The Comparison of three Non-hypothetical Valuation Methods: Choice Experiments, Contingent Valuation, and Experimental Auction

Jing Xie
jxie@ufl.edu
PhD Student
Food and Resource Economics Department
University of Florida

Zhifeng Gao
zfgao@ufl.edu
Assistant Professor
Food and Resource Economics Department
University of Florida

Selected Poster prepared for presentation at the Southern Agricultural Economics Association (SAEA) Annual Meeting, Orlando, Florida, 3-5 February 2013

Abstract:

This article examines the preferences revealed by three non-hypothetical experiments. We found that WTP estimates from the choice experiment are the highest, followed by that of contingent valuation methods, and then experimental auctions. Our results also suggest that the discrepancies among estimates from the valuation methods can come from the heterogeneity of respondents' price bargaining aggressiveness.

Keywords:

Consumer preferences, choice experiment, contingent valuation, auction, price bargaining aggressiveness

1. Introduction

The inconsistency in product valuations between hypothetical and non-hypothetical experiments has been examined thoroughly in the literature. A wealth of evidence has indicated that individuals tend to over-state the amount of money they are willing to pay in the state preference survey compared to the elicitation experiments with real money exchanges, and thus results in hypothetical bias.

Researchers in economics and market behaviors have proposed many ways to reduce hypothetical bias. One way is to change the survey design such as using cheap talk (Cummings and Taylor 1999), or add some questions in the hypothetical experiment to get more information about consumers' real attitudes (Johannesson, Liljas, and Johannesson 1998; Champ and Bishop 2001); another way is to combine the hypothetical experiment with non-hypothetical experiment to obtain the calibration factors of the hypothetical bias (Fox, Shogren, Hayes, and Kliebenstein 1998; Johannesson, Liljas, and Johansson 1998; Norwood and Lusk 2011); more often, researchers use non-hypothetical experiment directly in order to avoid hypothetical bias (Carlsson and Martinsson 2001; Cameron *et al.* 2002; Chang, Lusk and Norwood 2009; Lusk, Fields and Prevett 2008; Johansson-Stenman and Svedsäter 2008; Loomis *et al.* 2009).

While economic theory suggests that all of these non-hypothetical valuation methods are incentive compatible and should give equivalent outcomes in estimating consumers' willingness to pay (WTP)，Lusk and Schroeder (2006) and Gracia, Loureiro, and Nayag (2011) have found disparities between experimental auction (EA) and real choice experiments (RCE).

Lusk and Schroeder (2006) compared experimental auction and real choice experiment. They study consumers' preference for genetic steak, guaranteed tender steak, natural steak, and USDA Choice steak using two non-hypothetical experiments, EA and RCE. They compared the estimated demand elisticities from each method and confirmed that the disparity existed between EA and RCE—the auction bids were lower than the choice prices, and the demand elasticites calculated from the two experiment methods were also inconsistent.

Continuing with Lusk and Schroeder (2006)'s research, Gracia, Loureiro and Nayag (2011) compared the same experiments—EA and RCE. They used storable product (ham) with animal welfare labels in order to decrease possible temporal bias in shopping associated with the use of fresh product. And the subjects in their EA only bid on one product to avoid bid affiliation problems. Besides all the differences this article did in the experiments, they found the same results as Lusk and Schroeder (2006) that the auction bids were lower than the WTP estimates from the choice data.

Both of the studies show that auctions bids were significantly lower than the WTP estimated from RCE. These articles have discussed that this disparity might come from the differences in the mechanism between EA and RCE, however, they didn't specified where exactly the differences came from, nor did any tests of it. In addition, both studies do not include contingent valuation methods in their comparison, which are one of the most important valuation methods in market and non-market product valuation. Because all the three non-hypothetical experiments are widely used, it is important to compare the WTP elicited from them. If the disparity does exist, how and why it exists and which one is better are critical for researchers who want to accurately estimate the real market behaviors.

In this article, we hypothesize that the WTP differences in EA and RCE could come from the fact that these experiments emphasize price attribute differently.

Price plays quite different roles in RCE, RCVM, and EA when eliciting consumers' WTP value. In EA, the bidding process—"how much are you willing to pay" draws consumers' full attention to the price. In RCVM, price is also the most important attribute as participants are asked if they are willing to pay a certain amount of money. The difference is that in EA participants in RCVM can bid any price value they want while in RCVM participants choose to "accept" or "don't accept" a given price. In the choice experiment, however, participants are facing the choices with price and non-price attributes at the same time. As a result, in EA or RCVM consumers are focusing on price levels and more likely to bid or choose a lower price while participants in RCE are more likely to consider all the attribute levels at the same time and accept a higher price choice. If we treat the respondent in the experiments as a buyer and the experiment host as a seller in the market, in EA experiment participants are the price makers, in RCVM the experiment designer is the price maker, but the respondents have chance to bargain back (accept the price or not), and in RCE the seller is the price marker in each choice set, the only thing respondents can do is to choose one of the option without bargaining the price (they have the option to choose "none of products"). Obviously, the bargaining powers that the respondents can implemented in the three methods EA, RCVM and EA are decreasing.

