HYPOTHETICAL (NON)BIAS IN CHOICE EXPERIMENTS: EVIDENCE FROM FRESHWATER PRAWNS

Darren Hudson, Karina Gallardo, and Terry Hanson
Associate Professor, Mississippi State University
Ph.D. Student, Oklahoma State University
Associate Professor, Mississippi State University

Abstract

A mail survey choice experiment and in-store controlled experiment were conducted concurrently in the same location to test for hypothetical bias in choice experiments using a new product—freshwater prawns. Findings suggest that hypothetical bias in the mail survey was not present for the new product, but some hypothetical bias was detected for substitute products, depending on the choice of econometric estimation method. However, the general conclusion is the choice experiments are an effective means of assessing potential market demand for new products with little evidence for hypothetical bias.

HYPOTHETICAL (NON)BIAS IN CHOICE EXPERIMENTS: EVIDENCE FROM FRESHWATER PRAWNS

Conjoint analysis has become a popular method of estimating willingness to pay (WTP) for products and services in a variety of settings (Adamowicz et al.; Beggs, Cardell, and Hausman; Jayne et al.; Lusk, Roosen, and Fox; Unterschultz et al.). Their popularity is understandable given their consistency with Lancaster's utility theory (Louviere, Hensher, and Swait), the ability to handle a number of attributes simultaneously in a controlled, orthogonal experimental design (Lusk and Schroeder), and the ability to generate a large number of observations on choice from a relatively small number of respondents, thus lowering cost (Louviere, Hensher, and Swait). Despite their popularity in applied analysis, a number of critical questions remain open as to the validity of conjoint analysis in predicting actual behavior.

In terms of valuation, a critical question relates to the potential for hypothetical bias in choice experiment/conjoint questions (CE). Hypothetical bias has been widely documented in contingent valuation literature (List and Gallet; List and Shogren; Fox et al.). Unlike contingent valuation, CE questions are typically posed in a manner that is more like true choice environments, leading to a maintained hypothesis that CEs are less likely prone to hypothetical bias. Recent research by Carlsson and Martinsson and Lusk and Schroeder lend some support to this hypothesis. More specifically, Lusk and Schroeder find that hypothetical choice experiments result in WTP values that are statistically higher than WTP values from non-hypothetical choice experiments, but also find that the marginal differences in WTP between product choices are not different between hypothetical and non-hypothetical experiments. Carlsson and Martinsson's analysis focused only on marginal differences in WTP between product choices and found similar results to Lusk and Schroeder.

While both studies provide insight into the external validity of choice experiments, they are confined to a laboratory setting. While laboratory experiments do offer a high degree of control over decision variables of interest (Hudson), their sterile nature makes generalization to a complex real world difficult. That is, laboratory experiments are conducted "out of context," which may lead respondents to focus all attention on the decision task. Conversely, real shoppers are attempting to make a myriad of choices, which increases taxation on mental capacity and may lead to a different set of decisions.

To more fully understand the external validity of CEs, it would be useful to compare results of CEs with actual purchasing behavior. However, naturally occurring purchase data are subject to lack of control over prices to mimic the CE data, resulting in confusion over whether observed differences are artifactual or behavioral. One would need to have control over prices in the natural environment to truly test predictability of CEs with naturally occurring purchase data.

This paper outlines an experiment which combines a choice experiment with a controlled experiment in a grocery store, with the specific objective to test whether the CE has predictive capacity for real-world purchase decisions. Rather than conducting in-store auctions or "taste tests," this study places the product in the store where prices are controlled, but the shopper is unaware of the experimental design so that the shopping experience is identical to the shopper's ordinary experience. This approach necessarily means sacrificing some control over the external environment, but improves the realism of the experiment.

The product used in the analysis was the freshwater prawn, which is a product similar in appearance to marine shrimp. Prawns make an interesting subject for analysis because it is a relatively new product in the U.S. market, but is similar to existing products (marine shrimp and lobster), which allows for an interesting view of the predictive power of the CE in a setting where a new product is introduced into a market with existing substitutes. From the perspective of point estimates, we generally find that the CE performs well in predicting actual market outcomes in terms of consumer WTP, but predictive power is sensitive to the choice of estimation method.

Methods

The analysis centers on two related experiments conducted concurrently during January-March 2004 in the same location (Germantown, TN)[1]. First, we describe the in-store experiment. Second, we describe the CE survey. Finally, the procedures used in the analysis are presented.

