higher cost per usable response—for example, a 20% response rate would result in a cost of $18.60 per response, even higher than the BDM auction cost of $16.39. Given the apparent bias associated with survey results compared to the incentive compatible BDM auctions, the benefits of employing a "cheap" mail survey instead of experimental auctions to measure WTP are somewhat questionable. Having said that, the costs of conducting market research across a wider geographic region will increase dramatically if experimental auctions rather than mail surveys are chosen.

Results of the tobit and double-hurdle econometric models of WTP for each elicitation method and fictional brand are shown in Table2. Fifteen models were estimated (four brands plus Canada AAA for each of three elicitation methods). The tobit model was rejected in favor of the double-hurdle model in most (but not all) cases. *Like name* exerted a positive and statistically significant influence upon WTP in nearly every model, demonstrating the importance of careful development of brand name and logo, as well as product information, when measuring WTP for a new product.

Though not statistically significant in all cases, most of the demographic variable coefficients were of the expected sign. There was no strong expectation on the sign for gender (female = 0; male = 1); Feuz et al. (2004) found males willing to pay more than females for steaks with various quality attributes whereas

Lusk, Feldkamp and Schroeder (2004) discovered a higher WTP for females. This variable was only statistically significant in three models; in each case, males were found to have the higher WTP. An interesting observation from the cheap talk survey was that for females, the brand with the highest mean willingness-to-pay was Nature's Diamond, whereas for males it was Original Angus.

Respondent age was statistically significant in four models and exerted a negative influence upon WTP in each of those cases, indicating that older respondents were less receptive to the attributes associated with the brand name steaks than younger ones. This is in line with the findings of Feuz et al (2004) and Lusk, Feldkamp and Schroeder (2004) who also discovered a negative relationship between respondent age and WTP. Several other researchers, including Lusk et al (2001), found this relationship to be indeterminate.

As predicted by economic theory, respondent income exerted a pervasively positive impact upon WTP, although the relationship was only statistically significant in five of the models. This reaffirms the findings of Lusk and Fox (2002). Results for education, the final demographic variable, were mixed: a positive and statistically significant result was discovered for two of the models; in two others the coefficient was negative and significant. Perhaps most interestingly, the coefficient was positive and significant in the experimental auction and conventional survey models for the Nature's Diamond brand, but negative and significant in the cheap talk model. Other research has discovered similarly mixed results: Lusk, Feldkamp and Schroeder (2004) found the relationship between education and WTP to be inverse, whereas Lusk and Fox (2002) found it to be positive. Similar to the findings of the present research, Lusk et al (2001) discovered the variable to have opposite signs in different treatments.

Econometric results for *beef eaten* are among the most intriguing—the relationship between WTP and this variable is significant and positive in three cases and negative in two others. Each of the positive findings is for WTP measured using an experimental auction, whereas both the negative findings come from data gathered using the conventional survey treatment. Given the BDM auction has been demonstrated to be incentive compatible and that auction participants were actual grocery shoppers approaching the meat counter in a supermarket, it is probably logical to characterize the experimental auction results as more reliable. This may indicate that purchasing decisions made in an experimental auction vs. hypothetical environment have consequences that go beyond mere differences in WTP estimates; it may be the case that results from mailed surveys are highly questionable in terms of reliability for this type of good. Similarly, each of the three cases where a negative and significant relationship was discovered between WTP and *confidence* used experimental auction data in estimation, while the two instances where a positive relationship were found used survey data. Again, it may be appropriate to characterize the BDM auction results as superior to those from the mailed surveys.

The fact that coefficients for the same variables within a common model can possess statistically significant coefficients opposite in sign depending upon WTP elicitation method is a potentially important finding for researchers. Auction participants were active shoppers approaching the meat counter in a grocery store whereas survey recipients are most likely in their home completing the survey; thus the importance of market research using "active" vs. "passive" (or hypothetical) shoppers may be important. Understanding of factors affecting consumers' purchasing decisions should be of extreme importance to firms in a marketplace. Decisions on market segmentation and advertising expenditures can depend critically on this understanding. Results reported here suggest careful selection of marketing research tools is in order—quantitative analysis based upon biased data has the potential to lead to suboptimal decision making by firms.