To test whether price bargaining (PB) behavior is the key factor that can explain the disparities among the three methods. We measured the participant aggressiveness level in PB and divided them into different groups. If different price emphasis among experiments is the reason for the disparity in WTP estimates, the differences between the WTP estimates from participants with different PB behavior in EA should be larger than that in the experiments, where respondent has less opportunity to implement the PB power. .

This article could contribute to the body of literature in many ways. First, this article includes RCVM in the non-hypothetical valuation method comparison. Second and most importantly, this article proposes and tests a potential reason why discrepancies exist among these incentive compatible methods. This article is structured as follows. Section 2 describes the experimental design for each method. Section 3 illustrates the estimation model of each elicitation value mechanism. Section 4 reports the results derived from each method. The last section is the conclusion.

2. Experimental Design

General experimental design

To test consistency of the results among three methods, we carefully designed RCE, RCVM, and EA to be comparable. We kept the feature of different experiments as similar as possible by using the same products and setting price attribute ranges as close as possible. Moreover, we recruited the experiment participants randomly enough so that their demographic characteristics are statistically indifferent across all the groups.

Experiment subjects were recruited in front of local grocery stores in Changsha, China in 2012. The preferences of residents in Changsha cannot represent the preferences of whole residents in China; however, the focus of this study is the difference among three experiments. As long as we could control demographic statistics equivalence among those different experiments, conducting experiments in multiple cities is not a necessity. Subjects were offered ￥20 (about three us dollars) to participate in an "orange juice preference experiment." All participants claimed to eat orange juice at least occasionally. Each participant attended only one of the three experiments. They were asked to indicate their preference for different type of orange juice—100% Not From Concentrated Orange Juice (NFCOJ), 100% From Concentrated Orange Juice (FCOJ), and 10% Orange Juice Drink (OJD). The 100% NFCOJ is fairly new product in the China market and not available everywhere in normal grocery store yet, and so far, only a local brand and several imported brand (including Florida's Nature from the United States) produce this type of orange juice. In real market, the product size of NFCOJ is usually larger than FCOJ and OJD, but in experiments we keep the size of products the same across choices (500ml) since it is easy for respondents to compare and choose. By including new product in our experiment could help us understand that how consumers react differently between new products and products they are very familiar with (such as FCOJD and OJD).

The information of each type of orange juice was introduced to participants in the experiment (Appendix). After they read the information, they took about 10 to 20 minutes to do the experiment and complete the survey on socioeconomic and demographic characteristics as well as their orange juice shopping and eating behavior. In the last section of the survey, we also asked some questions about their PB attitudes and previous PB behaviors in order to measure the aggressiveness of the PB behavior.

We used the framework of the Fishbein behavioral intention model (Fishbein and Ajzen 1975) to construct the aggressiveness measurement of the bargaining intention. The Fishbein behavior intention model uses a series of questions about attitudes, subjective norm, intentions, and behaviors to measure individuals' behavior intentions. We constructed a series questions on individuals' personal bargaining attitudes, bargaining intentions, and individuals' bargaining competitiveness in purchasing both cheap and expensive products to measure individuals' aggressiveness level in PB. The details of these questions are reported in in Table 1 and Table 2.

Our research interest in this article is whether there are statistically significant differences in WTP elicited for the orange juice among RCE, RCVM, and EA. If we reject the null hypothesis, which means we find significant differences in WTP among these methods, we then test if the aggressiveness in PB is the reason for the WTP inconsistency.

RCE design

We designed our choice sets with two attributes: price and product types. The attribute levels are reported in Table 3, and an example of a choice set is provided in Table 4. The average price levels of the orange juice products were chosen to be consistent with prices in local grocery stores. To determine which choice sets to present to respondents, we used "brand" experiment design in which each type of orange juice was treaded as a factor and was varied at four price levels. This design generated in total 4^3=64 full factorial choice sets. From this full factorial we selected 10 saturated choice sets by using SAS, and the D-efficiency is 81.91%.

At the very beginning, subjects were instructed the process of RCE step by step. To ensure the elicitation mechanism is theoretically incentive compatible, respondents were told that after finishing the survey, they will randomly draw a number through 1 to 10 to determine the binding shopping scenario and purchase the product they chose in that scenario. If they choose "none of them," they will leave without purchasing any orange juice. Respondents were explicitly informed that actual payment would occur for the binding scenarios and they should evaluate each scenario carefully, as each scenario had equally chance of being binding.

RCVM design

There are many types of contingent valuation (CV) methods. Most popular ones are single-bounded dichotomous choice (SBDC) CV model and double-bounded dichotomous choice (DBDC) CV model. Several studies (Whitehead, Groothuis and Blomquist 1993, 1994; Eklöf and Karlsson, 1997) discussed the biasness issue in a SBDC-CV. Therefore in this articlestudy we used DBDC-CV.