Grocery Store Experiment

First, the grocery store experiment was conducted in a major grocery chain store in Germantown, which was one of the major grocery stores in that suburban community. The grocery store agreed to stock the freshwater prawns in their fresh seafood counter, allowing the researchers to set the price for the prawns on a weekly basis and collect data

[1] Germantown, TN is a suburb of Memphis, TN. Germantown was chosen for two important reasons. First, it represents a more affluent, suburban community that is the most likely target market for the freshwater prawns. Second, and most importantly, it was the location where the grocery chain agreed to test market the product.

on sales of the freshwater prawns as well as the competing products of marine shrimp and lobster.[2]

Regional grocery chain stores were contacted to determine a reasonable range of prices for marine shrimp. Given that marine shrimp and freshwater prawns are close in composition, look, taste, and texture, it was assumed that freshwater prawns would be priced similarly to marine shrimp. Data collected over a three week period at 3 regional grocery stores showed that marine shrimp prices typically ranged between $7 to $13 per pound, depending on variety and size. The freshwater prawns (FP) were divided into two size categories—large (23-45 count[3]) and jumbo (14-22 count)—which is consistent with most measures of marine shrimp.[4]

Five price levels were established for the FP, and randomly assigned to different weeks as shown in Table 1. The price range was established to encompass the normal range of price observed in the stores for shrimp. The experiment was scheduled for a 10 week period with two different products (heads on and off, described below), resulting in the need for 5 weeks of different prices. Thus, prices were varied from $5.99-$13.99/lb for large prawns in $2/lb increments.[5] Prices for large and jumbo were paired (so that the lowest price for large was paired with the lowest price for jumbo) for realism. The five

[2] The prawns were either donated or purchased from members of the U.S. Freshwater Prawn Growers Association and an agricultural experiment station. The grocery store kept all proceeds from the sale of the prawns. Thus, the grocery store had the incentive to treat the product as it would any other revenue generating product offered in their store.

[3] The "count" refers to the number of prawns per pound. So, a 14-22 count means there are between 14 and 22 prawns per pound. All references to products in this paper are to headless shrimp/prawns/lobster with shells on.

[4] In reality, counts of shrimp in different size categories vary by species of shrimp and the store that sells them. These count sizes are consistent with the grocery store used in the experiment and is typical for Gulf of Mexico shrimp.

[5] The jumbo size was also offered in the store, but the results presented here focus on the large size, which was the size used in the choice experiment.

pairs were then randomly assigned by week to the experiment, and the resulting order of price is shown in Table 1.

The FP products were offered in the grocery store over a ten-week period. Two product forms were offered—heads on and head off (tails only)—with each form offered in alternating weeks beginning with the head-on form.[6] This study focuses on the results relating to the heads off product form (that is, there are five weeks of data on sales of the heads off product).[7] Data were collected by the grocery store on sales of the FP product during the week, as well as sales and prices of both fresh and frozen shrimp and lobster. These data were used to calculate a weighted average price and standard deviation of price for each product sold, and a sample of 1,000 observations on average price were simulated assuming a normal distribution. This process generated a distribution of prices for each product from which comparisons could be made to the choice model results.

CE Survey

A mail survey was conducted in the Germantown community during the same period as the in-store experiment was being conducted. A random sample of 2,000 names from the Germantown community (the same zip code as the store) was purchased from a commercial marketing firm. A Dillman three-wave design was used—survey, then reminder card, then survey—to mitigate no response bias (Dillman; Pennings, Irwin, and Good; Hudson et al.). The survey collected basic data on consumption patterns and attitudes toward seafood, demographic variables, as well as the choice experiment.

[6] Product forms were alternated to give time for shoppers to "forget" the prices during the previous period to reduce attempts to predict the pricing pattern.
[7] The study focuses on the heads off form results because these are the data that are consistent with the CE experiment.

The CE was constructed in a manner similar to Lusk and Schroeder, whereby respondents faced a series of choices on product type—in this case, prawns, shrimp, and lobster—where only the price of the product was allowed to vary.[8] An example of a CE scenario is shown in Figure 1. The same set of prices was used for the CE as for the grocery store experiment to insure comparability.[9] Thus, there were five price levels for each product. Because of the large number of potential choice sets ($3^5 = 243$), a fractional factorial design was used. The fractional factorial represents a subset of the full factorial that minimizes the number of choice sets as well as the correlation among the attributes (see Kuhfled, Tobias, and Garrett or Lusk and Schroeder). The result was 25 choice sets, but with 25 choice sets, respondent fatigue may still be a problem (Bradley and Daly). Thus, the 25 choice sets were randomly blocked into two different sets—one with 12 and one with 13 choice sets. These two different versions were randomly assigned to individuals, resulting in 1,000 people initially receiving version 1 and 1,000 people initially receiving version 2.