Conclusions

The three objectives of the paper were to determine the effect of elicitation method upon WTP for brand name steaks in Canada, to quantify the effects of cheap talk in reducing hypothetical bias, and to model the factors affecting WTP for

four fictional brands representing various steak attributes. Data from experimental auctions as well as cheap talk and "conventional" treatments of a mail survey were used in a hedonic model of consumer WTP.

Results of the research suggest that for each of the four fictional brands (as well as the generic Canada AAA steak), average stated WTP from the "conventional" mail survey treatment was the highest, followed by that from the cheap talk treatment. Average WTP from the incentive compatible BDM auctions was the lowest. Although a cheap talk script does result in lower stated WTP than the "conventional" treatment in a mail survey, stated WTP is still much higher than that elicited using incentive compatible BDM experimental auctions. This indicates that significant bias remains, even when a cheap talk script is used. It was also discovered that there exists a relatively small per-response cost difference between the two methods if non-responses by survey recipients are taken into account.

Findings from tobit and double-hurdle econometric modeling suggest consumer preference for brand name and logo results in higher WTP for steaks bearing brands that represent various desirable steak attributes. Demographic variables were found in several cases to exert statistically significant effects upon WTP that were mixed in some cases, but largely in agreement with both economic theory and the findings of previous researchers.

It was discovered that data gathered using different WTP elicitation methods can generate conflicting results within a common model. Respondents' self-assessed confidence in selecting beef was found to exert a negative effect upon WTP using experimental auction data, but a positive effect (though only for Canada AAA steak) using survey data. Similarly, the frequency with which respondents eat beef exerted a positive effect upon WTP using data from experimental auctions but negative effects when survey data were used. This is potentially important to not only experimenters, but also to firms hoping to use quantitative analysis of market research data for strategic purposes.

A number of opportunities exist to extend this work. Similar comparisons of WTP data from different elicitation methods should be used to determine whether the findings of this research are robust across experimental methods. This research used experiments in the field to measure WTP from auctions and compare it to results from a mail survey; perhaps experiments in a laboratory setting could be used to reaffirm these findings. Alternative products could also be used—it would be worthwhile to explore the extent to which various factors affect WTP for other types of goods, and whether there are similar effects of treatment upon stated WTP levels for alternate products.

References

Amemiya, T. 1973. "Regression analysis when the dependent variable is truncated normal." *Econometrica* 41(6): 997-1016.

Becker, G.M., M.H. DeGroot and J. Marschak. 1964. "Measuring utility by a single-response sequential method." *Behavioural Science* 9(3): 226-32.

Brown, T.C., P.A. Champ, R.C. Bishop and D.W. McCollum. 1996. "Which response format reveals the truth about donations to a public good?" *Land Economics* 72(1): 152-66.

Court, A.T. 1939. "Hedonic price indexes with automotive examples." The Dynamics of Auto Demand. General Motors Corporation, pp. 98-119.

Cragg, J.G. 1971. "Some statistical models for limited dependent variables with application to the demand for durable goods." *Econometrica* 39(5): 829-44.

Cummings, R.G. and L.O. Taylor. 1999. "Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method." *American Economic Review* 89(3):649-65.

Feldkamp, T.J., T.C. Schroeder and J.L. Lusk. 2005. "Determining consumer valuation of differentiated beef steak quality attributes." *Journal of Muscle Foods* 16(1): 1-15.

Feuz, D.M., W.J. Umberger, C.R. Calkins and B. Sitz. 2004. "U.S consumers' willingness to pay for flavor and tenderness in steaks as determined with an experimental auction." *Journal of Agricultural and Resource Economics* 29(3): 501-16.