The hard part for designing an incentive compatible DBDC-CV (we will refer it as RCVM in the rest of this article) is to provide individuals incentive to reveal their true value instead of choosing lowest price if they are willing to buy it. In a traditional hypothetical process of double bounded dichotomous choice contingent valuation mechanism, the respondents are first asked if they are willing to pay a middle price. If the answer is "yes," they will be asked a higher price; if the answer is "no," they will be asked a lower price. In the RCVM with the real purchasing requirement, the best choice for individuals who want to buy the product is always say "no" to the first question and say "yes" to the second question (lower price). As the result, subjects can always purchase the product at the lowest price in the question by saying "no" to the first question and say "yes" to the second question. In this case, the experiment provides no incentive for them to reveal their true value of the product.

One way to solve this problem is to introduce "the secret market price" into the experiment. This way makes sure that subjects cannot always purchase the product for the lowest price bundle. In our experiment, they were informed about the process of the experiment and how to draw a "secret market price," and they've been told that the best strategy for them is to answer "yes" if the price is lower than their true value at the first stage. An example of RCVM question is provided in Table 5. The RCVM were conducted according to the following steps:

Step1. Subjects were asked if they were willing to pay a certain amount of money to purchase a type of orange juice. They can answer "Yes" if they thought this price was lower than or as same as their true value; answer "No" if they thought the price was higher than their true WTP.

Step 2. If the subject's answer was "Yes," he/she was asked a higher value that if he/she was willing to pay; if his/her answer was "No," he/she was asked a lower value that if he/she was willing to pay. If in both cases, his/her answer was "No", he/she would not purchase this product.

Step 3. After subjects finished the survey, we randomly drew a type of orange juice as the binding product.

Step 4. The subject randomly drew a number from a bowl as a "secret market price" for the binding orange juice. If the market price was higher than his/her accepted price, he/she could not purchase the product; if his/her acceptance price was higher than the market price, then he/she will purchase the orange juice for a price equal to the market price.

Respondents were explicitly informed that actual payment would occur for the binding scenarios and they should evaluate each scenario carefully, as each scenario had equally chance of being binding.

EA design

Among all the auction methods, we choose Becker, DeGroot, Marshak (BDM) experiment. BDM is a common and easy method for eliciting the willingness to pay. Under the BDM, an individual reports a bid for an item; the item’s price is then randomly drawn (respondents do not know the price range). If the bid is above the price, the individual receives the good and pays the drawn price. If the bid is below the price, the individual does not receive the good and pays nothing. The incentive of truth-telling in this mechanism is that truth-telling is a dominant strategy and therefore it is independent of risk attitudes and whether the individual is an expected utility maximizer.

Many studies show that the BDM is incentive compatible for non-random goods (Davis and Holt 1993; Rutström 1998; Irwin *et al*. 1998; Noussiar, Robin, and Ruffieux 2004; Shogren *et al.* 2001). And BDM is the easiest experiment auction to conduct since the respondents are randomly picked in front of a grocery store gate and it is hard to gather them together at the same time. Another reason we used BDM instead of other auction mechanism is that BDM is very easy to understand by participants. It is an individual decision-making mechanism instead of group decisions. Lusk and Schroeder (2006) mention that the auction mechanisms such as second price auction are unfamiliar to most individuals, and Plott and Zeiler (2005) also show that without significant training and experience, misperceptions could affect the valuation methods. Comparing to 2nd price or random nth price auction, BDM is easier for participant to understand and conduct, and they don’t have to worry about other individuals’ preference and bidding prices. Lusk, Alexander, and Rousu (2007) discuss a potential problem for BDM: the bids of people with relatively high values tend to have less deviation then that of people with relatively low value. But this problem would be occurred if the sample size the fairly large, and easily understandable process of BDM could do us a favor in reducing the misperception bias. The auctions were conducted according to the following steps:

Step 1. Subjects were asked to read the products information and experiment instruction carefully.

Step 2. Subjects wrote down the most they are willing to pay for each type of orange juice. If they don't want to purchase a certain type of product, they can fill ￥0 for this product.

Step 3. After subjects finish the survey, we randomly drew a type of orange juice as the binding product.

Step 4. Subject randomly drew a number from a bowl as a "secret market price" for the binding orange juice. When his/her bid for the binding orange juice was equal or higher than the price from the bowl, he/she purchased the orange for the market price; when the bid price was lower than the market price, subject couldn't purchase the product.

3. Model and Specification

The auction bids are continuous while the choices in RCE and RCVM are discrete. To make the result comparable, we convert the results from RCE and RCVM to continuous measurement—WTP values, so that we can compare them to the average bids from EA data.

WTP value from RCE

RCE is based on random utility theory (Hanemann 1984; Hanley *et al*. 1998; Hanley, Wright, and Adamowicz 1998). To determine the WTP values for each orange juice product from the RCE and RCVM and allow for heterogeneity in valuations, we use random parameter logit model. The utility level of the ith product for the nth respondent can be written as:

(1) $U_{ni} = V_{ni} + \varepsilon_{ni} = \alpha_{ni} + \beta_n p_i + \varepsilon_{ni}$.

where V_{ni} is the deterministic and ε_{ni} is the stochastic portion of utility, p_i is the price in the choice set or the bid offered in the contingent valuation, α_{ni} is the intrinsic preference of respondent that captures all the non-price attributes of product i, and β_n is the marginal utility of price.