The combined data from the grocery store and the CE are unique in a number of respects. First, the researchers had direct control of the product price in both the natural environment (the grocery store) as well as in the CE. As Lusk and Schroeder and Shogren et al. note, this is a preferred method of testing external validity, but is often difficult to get grocery stores to agree to participate due to the proprietary nature of the data. In this case, we had full cooperation of the grocery store, leading to a direct test of

[8] Each category was 1 pound of product, with shrimp and prawns being the same count size (23-45 count). Thus, the stated price in CE experiment was on a per pound basis. This was designed to match the count sizes in the in-store experiment for the "large" category.

[9] Here, the FP prices were the same as used in the store. The price ranges for shrimp and lobster were consistent with the store, but not all prices used in the survey were observed in the store over the test period. Shrimp prices were comparable, but lobster prices in the store were concentrated at the high end of the range used in the survey. This concentration for lobster prices ultimately had some impact on the results for lobsters as will be discussed in the results.

external validity. One can certainly argue that there are other grocery stores in the area, thus offering consumers a choice not captured in the CE. However, no other grocery stores in the area were offering freshwater prawns. A second important feature of these data is that they were collected at the same time in the same location. This prevents confounding potential seasonal or location effects.

At the same time, there is some difficulty in using this procedure as well. Unless demographic data of the grocery store shoppers is collected, it is difficult to know whether differences arising between CE and grocery store results arise from differences in sample or hypothetical bias.[10] As it is impractical to collect demographic data from each individual (as well as it may provide them information about the fact they are being studied, thus influencing their behavior), it is difficult to work around this issue. However, given that we do have the demographic data from the sample, which is confined to a specific community (more specifically, a zip code within the community where the store was located) with given demographic characteristics, we will assume that for the purpose of this analysis that patrons to this grocery store represent a random sample from the community at large so that consistency between the mail sample and community demographics proxies for consistency between in store and mail samples.

Data Analysis

Responses to CE questions were analyzed according to random utility theory, which holds that utility is given by:

(1) $$U_{ij} = V_{ij} + \varepsilon_{ij} \,,$$

[10] It may be possible for many products to collect demographic data from a "shopper's card" or some other device. However, in this case, the seafood department only inserts a "seafood" UPC code on purchases from the fresh counter, so it is impossible to trace what "seafood" products were being purchased and match them directly to demographic data. Data for this analysis was collected directly by the seafood department, and no linking to shopper information was possible.

where U_{ij} is utility for the i[th] consumer choosing the j[th] product, V_{ij} is the deterministic portion of utility for i and j, and ε_{ij} is the random component of utility. If we assume that consumers wish to maximize the unconditional indirect utility: $V_i = \max\lfloor V_1, V_2, ..., V_j \rfloor$, consumers will only choose product j if $U_{ij} \geq U_{ik}$. The probability that consumer i chooses alternative j from a set of k alternatives is given by:

$$\text{(2)} \qquad \Pr(j\, is\, chosen) = \Pr\{V_{ij} + \varepsilon_{ij} \geq V_{ik} + \varepsilon_{ik}; \forall\, k \in C_i\}$$

where C_i is the set of all consumer choice alternatives {C = prawns, shrimp, lobster, none}.

The most common method of estimating parameters for this model is the multinomial logit/conditional logit (CL) approach, which assumes that the error terms on utility are iid with a Type I extreme value distribution. Given these assumptions, the probability that consumer i chooses alternative j is modeled as (or the conditional indirect utility function):

$$\text{(3)} \qquad P(j\, is\, chosen) = \frac{e^{V_{ij}}}{\sum_{k \in C} e^{V_{ik}}}.$$

However, the CL approach suffers from the assumption of independence of irrelevant alternative (IIA), or that model errors are independently distributed across model alternatives. Several alternative functional forms exist which relax the IIA assumption, although in different ways. The heteroskedastic extreme value (HEV) model assumes that errors are independently, but not identically, distributed across the alternatives (prawns, shrimp, and lobster).