Froehlich, E.J., J.G. Carlberg and C.E. Ward. 2009. "Willingness-to-pay for fresh brand name beef." *Canadian Journal of Agricultural Economics* 57(1): 119-37.

Goodman, A.C. 1998. "Andrew court and the invention of hedonic price analysis." *Journal of Urban Economics* 44(2): 291-8.

Griliches, Z. 1958. "The demand for fertilizer: an econometric reinterpretation of a technical change." *Journal of Farm Economics* 40(3): 591-606.

Lancaster, K.J. 1966. "A new approach to consumer theory." *Journal of Political Economy* 74(2): 132-57.

Ladd, G.W. and M.B. Martin. 1976. "Prices and demands for input characteristics." *American Journal of Agricultural Economics* 58(1): 21-30.

List, J.A. 2001. "Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards." *American Economic Review* 91(5): 1498-1507.

List, J.A. 2003. "Using random *n*th price auctions to value non-market goods and services." *Journal of Regulatory Economics* 23(2):193-205.

Loomis, J., T. Brown, B. Lucero and G. Peterson. 1997. "Evaluating the validity of the dichotomous choice question format in contingent valuation." *Environmental and Resource Economics* 10(2): 109-23.

Loomis, J.B., A. Gonzalez-Caban, and R. Gregory. 1996. *A Contingent Valuation Study of the Value of Reducing Fire Hazards to Old-Growth Forests in the Pacific Northwest.* U.S. Dept. of Agriculture, Forest Service, Pacific Southwest Research Station, Albany, CA.

Lusk, J.L., J. Fox, T. Schroeder, J. Mintert and M. Koohmaraie. 2001. "In-store valuation of steak tenderness." *American Journal of Agricultural Economics* 83(3): 539-50.

Lusk, J.L. and J.A. Fox. 2002. "Consumer demand for mandatory labeling of beef from cattle administered grown hormones or genetically modified corn." *Journal of Agricultural and Applied Economics* 37(1): 27-38.

Lusk, J.L. and J.A. Fox. 2003. "Value elicitation in retail and laboratory environments." *Economics Letters* 79(1): 27-34.

Lusk, J.L. "Effectiveness of Cheap Talk On Consumer Willingness-To-Pay for Golden Rice." *American Journal of Agricultural Economics* 85(November 2003):840-56.

Lusk, J.L., T. Feldkamp and T.C. Schroeder. 2004. "Experimental auction procedure: Impact on the valuation of quality differentiated goods." *American Journal of Agricultural Economics* 86(2): 389-405.

Lusk, J.L. and J.R. Shogren. 2007. *Experimental Auctions: Methods and Applications in Economic and Marketing Research.* Cambridge: Cambridge University Press.

Neill, H.R., R.G. Cummings, P.T. Ganderton, G.W. Harrison and T. McGuckin. 1994. "Hypothetical surveys and real economic commitment." *Land Economics* 70(1): 145-54.

Nesheim, L. 2006. "Hedonic price functions." Centre for Microdata Methods and Practice working paper CWP18/06, The Institute for Fiscal Studies, Department of Economics, University College of London.

Ness, M.R. and H. Gerhardy. 1994. "Consumer preferences for quality and freshness attributes of eggs." *British Food Journal* 96(1): 26-34.

Tobin, J. 1958. "Estimation of relationships for limited dependent variables." *Econometrica* 26(1): 24-36.

Umberger, W.J. and D.M. Feuz. 2004. "The usefulness of experimental auctions in determining consumers' willingness-to-pay for quality differentiated products." *Review of Agricultural Economics* 26(2): 170-85.