The probability for consumer n choosing choice i under the random parameter assumption is:

(2) $P_{ni} = \int \frac{e^{V_{ni}}}{\sum_{k=1}^{J} e^{V_{nk}}} f(\alpha)\mathrm{d}\alpha$

where $f(\alpha)$ is the density function of random parameter α. The probability is a weighted average of the logit formula evaluated at different values of α, with the weights given by the density of $f(\alpha)$.

We estimated the model using the maximum simulated likelihood method. By assuming α as lognormal distributed coefficients, the estimate of the mean WTP for product i is obtained from $WTP_j = \bar{\alpha}/\beta_p$ (Revelt and Train 1998). The WTP is lognormal distributed as well.

WTP value from RCVM

Following Hanemann *et al.* (1991), let $i = 1, \ldots, N$ be the index for each respondent in the sample, the determinants of WTP as a vector, x_i, and assume a linear functional form for the WTP equation, the true WTP y_i^* can be written as:

(3) $y_i^* = x_i'\beta + u_i$,

where β is a vector of parameters, and u_i is a random error assumed to be normally distributed with mean 0 and standard deviation σ. Let B_i be the initial price, B_i^H be the higher second price when subject responds "Yes" to the first question, and B_i^L be the lower second price when subject responds "No" to the first question. There are four possible outcomes for each respondent: (i) both answers are "yes" (YY); (ii) both answers are "no" (NN); (iii) a "yes" followed by a "no" (YN); and (iv) a "no" followed by a "yes" (NY). Thus we have four indicator function $I_i^{YY}, I_i^{NN}, I_i^{YN}$, and I_i^{NY} such that:

(4) $I_i^{YY} = \mathbf{1}$(ith respondent's response is "yes-yes")

$I_i^{NN} = \mathbf{1}$(ith respondent's response is "no-no")

$I_i^{YN} = \mathbf{1}$(ith respondent's response is "yes-no")

$I_i^{NY} = \mathbf{1}$(ith respondent's response is "no-yes")

The log-likelihood function of RCVM model is:

$$(5)\ \ln L = \sum_{i=1}^{N}\left\{I_i^{YY}\ln\left[1-\Phi\left(\frac{B_i^H - x_i'\beta}{\sigma}\right)\right] + I_i^{YN}\ln\left[\Phi\left(\frac{B_i^H - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{B_i - x_i'\beta}{\sigma}\right)\right] + I_i^{NY}\ln\left[\Phi\left(\frac{B_i - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{B_i^L - x_i'\beta}{\sigma}\right)\right] + I_i^{NN}\ln\Phi\left(\frac{B_i^L - x_i'\beta}{\sigma}\right)\right\}$$

4. Results

Participants in the experiment were recruited in June, 2012 in local grocery stores in Changsha, China. Participants were offered ￥20 (about $3) to compensate their time consuming in the experiment. Each participant attended (were randomly assigned) only one of the three experiments. In total, 321 individuals agreed to participate in the experiments, and 290 of them (90.3%) have completed the experiments. Among these individuals, 107 completed the RCE, 107 completed the CVM, and 76 completed the EA. The null hypothesis of equality of means for demographic variables such as gender, age, and household income in the RCE, RCVM, and EA, cannot be rejected at any standard significance level, ensuring that the discrepancies, if it exists, are not coming from the demographic differences across experiments.

Random Parameter Logit Model results from RCE data

Table 6 reports the results of the random parameter logit estimates obtained from the RCE data. The Ransom Parameters Logit Model (RPL) using a pandel data structure to take into account the fact that each individual made 10 choices. All estimations were conducted using STATA 12.0 assuming that price is a fixed coefficient and the coefficients for three orange juice products are random following lognormal distribution. We assumed the coefficients of NFCOJ, FCOJ, and OJD following lognormal distribution because it is nature to assume that respondents who occasionally consume orange juice value orange juice positive comparing to "consuming nothing," thus the normality assumption of the random parameters is not reasonable here. The estimated price parameters in the first column of Table 6 are the mean (β) and standard deviation (sd) of the natural logarithm of the random parameters. The mean and standard deviation of the coefficient itself are given by $\exp(\beta + sd^2/2)$ and $\exp(\beta + sd^2/2)\sqrt{\exp(sd^2) - 1}$, respectively (Train 2003). We report the calculated mean and standard deviation of the coefficients in the third column of Table 6 as well. As we can see, the coefficient estimate of price is significantly negative, and all the coefficient estimates of orange juice are significantly positive, indicating that respondents are having positive value of orange juice drink comparing to consuming noting. Moreover, the coefficient estimate of NFCOJ is the highest, followed by FCOJ, and the coefficient of OJD is the least, implying that consumers prefer NFCOJ to other two types of orange juice. The coefficient estimates of all the standard deviation are significantly, reflecting the facts that consumers have heterogeneous preferences.