From a slightly different perspective, the multinomial probit (MNP) relaxes the IIA assumption by assuming that the errors across alternatives are normally distributed.

To operationalize the MNP model, we assume that all off-diagonal covariances are zero and hold the variance of the alternatives equal to one. This produces a model very similar to the HEV model except that errors are distributed normally rather than extreme value. Yet another method of relaxing the IIA assumption is through the random parameters logit (RPL) model (Revalt and Train). Here, taste parameters are assumed to be random within the population with a given distribution (in this case, normal). We allow the alternative specific constants to vary randomly within the population and hold the price parameters fixed.[11]

Finally, the inclusion of the "none" category as shown in Figure 1 allows respondents to "opt out" or choose not to purchase. This may result in a sample selection problem in that we have observations on respondents to the survey who choose not to purchase, but do not have the same data from the grocery store. To account for this possibility, we formulated a nested logit model. One branch represents the "no-buy" decision, or the "none" in the choice set. The second branch represents the "buy" decision. Under this, the limbs are a choice between prawns, shrimp, and lobster. Thus, the model generates results predicting the first stage choice of buy/no-buy as well as the second stage choice among alternative products. To operationalize this model, we held lobster as the base case, resulting is WTP estimates for prawns and shrimp.

Estimation of point estimates for product WTP from each model is a straightforward calculation:

$$(4) \qquad WTP_j = \left| \frac{\alpha_j}{\beta_j} \right|,$$

[11] In the RPL model, if the standard deviations in the population are equal to zero, the RPL is equivalent to the CL model. Holding the price parameters constant insures that the WTP distribution derived from equation 4 is normally distributed.

where α_j is the alternative (product) specific constant term and β_j is the price response coefficient for product j. WTP_j represents a point estimate of WTP for product j, but does not provide an estimate of variance.[12] The procedure proposed by Krinsky and Robb is used to bootstrap confidence intervals around WTP_j to facilitate statistical tests. The variance-covariance of α_j and β_j was used to generate a bivariate normal density on WTP_j using 1,000 simulated observations. A 95% confidence interval on $\overline{WTP_j}$ was constructed from these simulated observations. These bootstrapped distributions on WTP were then used to compare against the distributions of grocery store prices derived from the in-store experiment.

To examine differences between WTP distributions, we follow the combinatorial procedure employed by Lusk and Schroeder which follows the work of Poe, Giraud, and Loomis. The combinatorial approach takes the difference between the ith element of one distribution and every element of the second distribution. For example, it takes the difference between the 1st element of the grocery store distribution of prawn prices and each of the 1,000 simulated WTP values from the CL model. In this manner, the procedure constructs every possible difference between the two distributions (1,000*1,000 = 1 million differences). Within this distribution, the percentage of observations greater than zero is the unbiased, non-parametric p-value that the mean of the first WTP distribution is statistically greater than the mean of the second WTP distribution (Poe, Giraud, and Loomis).

[12] The WTP calculation presented here assumes no uncertainty about the product in question. Adamowicz et al. discuss alternative calculations when consumers may have some uncertainty about the product under question. Given that this product is a straightforward food product, little consumer uncertainty was expected. Further, respondents in the mail survey were provided a brochure with nutritional facts and pictures of prawns to reduce uncertainty about the product (the same information was made available in the store).

Results

Sample Characteristics

Of the 2,000 original surveys mailed, 91 were returned with incorrect addresses, leaving an effective sample of 1,909. Of these, 550 were returned (response rate = 28.8%), but 523 were usable (usable response rate = 27%). While somewhat lower than desired, the response rate was still within the acceptable norm for mail surveys (Dillman). The demographic characteristics were compared to the U.S. Census for the Germantown community (Table 2). As can be seen, the mail sample tracks census data generally well, especially for income and ethnicity, with higher education somewhat overrepresented.[13] However, males appear to be overrepresented in the sample compared with the general population. Given the disparity in male representation in the sample versus the general population, data from the survey were weighted by the proportion of males in the sample to proportion of males in the population (ratio = 1.2671). The weighted data were used in the econometric analysis.