Table 2. Tobit and double-hurdle econometric model results, all treatments

	Canada AAA			Prairie Prime			Tender Grill			Nature's Diamond			Original Angus		
	EA[a]	CT[b]	Conv.[a]	EA[a]	CT[a]	Conv.[a]	EA[b]	CT[a]	Conv.[b]	EA[a]	CT[a]	Conv.[a]	EA[a]	CT[a]	Conv.[a]
Intercept	0.22	1.01*	-1.52	-0.69	-0.54	0.10	0.85*	2.46	1.68**	-3.34*	2.67**	-3.13	-0.71	1.34	-0.19
	(0.74)	(0.58)	(1.54)	(1.05)	(1.29)	(1.74)	(0.45)	(1.62)	(0.53)	(1.75)	(1.07)	(2.34)	(0.70)	(1.74)	(1.58)
Beef Eaten	0.19**	0.10	-0.02	0.10	0.06	-0.24*	0.12**	0.03	-0.12*	0.06	-0.11	0.19	0.24**	-0.18	-0.11
	(0.08)	(0.07)	(0.10)	(0.11)	(0.10)	(0.14)	(0.06)	(0.11)	(0.07)	(0.16)	(0.10)	(0.12)	(0.09)	(0.13)	(0.12)
Like Name	n/a	n/a	n/a	0.34**	0.04	0.12	0.26**	0.03	0.42**	0.39**	0.34**	0.52**	0.24**	0.60**	0.60**
	n/a	n/a	n/a	(0.14)	(0.09)	(0.14)	(0.06)	(0.12)	(0.07)	(0.15)	(0.08)	(0.13)	(0.11)	(0.14)	(0.17)
Confidence	-0.10	0.10*	0.19*	-0.26**	-0.08	0.15	-0.11*	-0.11*	0.06	-0.27**	0.06	0.01	-0.10	0.04	0.04
	(0.07)	(0.06)	(0.10)	(0.11)	(0.08)	(0.11)	(0.06)	(0.09)	(0.06)	(0.14)	(0.08)	(0.12)	(0.08)	(0.10)	(0.10)
Gender	0.18	0.17	0.04	0.73**	0.03	0.22	0.21	-0.02	0.44**	0.13	-0.16	-0.09	0.58**	0.26	-0.03
	(0.19)	(0.18)	(0.30)	(0.26)	(0.26)	(0.34)	(0.15)	(0.30)	(0.17)	(0.36)	(0.24)	(0.39)	(0.21)	(0.30)	(0.31)
Age	-0.11	-0.22**	-0.02	-0.16	0.03	0.04	-0.07	-0.45**	-0.31**	0.04	-0.24	0.11	-0.18*	-0.12	-0.20
	(0.09)	(0.07)	(0.23)	(0.14)	(0.17)	(0.29)	(0.05)	(0.23)	(0.07)	(0.18)	(0.15)	(0.26)	(0.09)	(0.25)	(0.22)
Income	0.20**	0.26**	0.49	0.01	0.59**	-0.20	0.02	0.58**	-0.01	0.06	0.43**	0.20	0.14	0.10	0.31
	(0.09)	(0.08)	(0.23)	(0.15)	(0.20)	(0.23)	(0.06)	(0.21)	(0.07)	(0.23)	(0.15)	(0.23)	(0.11)	(0.20)	(0.21)
Education	0.01	-0.09	0.06	0.16	-0.16	0.20	-0.07	-0.32**	-0.01	0.76**	-0.24**	0.56**	0.09	0.27	0.06
	(0.08)	(0.07)	(0.19)	(0.17)	(0.13)	(0.25)	(0.05)	(0.14)	(0.06)	(0.36)	(0.11)	(0.27)	(0.08)	(0.17)	(0.15)

Notes: Standard errors are given in parentheses. double and single asterisks denote statistical significance at the 5% and 10% levels, respectively. a indicates the double-hurdle model was used. b indicates the tobit model was used. EA denotes model results using data from the BDM experimental auction; CT denotes model results using data from the mail survey that incorporated a cheap talk script; Conv. denotes model results using data from conventional mail survey (i.e. no cheap talk script).