WTP values from three experiments

The calculated WTP values from all the three experiments are reported in Table 7. The first column reports the WTP value for three types of orange juice from RCE data. In RCE, using RPL we estimated the WTP values for NFCOJ, FCOJ, and OJD are ￥20.3, ￥16.3, and ￥7.9 respectively. However, these numbers dramatically dropped to ￥16.5, ￥9.2, and ￥3.2 in RCVM, and dropped further in EA to ￥11.9, ￥8.1, and ￥3.6, respectively. In most cases, estimates of WTP values in RCE is the highest, followed by RCVM, then by EA except one scenario that the average bids of OJD is a bit higher than the estimate WTP for OJD from RCVM data. This result is consistent with what Lusk and Schroeder (2006) and Gracia, Loureiro, and Nayag (2011) found in their studies that estimates of WTP from RCE data were significantly higher than bidding value from EA data, and this result also confirms our expectation that the order of WTP value (from high to low) is RCE, RCVM, and EA because the last two experiments are focusing on price attributes more.

Aggressiveness in PB

As we discussed before, the procedure of doing RCE, RCVM, and EA are quite different. In RCE, respondents are presented with some choice sets, and in each choice set, there are several choice options with different types of products. Respondents in RCE are deciding which option they want most by considering different level of attributes of each product. In EA, the story is totally different as respondents bid a value as the highest value they are willing to pay for each product, and "price" is in the center of the purchasing decision process. RCVM, comparing to EA and RCE, is like EA as price is also the focus of the process, but is also like RCE that RCVM presents the price choices to respondents to choose.

In this study, we propose that: consumers are heterogeneous in the aggressiveness level of PB; when there is a chance for them to bargain the price, some aggressive respondents will offer comparatively low price and push the estimate WTP value lower than their true WTP. Moreover, as we discussed before, the process of EA, respondent have the market power to bid whatever value they want, in CVM, respondents are no longer bidding whatever value they want but accept or leave with certain price level of the product (they have chance to counter the offer in double bounded CVM), and in RCE, respondents can only choose product within the choice sets. Obviously, in terms of deciding the price, respondents have more power in the EA than in the RCE. When respondents have more market power, they can be more aggressive in terms of the bidding a fairly low price for the product.

To test this idea, we need to identify and measure the aggressiveness level of respondents. Following Lee (2000), we asked a sequence questions to measure the aggressiveness level of respondents. Table 1 and Table 2 are questions we used after the experiments to measure respondents' aggressiveness in PB. We measured individuals bargaining attitude by the first four questions in Table 1, bargaining intention by the last two questions in Table 1, and bargaining competitiveness for low- and high-priced products in Table 2. We didn't include subjective norm of the PB here because here we don't need to compare culture difference between China and the United States as Lee (2000) did. All the choices of the questions were placed in order: 1 to 5 denote from extreme high aggressive to extreme low aggressive, thus when we analyzed choice results, the lower score, the higher the aggressiveness level the respondent has.

Using the K-means cluster method to cluster all the six measurements in Table 1 and Table 2, we categorized respondents of each experiment into low, middle, and high aggressive groups. The summary of the sum value of all the six measurements is reported in Table 8. As we can see, for all the high aggressive groups (from RCE, RCVM, and EA), the average of the total value of six measurement in each experiment were lower than 16.3, means most of them chose very aggressive choices in the questions; for all the low aggressive groups, the average value of the sum were above 24, means most people chose less aggressive choices. Here we only report the summary of the sum of all the measurements instead of summary each measurement, but all the measurements have consistent characteristics that the high aggressiveness group has the lowest mean value, and the low aggressiveness group has the highest mean value for each measurement.

Compare WTP values across aggressiveness levels

Table 9 reports the estimates WTP values from RCE data, average bids from EA, and estimates WTP values from RCVM data across aggressiveness levels in PB. As we discussed above, the difference mechanism of the experiments provide subjects different ability to bargain the price they are willing to pay. The first panel in Table 9 reports the WTP values of each group (low aggressive, middle aggressive, and highly aggressive groups) from RCE data. We can see that for all three different orange juice products, WTP value in highly aggressive group are lower than that in low aggressive group. We can find the same patterns in data from RCVM and EA data. All the t-tests for the mean difference between low aggressive group and highly aggressive group of each product in each experiment significantly reject the non-hypothesis that the mean are equal from different groups. This result indicates that the aggressiveness in PB significantly affects the estimate results of WTP value—highly aggressive participants bid/choose lower price in the experiments.

Most importantly, different experiments provided different gaps between low and highly aggressive groups. The last column of Table 9 reports the change proportion of the mean WTP between low and highly aggressive groups. In RCE, the change percentage ($(WTP_{low} - WTP_{high})/WTP_{high}$) was 15.1%, 14.5%, and 9.7% for NFCOJ, FCOJ, and OJD, respectively. These numbers extended to 38.3%, 14.5%, and 52.9% from RCVM data, and 24.2%, 27.6%, and 35.6% from EA data. The wider gaps between low and highly aggressive groups in EA and RCVM indicate that these two experiments provide more ability to participants to bargain the price, which pushes the average WTP values lower than it from RCE data.