Grocery Store Results

Overall, 36 pounds of large size prawns were sold in the fresh seafood counter over the five week period (compared with 122.5 pounds of marine shrimp in the fresh counter and 286.5 pounds in the frozen section and 132 pounds of lobster). The weighted average prices of prawns, shrimp, and lobster with their associated 95% confidence intervals are shown in Table 3. As can be seen, prawn weighted average prices are higher, but not

[13] The table shows large disparities in age, but age is difficult to judge because the sample was restricted to individuals with mailing addresses, which necessarily precludes children whose numbers are reflected in the census. Further, education can be somewhat misleading as the mail sample uses categories to approximate years of education.

statistically different, than marine shrimp average prices, reinforcing the *a priori* expectation that prawns are viewed as close substitutes for other shellfish products.

CE Results

Table 4 shows the results of the CL and RPL models. All alternative specific constants (ASCs) are significantly different from zero, indicating that all products were preferred to "none." Additionally, all price coefficients are negative and statistically different than zero, indicating that increases in price lead to a decreased probability of choice. As noted in Table 4, the standard deviations of the ASC distributions for the RPL model are all not significantly different from zero, suggesting no significant heterogeneity in taste preferences. As such, the parameter estimates from the RPL closely resemble those from the CL, with some refinement in the standard errors.

Table 5 shows the results of the HEV and MNP models. Again, all relevant parameter estimates are statistically significant with expected signs. Because the scale parameter is inversely related to the error variance of the choice, results suggest more noise in the prawn alternative than shrimp or lobster, which might be expected for a new product. However, all alternatives have more noise than the "none" alternative. The MNP model yields similar results to the HEV model, holding all standard deviations equal to 1. This assumption necessarily forces all of the standard deviations of the error to be the same, which is contraindicated by the HEV model, but the differences in standard deviation do not appear large (a fact further supported by the RPL model).

Finally, Table 6 shows the results of the NL model. Again, all parameter coefficients were statistically significant and of the expected sign. Here, only prawns and shrimp are estimated in the alternative choice equation for identification purposes. In addition, increases in price levels significantly reduced the probability of a "buy" as indicated by the branch choice equation.

Based on these results, mean WTP and confidence intervals were bootstrapped using the Krinsky-Robb procedure (Table 7). Results show some differences in WTP estimates, which is not surprising given the different assumptions made on the error structure. Generally, results of the CL and RPL are consistent with one another, as are the results of the HEV and MNP models. Overall, the CL and RPL models produced WTP estimates that were lower than the HEV and MNP models. The NL model appeared to fall between the two groupings.

Comparisons

Table 8 shows the comparison between the grocery store results and the WTP estimates from the choice experiment using the Poe, Giraud, and Loomis combinatorial approach. For prawns, there appears to be no statistical evidence of hypothetical bias between the survey results and the grocery store results for any model form. However, for shrimp, the CL and RPL models appears to produce estimates that were statistically lower than the grocery store results, but the HEV and MNP models did not. This is consistent with the prawn results in that the HEV and MNP models did more precisely predict the grocery store results as indicated by the much higher p-values. Interestingly, the nested logit model showed no signs of hypothetical bias for either product, which may indicate that by accounting for potential sample selection, overall estimates were improved.

For lobster, all models showed the existence of hypothetical bias. This result is likely an artifact of the grocery store data. That is, the CE encompassed a wide range of potential prices for lobster, but prices in the store were clustered at the high end of that distribution, generating a relatively tight simulated distribution for store prices. The other products experienced a wider range of prices in the store—prawns by design, and shrimp by virtue of the natural change in prices over the time period of the experiment.

Discussion

The finding of no hypothetical bias in this paper is certainly heartening to those that employ this method and reinforces the widespread use of this approach in the private sector to make predictions about potential new product demand and pricing. These results appear to run counter to those of Lusk and Schroeder in that they found evidence of significant hypothetical bias in the levels of WTP. However, their study was confined to a laboratory setting. While the lab certainly offers a degree of control not found in a natural environment, the sterile nature of the lab places the shopper "out of context," which may lead to a different conclusion as compared to the situation where the shopper is making actual choices in their natural context. The present study, while suffering some loss of control, extends the debate about potential bias into the natural shopping environment and reaches a somewhat different conclusion.

A shortcoming of this analysis that may lead to some differences in results is the relatively small sample size in the grocery store. The study was only conducted over a ten week period (with five weeks of observations on the product in question). While this represents a substantial period of time for such a controlled experiment in a store and prawn purchases over this period were comparable to other competing seafood products in the fresh counter, it remains an open question whether longer periods of time would have resulted in different weighted average prices and premiums. There is likely some seasonality in seafood purchases. However, the mail survey was conducted at the same time as the grocery store experiment to control for any "seasonal bias" that might be in the consumer's mind.

Conclusions

This paper has examined the issue of hypothetical bias in choice experiments using choice experiment data and a controlled experiment in a grocery store to determine the predictive capacity of the CE for real world outcomes. While previous studies have examined the issue of hypothetical bias, this paper adds the unique feature of having a controlled, real-world experiment conducted concurrently with the CE at the same location.

We find no statistical evidence that hypothetical bias exists for the product in question in the level of WTP predicted by the CE relative to the grocery store weighted average prices. However, it appears that the conclusions reached on shrimp are sensitive to the choice of estimation method. This observed variance in results across functional forms is no surprise, but reinforces the importance of functional form choice to insure the most applicable results.

This paper highlights the need for additional work in this area. To improve upon this approach, it would be desirable to obtain demographic data of the shoppers in the store who purchase the products. Due to the proprietary nature of such data, it may difficult to obtain, but would certainly allow for a richer analysis of preferences in comparison to hypothetical surveys. Further, it would enhance the robustness of the experiment if the in-store experiment allowed for control over all relevant prices. Here, we controlled prawn prices directly, but were forced to observe prices for other products. While this posed no serious problem for the shrimp, it did appear to affect the results for lobster, leading to decreased confidence in being able to analyze cross-price effects. Nevertheless, this study does provide evidence for the predictive capacity of CEs as a market research device.

References

Adamowicz, W., R. Boxall, M. Williams, and J. Louviere. "Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation." *American Journal of Agricultural Economics*, 80(1998): 64-75.

Allenby, G. and J. Ginter. "The Effects of In-Store Displays and Feature Advertising on Consideration Sets." *International Journal of Research in Marketing*, 12(1995): 67-80.

Beggs, S., S. Cardell, and J. Hausman. "Assessing the Potential Demand for Electric Cars." *Journal of Econometrics*, 16(1981): 1-19.

Bhat, C. "A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice." *Transportation Research B*, 29(1995):471-483.

Bradley, M. and A. Daly. "Use of the Logit Scaling Approach to Test for Rank-Order and Fatigue Effects in Stated Preference Data." *Transportation*, 21(1994): 167-184.

Carlsson, F. and P. Martinsson. "Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments." *Journal of Environmental Economics and Management*, 41(2001): 179-192.

Dillman, D. *Mail and Telephone Surveys: The Total Design Method*. John Wiley, New York, NY, 1978.

Fox, J., J. Shogren, D. Hayes, and J. Kliebenstein. "CVM-X: Calibrating Contingent Values with Experimental Auction Markets." *American Journal of Agricultural Economics*, 80(1998): 455-465.

Hudson, D. "Problem Solving and Hypothesis Testing Using Economic Experiments." *Journal of Agricultural and Applied Economics*, 35(2003): 337-347.

Hudson, D. and J. Lusk. "Risk and Transactions Cost in Contracting: Results of a Choice-Based Experiment of Producer Preferences for Contract Attributes." *Journal of Agricultural and Food Industrial Organization*, 2(2004): Article 2, http://www.bepress.com/jafio/vol2/iss1/art2.

Hudson, D., D. Hite, L. Seah, and T. Haab. "Telephone Presurveys, Sample Selection, and Non-Response Bias to Mail and Internet Surveys in Economic Research." *Applied Economics Letters*, 11(2004): 237-240.

Jayne, T., L. Rubey, F. Lupi, D. Tschirley, and M. Weber. "Estimating Consumer Response to Food Market Reform Using Stated Preference Data: Evidence for Eastern and Southern Africa." *American Journal of Agricultural Economics*, 78(1996): 820-824.

Kuhfeld, W., R. Tobias, and M. Garrett. "Efficient Experimental Design with Marketing Research Applications." *Journal of Marketing Research*, 31(1994): 545-557.

Lancaster, K. “A New Approach to Consumer Theory.” *Journal of Political Economy*, 74(1966): 132-157.

List, J. and C. Gallet. “What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values.” *Environmental and Resource Economics*, 20(2001): 241-254.

List, J. and J. Shogren. “Calibration of the Differences Between Actual and Hypothetical Valuations in a Field Experiment.” *Journal of Economic Behavior and Organization*, 37(1998): 193-205.