The reason that there are discrepancies of WTP values among low aggressive groups across three experiments also can be explained by the mechanism differences among experiments. “Low aggressive” does not mean respondents are not aggressive in PB at all, in contrast, since not all of them were choosing the last option in the aggressiveness questions (Table 1 and Table 2), they might also bargain the price if there is a chance to do it. The only difference between low aggressive and highly aggressive people is, the low aggressive people is less likely to bargain or bargain less margin of the price, but does not mean they won’t bargain at all. So when it comes to the experiment like RCVM or EA, low aggressive individuals would also bid/choose lower price than they could do in RCE.

Another interesting founding is that when we compare low and middle aggressive groups, in RCVM the WTP values of middle aggressive group are slightly higher than the WTP values low aggressiveness group for FCOJ and OJD, but the gaps are very small. It could indicate that the difference between low aggressive and middle aggressive group is not that big in RCVM data. However, the significantly drop of the estimate WTP values in low aggressive group in RCVM data indicates that the aggressiveness in PB still matters.

5. Conclusion

Our study conducted and compared the estimate WTP values from three different incentive compatible experiments, RCE, RCVM, and EA. And we also analyzed the individuals’ aggressiveness levels in PB in the experiments and compared the WTP values by aggressiveness groups. We found that 1) WTP values from RCE data were the highest, followed by WTP values from RCVM, and the least, average bids from EA, which is consistent with the results in Lusk and Schroeder (2006) and Gracia, Loureiro, and Nayga (2011); 2) by grouping respondents into low aggressive, middle aggressive, and highly aggressive groups, we found that WTP values were significantly lower in highly aggressive groups; 3) moreover, the gaps of WTP values between low and highly aggressive group in RCVM and EA is higher than the gaps in RCE, indicating that highly aggressive people in RCVM and EA showing more aggressiveness than the highly aggressive people in RCE. Our task in this study is not to confirm which experiment is the best to reveal consumers’ WTP, but to illustrate that the different mechanism of experiments could trigger consumers behavior differently.

References:

1. Cameron, T., G. Poe, R. Ethier, and T. Schulze. 2002. Alternative Non-Market Value-Elicitation Methods:Are the Underlying Preferences the Same? *Journal of Environmental Economics and Management* 44: 391–425.
2. Carlsson, F., and P. Martinsson. 2001. Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments? *Journal of Environmental Economics and Management* 41: 179–192.
3. Champ, P. A. and R. C. Bishop. 2001. "Donation Payment Mechanisms and Contingent Valuation: An Empirical Study of Hypothetical Bias." *Environmental and Resource Economics* 19 (4): 383-402.
4. Chang, J. B., J. L. Lusk, and F. B. Norwood. 2009. "How Closely do Hypothetical Surveys and Laboratory Experiments Predict Field Behavior?" *American Journal of Agricultural Economics* 91 (2): 518-534.
5. Cummings, R. G. and L. O. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *The American Economic Review* 89 (3): 649-665.
6. Davis, D. D. and C. A. Holt. 1993. *Experimental Economics* Princeton Univ Press.
7. Eklöf, J. and S. Karlsson. 1997. Testing and Correcting for Sample Selection Bias in Discrete Choice Contingent Valuation Studies, Working Paper No. 171. Stockholm School of Economics, Sweden
8. Fishbein, M., and Ajzen, I. 1975. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, MA: Addison-Wesley.
9. Fox, J. A., J. F. Shogren, D. J. Hayes, and J. B. Kliebenstein. 1998. "CVM-X: Calibrating Contingent Values with Experimental Auction Markets." *American Journal of Agricultural Economics* 80 (3): 455-465.
10. Gracia, A., M. L. Loureiro, and R. M. Nayga Jr. 2011. "Are Valuations from Nonhypothetical Choice Experiments Different from those of Experimental Auctions?" *American Journal of Agricultural Economics* 93 (5): 1358-1373.
11. Hanemann, M., 1984. Discrete choice models of consumer demand. *Econometrica* 52: 541-561
12. Hanemann, M., Loomis, J., Kanninen, B., 1991. Statistical efficiency of double-bounded dichotomous choice contingent valuation*. American Journal of Agricultural Economics* 73 (5): 1255–1263.
13. Hanley, N., MacMillan, D., Wright, R., Bullock, C., Simpson, I., Parsisson, D., Crabtree, B., 1998. Contingent valuation versus choice experiments: estimating the benefits of environmentally sensitive areas in Scotland. *Journal of Agricultural Economics* 49: 1-15.
14. Hanley, N., Wright, R., Adamowicz, W., 1998. Using choice experiments to value the environment. *Environmental and Resource Economics* 11: 413-428.
15. Irwin, J.R., G.H. McClelland, M. McKee, W.D. Schulze, and N.E. Norden. 1998. “Payoff Dominance vs. Cognitive Transparency in Decision Making.” *Economic Inquiry* 36: 272-85.
16. Johannesson M., Liljas, B., and Johansson P.O., 1998. "An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions," *Applied Economics*, Taylor and Francis Journals 30(5): 643-647.
17. Johansson-Stenman, O. and H. Svedsater. 2008. "Measuring Hypothetical Bias in Choice Experiments: The Importance of Cognitive Consistency." *The B.E. Journal of Economic Analysis & Policy* 8 (1): 1898.
18. Lee, D.Y. 2000. “Retail Bargaining Behavior of American and Chinese Customers.” European Journal of Marketing, 34:190-206.
19. Loomis, J., P. Bell, H. Cooney, and C. Asmus. 2009. A Comparison of Actual and Hypothetical Willingness to Pay of Parents and
20. Non-Parents for Protecting Infant Health: The Case of Nitrates in Drinking Water*. Journal of Agricultural and Applied Economics* 41(3): 697–712.
21. Lusk, J. L., C. Alexander, and M. Rousu. 2007. "Designing Experimental Auctions for Marketing Research: The Effect of Values, Distributions, and Mechanisms on Incentives for Truthful Bidding." *Review of Marketing Science* 5 (1): 1-30.
22. Lusk, J., D. Fields, and W. Prevett. 2008. An Incentive Compatible Conjoint Ranking Mechanism. *American Journal of Agricultural Economics* 90(1): 487–498.
23. Lusk, J. L. and T. C. Schroeder. 2006. "Auction Bids and Shopping Choices." *The BE Journal of Economic Analysis & Policy*.

24. Noussair, C., S. Robin and B. Ruffieux. 2004. “Do Consumers Really Refuse To Buy Genetically Modified Food?” *Economic Journal* 114:102-121.
25. Plott C. R. and Zeiler, K. (2005) “The Willingness to Pay-Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions, and Experimental Procedures for Eliciting Valuations” *The American Economic Review* 95(3): 530-545
26. David Revelt & Kenneth Train, 1998. "Mixed Logit With Repeated Choices: Households' Choices Of Appliance Efficiency Level," The Review of Economics and Statistics, MIT Press, 80(4): 647-657.
27. Rutström, E.E. “Home-Grown Values and Incentive Compatible Auction Design.” *International Journal of Game Theory* 27(1998):427-41.
28. Shogren, J.F., Cho, S., Koo, C., List, J., Park, C., Polo, P., Wilhelmi, R. (2001). “Auction mechanisms and the measurement of WTP and WTA.” *Resource and Energy Economics* 23: 97-109.
29. Train, K. 2003. "Discrete Choice Methods with Simulation," Online economics textbooks, SUNY-Oswego, Department of Economics, number emetr2.
30. Whitehead, J. C., P. A. Groothuis and G. C. Blomquist (1993), ‘Testing for Non-Response and Sample Selection Bias in Contingent Valuation: Analysis of a Combination Phone/Mail Survey. *Economics Letters* 41: 215–220.
31. Whitehead, J. C., P. A. Groothuis, T. J. Hoban and W. B. Clifford (1994), ‘Sample Bias in Contingent Valuation: A Comparison of the Correction Methods’, *Leisure Science* 16: 249–258.

Appendices:

Information about three types of orange juice or organic drink:

1) Not from Concentrated Orange Juice (NFCOJ): Is orange juice processed and pasteurized by flash heating immediately after squeezing the fruit without removing the water content from the juice. No additional water or other ingredients are added in 100% NFCOJ. There are only a few NFC orange juice products in the Chinese market such as Paisengbai NFC orange juice and some imported brands such as NFC orange juice from Florida and Australia. Now the price of a bottle of 250ml 100% NFCOJ ranges from ￥5 to ￥12.

2) From Concentrated Orange Juice (FCOJ): Is orange juice obtained from concentrated juice (COJ) that is reconstituted with water. FCOJ is orange juice made by removing, through evaporation, the water from the orange juice of fresh, ripe oranges that have been squeezed in extraction machines. No other ingredients are added in 100% FCOJ except for the same amount of water that was evaporated. So far, FCOJ has the biggest market share in China. For example, Huiyuan 100% FCOJ, Farmer's Orchard 100% FCOJ, and Great Lake 100% FCOJ are very common in the market. The price for a bottle of 450ml 100% FCOJ ranges from ￥4 to ￥8.

3) Orange Juice drink (OJD): Is sweetened beverage that is made of diluted fruit juice containing no less than 10% orange juice with other ingredient such as sweetener added. OJD is also very popular in the orange juice drink market. You can find OJD in the market very easily. Minute Maid, Uni President, and Master Kong are the common brands which carry orange juice drinks. The price for a bottle of 450ml OJD ranges from ￥1 to￥5.

Tables:

Table 1 Bargaining Attitudes and Intention

People always have different attitude towards bargaining behavior. Please indicate how you agree (or disagree) with the following statements.