Louviere, J., D. Hensher, and J. Swait. *Stated Choice Methods: Analysis and Application*. Cambridge University Press, Cambridge, United Kingdom, 2000.

Lusk, J. and T. Schroeder. “Are Choice Experiments Incentive Compatible: A Test With Quality Differentiated Beef Steaks.” *American Journal of Agricultural Economics*, 86(2004): 467-482.

Lusk, J., J. Roosen, and J. Fox. “Demand for Beef from Cattle Administered Growth Hormones or Fed Genetically Modified Corn: A Comparison of Consumers in France, Germany, and the United Kingdom.” *American Journal of Agricultural Economics*, 85(2003): 16-29.

Pennings, J., S. Irwin, and D. Good. "Surveying Farmers: A Case Study." *Review of Agricultural Economics*, 24(2002): 266-277.

Revalt, D. and K. Train. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *Review of Economics and Statistics*, 80(1998): 647-657.

Unterschultz, J., K. Quarainie, M. Veeman, and R. Kim. "South Korean Hotel Meat Buyers' Perceptions of Australian, Canadian, and U.S. Beef." *Canadian Journal of Agricultural Economics*, 46(1998): 53-68.

Table 1. Randomly Assigned Prices for Freshwater Prawns, Heads Off, Germantown, TN, 2004.

Size			Weekly Price in $/pound				
			Week 1[a]	Week 3	Week 5	Week 7	Week 9
Large	10-20 grams each	23-45 count	9.99	13.99	11.99	5.99	7.99

[a] Products were offered every other week. A different product form (heads on) was offered in the other weeks.

Table 2. Comparison of Response and Census Demographic Characteristics, Germantown, TN, 2004.

Demographics	**Survey n=523 %**	**US Census N=37,348 %**
Age of Respondent		
Under 35	4.76	40.40
Between 35 and 50	22.78	16.90
Between 50 and 65	49.48	33.50
Older than 65	22.98	9.30
Gender		
Percent Male	61.71	48.70
Household Income		
Less than $25,000	3.83	6.10
$25,000 - $50,000	10.81	12.80
$50,000 - $75,000	15.54	18.10
$75,000 - $100,000	19.14	16.30
$100,000 or more	50.68	46.80
Education		
Less than high school	0.40	2.00
High school	4.18	11.00
Some college	12.75	22.40
Completed college	42.63	42.30
Beyond B.S. degree	40.04	22.40
Ethnicity		
Caucasian	95.09	92.90
African American	1.02	2.30
Native American	0.61	0.20
Asian	2.04	3.50
Hispanic	0.61	1.10
Other	0.61	--

Table 3. Comparison of Weighted Average Prices of Shellfish Products Sold in the Fresh Counter, Germantown, TN, 2004.

Size	**Freshwater Prawns**	**Marine Shrimp**[a]	**Lobster**[b]
	$/lb.	$/lb.	$/lb.
Large	9.49	8.72	12.70
	[7.73-11.25]	[8.37-9.07]	[11.46-13.93]

[a] Marine shrimp weighted average prices are based only on sales in the fresh counter.
[b] Lobster was only sold in one size, but the price is in $/lb.
[c] Numbers in brackets are 95% confidence intervals.

Table 4. Conditional Logit and Random Parameters Logit Model Results for CE Responses, Germantown, TN, 2004.

Parameter	**Conditional Logit**		**Random Parameters Logit**	
	Estimate	**Standard Error**	**Estimate**	**Standard Error**
ASC_Prawns[a]	5.054*	0.197	5.054*	0.189
ASC_Shrimp	5.405*	0.197	5.405*	0.188
ASC_Lobster	4.212*	0.201	4.212*	0.194
P_Prawns[b]	-0.683*	0.027	-0.683*	0.026
P_Shrimp	-0.680*	0.026	-0.680*	0.024
P_Lobster	-0.508*	0.024	-0.508*	0.024
S_ASC_Prawn[c]	--		0.002	0.051
S_ASC_Shrimp	--		0.006	0.048
S_ASC_Lobster	--		0.001	0.046
Log-Likelihood	-3595.357		-3595.347	

* Statistically significant at the $p < 0.001$ level.
[a] ASC refers to the alternative (product) specific constant for that product.
[b] P refers to the product specific price.
[c] S_ASC refers to the derived standard deviation of the parameter distribution for the random parameters logit model.

Table 5. Heteroskedastic Extreme Value and Multinomial Probit Model Results for CE Responses, Germantown, 2004.