	1) Strongly Agree	2) Agree	3) Neutral	4) Disagree	5) Strongly Disagree
1. Bargaining gives me the pleasure of shopping.					
2. Bargaining makes my life interesting.					
3. Sometimes it is not about money, if I can get discount by bargaining, I enjoy it and feel happy.					
4. I feel comfortable when I bargain.					
5. Whenever I go shopping I would try to bargain if bargaining is possible.					
6. I will bargain during my next shopping trip if bargaining is possible.					

Table 2 Bargaining Competitiveness

Suppose you want to buy the following products. If bargaining is possible, please indicate the price level that closest to your bargaining price:

1. A simple cotton T-shirt, price ￥20	1)￥10	2)￥13	3)￥15	4)￥17	5)￥20
2. A regular desk computer with all the common features you need, price ￥4000	1)￥3000	2)￥3300	3)￥3500	4)￥3700	5)￥4000

Table 3 Attributes for choice experimental design

Product	Price levels
NFCOJ, 500ml	￥17, ￥21, ￥25, and ￥29
FCOJ, 500ml	￥6, ￥8, ￥10, and ￥12
OJD, 500ml	￥2, ￥2.5, ￥3, and ￥3.5

Table 4 An example of RCE choice set

In these 4 choices, I would choose…

A. 500ml 100% NFCOJ ￥21	B. 500ml 100% FCOJ ￥8	C. 500ml 10% OJD ￥3	D. None of these

Table 5 An example of RCVM question

Are you willing to pay ￥10 for a 500ml bottle of 100% FCOJ?

A. Yes B. No

If your answer is "Yes," are you willing to pay ￥12 for it?

A. Yes B. No

If your answer is "No," are you willing to pay ￥8 for it?

A. Yes B. No

Table 6 Random Parameter Logit Estimates from RCE Data: Means and Standard Deviations of Random Parameters

Variable	ln(coefficient)	St. Err. of ln(coefficient)	Coefficient	Standard Error[a]
Price			-0.344*	0.031
Means of Random Parameters				
NFCOJ	1.895*	0.105	6.974*	0.690
FCOJ	1.709*	0.067	5.611*	0.378
OJD	0.8913*	0.112	2.720*	0.314
Standard Deviations of Random Parameters				
NFCOJ	0.306*	0.038	2.182*	0.264
FCOJ	0.179*	0.037	1.011*	0.232
OJD	0.468*	0.107	1.345*	0.415
Number of Observations			760	
Log likelihood			-659.423	

Note: 1. * represents statistical significance at the 0.01 level.

2. Standard errors are calculated by Delta Method.
3. We used 100 Halton draws in the Maximum likelihood Estimation.

Table 7 Compare WTP from EA, RCE, and RCVM

WTP	RCE	RCVM	EA
NFCOJ	20.262	16.464	11.850
	(1.141)[a]	(1.21)[a]	(7.394)[a]
FCOJ	16.302	9.191	8.089
	(0.901)	(0.39)	(3.698)
OJD	7.903	3.174	3.556
	(0.992)	(0.21)	(1.825)
No. of observations	760[b]	107	107

Note: [a] Numbers in the parentheses are standard error calculated by Delta method.

[b] Number of observations is 760 (=76 individuals ×10 choices each)

Table 8. Summary of Aggressiveness Index in Each Experiment Data

	RCE	RCVM	EA
Low Aggressive	24.409[a]	24.086	25.241
	(2.518)[b]	(3.043)	(3.124)
Middle Aggressive	19.333	18.131	19.974
	(2.056)	(1.961)	(1.716)
High Aggressive	16.167	15.152	16.225
	(3.146)	(2.093)	(2.587)

Note: [a] Mean of the aggressiveness index
[b] The numbers in the parentheses are standard deviation calculated by conventional manner.

Table 9. WTP values by Aggressiveness Level in PB

		RCE		
	Low Aggressive	Middle Aggressive	Highly Aggressive	Change
WTP for NFCOJ	21.355	20.775	18.555	15.1%
	(1.903)	(1.770)	(1.945)	
WTP for FCOJ	17.024	17.005	14.872	14.5%
	(1.856)	(1.388)	(1.504)	
WTP for OJD	8.706	8.527	7.934	9.7%
	(1.659)	(1.376)	(1.581)	
Number of observations	220	300	240	
		RCVM		
	Low Aggressive	Middle Aggressive	Highly Aggressive	% Change
WTP for NFCOJ	18.829	16.490	13.612	38.3%
	(1.89)	(1.82)	(2.69)	
WTP for FCOJ	9.351	10.106	8.101	14.5%
	(0.76)	(0.77)	(0.54)	
WTP for OJD	3.933	4.018	2.573	52.9%
	(0.72)	(0.76)	(0.17)	
Number of observations	35	38	33	
		EA		
	Low Aggressive	Middle Aggressive	Highly Aggressive	% Change
Average bids for NFCOJ	13.448	11.711	10.825	24.2%
	(8.249)	(6.102)	(7.838)	
Average bids for FCOJ	9.138	8.250	7.175	27.6%
	(4.340)	(3.830)	(2.834)	
Average bids for OJD	4.086	3.724	3.0123	35.6%
	(2.049)	(1.638)	(1.719)	
Number of observations	28	39	40	