Parameter	Heteroskedastic Extreme Value		Multinomial Probit	
	Estimate	Standard Error	Estimate	Standard Error
ASC_Prawns[a]	3.196*	0.200	2.788*	0.158
ASC_Shrimp	3.472*	0.222	3.042*	0.176
ASC_Lobster	3.221*	0.246	2.772*	0.173
P_Prawns[b]	-0.362*	0.032	-0.327*	0.026
P_Shrimp	-0.380*	0.034	-0.340*	0.027
P_Lobster	-0.364*	0.039	-0.304*	0.026
Scale Parameters/Standard Deviations[c]				
Prawns	1.898*	0.208	1.000	--
Shrimp	1.770*	0.201	1.000	--
Lobster	1.192*	0.135	1.000	--
None	1.000	--	1.000	--
Log-Likelihood	-3613.025		-3577.917	

* Statistically significant at the $p < 0.001$ level.
[a] ASC refers to the alternative (product) specific constant for that product.
[b] P refers to the product specific price.
[c] Parameter estimates are scale parameters for the heteroskedastic extreme value model and standard deviations for the multinomial probit model, which were held constant at 1 for estimation.

Table 6. Nested Logit Model Results for CE Responses, Germantown, TN, 2004.

Parameter	Estimate	Standard Error
ASC_Prawns[a]	5.951*	0.233
ASC_Shrimp	7.231*	0.240
P_Prawns[b]	-0.785*	0.031
P_Shrimp	-0.859*	0.030
Attributes for the Buy/No-Buy Decision		
ASC_No-Buy	-1.101*	0.200
P_Lobster	-0.605*	0.016
Inclusive Value Parameters		
Buy	0.436*	0.046
No-Buy	1.000	--
Log-Likelihood	-3855.499	

* Statistically significant at the $p < 0.001$ level.
[a] ASC refers to the alternative (product) specific constant for that product.
[b] P refers to the product specific price.

Table 7. Bootstrapped Mean WTP and 95% Confidence Intervals of CE Model Results Using the Krinsky-Robb Procedure.

Product	CL[a]	RPL	HEV	MNP	NL
Prawns	$7.40 [7.26-7.54][b]	$7.39 [7.26-7.54]	$8.83 [8.40-9.32]	$8.53 [8.15-8.96]	$7.58 [7.44-7.71]
Shrimp	$7.91 [7.79-8.05]	$7.95 [7.82-8.08]	$9.16 [8.75-9.63]	$8.95 [8.60-9.36]	$8.42 [8.34-8.56]
Lobster	$8.29 [8.12-8.46]	$8.29 [8.13-8.46]	$8.88 [8.44-9.45]	$9.13 [8.73-9.59]	--

[a] Abbreviations refer to CL (conditional logit), RPL (random parameters logit), HEV (heteroskedastic extreme value), MNP (multinomial probit), and NL (nested logit).
[b] Numbers in brackets are 95% confidence intervals derived from the bootstrapping of 1,000 observation on the WTP from the model results using the Krinsky-Robb procedure.

Table 8. One-sided P-Values from the Combinatorial Method Comparison of WTP Distributions and Grocery Store Weighted Average Price Distributions.

	CL[a]	RPL	HEV	MNP	NL
			--- Prawns ---		
Store	0.16[b]	0.16	0.36	0.38	0.18
			--- Shrimp ---		
Store	0.03	0.82	0.65	0.03	0.24
			--- Lobster ---		
Store	0.00	0.01	0.01	0.00	--

[a] Abbreviations refer to CL (conditional logit), RPL (random parameters logit), HEV (heteroskedastic extreme value), MNP (multinomial probit), and NL (nested logit).
[b] P-values represent the p-value of a one-sided test of grocery store > model WTP. The one-sided p-value of WTP > grocery store is simply 1 – p-value reported in the table. A two-sided test for statistical difference in the distributions is simply 2 * p-value in the table (Poe, Giraud, and Loomis).

Attribute	Farm-Raised Freshwater Prawns, 1 lb. Count: 23-45	Wild-Caught Marine Shrimp, 1 lb. Count: 23-45	Wild Caught Marine Lobster 1 lb.	None
Price	$5.99	$5.99	$6.99	
	⇩	⇩	⇩	⇩
I would choose..				

Figure 1. Example Choice Set, Freshwater Prawn Experiment, Germantown, TN, 2004.