

RESEARCH ARTICLE

# Bidding behavior in auctions versus posted prices: comparisons of mean and marginal effects

Shang Wu<sup>1</sup>, Jacob R. Fooks<sup>2</sup>, Tongzhe Li<sup>3</sup>, Kent D. Messer<sup>4\*</sup>  and Deborah A. Delaney<sup>5</sup>

<sup>1</sup>Center for Experimental and Applied Economics, University of Delaware, Newark, DE, USA, <sup>2</sup>Cigna Inc., Wilmington, DE, USA, <sup>3</sup>Food, Agricultural & Resource Economics, University of Guelph, Guelph, ON, Canada, <sup>4</sup>Department of Applied Economics and Statistics, University of Delaware, Newark, DE, USA and <sup>5</sup>Department of Entomology and Wildlife Ecology, University of Delaware, Newark, DE, USA

\*Corresponding author. Email: [messer@udel.edu](mailto:messer@udel.edu)

(Received 6 May 2020; revised 11 March 2021; accepted 12 March 2021)

## Abstract

Economic experiments have been widely used to elicit individuals' evaluation for various commodities. Common elicitation methods include auction and posted price mechanisms. A field experiment is designed to compare willingness-to-pay (WTP) estimates between these two mechanisms. Despite both of these formats being theoretically incentive compatible and demand revealing, results from 115 adult consumers indicate that WTP estimates obtained from an auction are 32–39 percent smaller than those from a posted price mechanism. A comparison in statistical significance shows that auctions require a smaller sample size than posted price mechanisms in order to detect the same preference change. Nevertheless, the signs of marginal effects for different product characteristics are consistent in both mechanisms.

**Keywords:** auction experiments; economic valuation; homegrown values; posted price markets; willingness to pay

**JEL codes:** D44; D12; C93

## Introduction

Economists frequently use auctions in experimental economics settings to measure consumers' preference for goods and services (List and Gallet 2001; Lusk, Feldkamp, and Schroeder 2004; Lusk and Shogren 2007). From a theoretical standpoint, most of these auctions (such as Vickrey, English, and Becker–DeGroot–Marschak (BDM)) are incentive compatible, meaning that participants have the incentive to bid their true willingness to pay (WTP)—the maximum amount they would be willing to spend on a product in a real market environment. Since a bid obtained from an auction is a point estimate of WTP, auctions are an attractive method as the data they generate is

easier to work with econometrically and provide more efficient estimates than the information obtained from other methods such as yes/no decision in a posted prices<sup>1</sup> format. Thus, it has become natural to emphasize auctions as a first-line valuation tool.

However, in the real world, consumers generally make purchasing decisions in a posted price setting, such as a grocery store or Amazon.com. In this environment, participants are price-takers, where they make yes/no decisions on whether or not to purchase an item at a given “posted” price. Therefore, what is implicit in the decision to use auction to understand consumer WTP is that, while some error in the value elicitation process may be inevitable, the WTP estimates from these auctions have applicability to decisions made in the more common postprice markets. However, since typical consumers rarely use auctions as their primary shopping method, an open question is whether consumers behave in an auction the way that is consistent with how they behave in a posted price market.

Researchers have been using posted price questions both in laboratory and in field settings to mimic a more common market setting. In this design, participants are asked a question: “Are you willing to purchase this item at \$A?” Participants will spend \$A to purchase the item if they choose “Yes”, while they will not get the item nor pay anything if they choose “No”. Since this framing of the purchase question resembles decisions consumers make every day about purchasing items at different posted prices, the design is easy for participants to understand. However, a disadvantage is that the experiment does not elicit the exact WTP for each participant—instead, it only indicates if WTP is above or below a certain value. Consequently, the mechanism is less statistically efficient and requires large sample sizes to produce the same level of precision as other methods (Loomis *et al.* 1997) such as auctions. Thus, a natural question arises: How do these two mechanisms compare?

In this research, we compare the estimated WTP from a variation of an English auction (Bernard 2006) and posted price questions. The experiment provided adult participants the opportunity to purchase jars of honey under these two mechanisms. Potential explanations of the discrepancy were also investigated. Furthermore, we compare the difference in marginal effects drawn from these two mechanisms.

Our results suggest that estimated mean WTP in auction is smaller than the posted price mechanism WTP in the range of 32–39 percent. We then seek to explain this result by testing different possible explanations. We found no evidence of anchoring effects. We did find evidence suggesting that the cause of low WTP estimates from auctions is associated with consumers’ lack of familiarity with auctions, as well as the lack of engagement of off-margin bidders. In terms of the marginal effects of different product attributes, we find that the auction and posted price mechanisms provide consistent signs, which indicates that consumer preferences for different product attributes do not vary with the elicitation methods. While the signs of coefficients are consistent, the significance level is much higher in auctions. Therefore, a posted price mechanism requires a larger sample size to detect the same preference change.

## Literature review

### *Auctions and posted price*

Approaches involving incentive-compatible auction mechanisms (e.g., Vickrey, English, BDM, and random  $n^{\text{th}}$  price) are widely used in experimental economics research to

<sup>1</sup>Or sometimes referred to as dichotomous choice or take-it-or-leave-it (TIOLI).

elicit values for consumer WTP as they provide a point estimate of WTP for each participant (Vickrey 1961; Becker, DeGroot, and Marschak 1964; Shogren et al. 2001). An auction is considered to be theoretically incentive compatible if the dominant strategy for participants is to bid their true values.

In the context of private-value auctions, where each participant knows what the item is worth to her but is uncertain of its value to other participants, both Vickrey and English auctions are theoretically incentive compatible (Vickrey 1961). It has also been shown that the Vickrey auction and English auction generate roughly equivalent results (Lucking-Reiley 1999). This study implements an English auction with the sealed bids feature of the Vickrey auction (Bernard 2006; Dillaway et al. 2011). The auction is theoretically incentive compatible, and Bernard (2006) demonstrated empirically that the participants were able to better learn and retain the dominant strategy of this auction mechanism, compared with traditional English and Vickrey auctions.

Economists have also adopted various posted price mechanisms in evaluation studies. For example, double-bounded posted price questions are widely used to elicit consumer WTP for new technologies (Li and McCluskey 2017), and researchers have studied how to best implement such a mechanism (e.g., Yoo and Yang 2001). In experimental settings with real monetary incentives, a single-bounded posted price format has become popular recently (Venkatachalam 2004; Li, Keciniski, and Messer 2017). This is mainly because posted price choice activities are easy to implement, especially in field experiments that usually happen in real marketplaces surrounded by many distractions.

### *Comparisons of posted prices and auctions*

The results are mixed from the few studies that have empirically compared relative WTP from posted price and auctions using incentive-compatible methods. In a recent article, Berry, Fischer, and Guiteras (2020) compared WTP elicited from BDM and TIOLI (take-it-or-leave-it or posted price) mechanisms in a field setting in Ghana. It is reported that the average WTP for the baseline in the BDM was approximately 17 percent lower than that from TIOLI. Furthermore, the WTP estimates from the BDM were consistently lower than those from TIOLI in the actual rounds as well as the practice rounds, where different items were used. Frykblom and Shogren (2000) compared a nonhypothetical posted price question to a Vickrey auction using a market good and did not find significant differences in resulting WTP estimates of the two methods. However, that experiment lacked training and practice rounds for the participants, which helps the participants to understand that the dominant strategy is to bid their true value (Lusk, Feldkamp, and Schroeder 2004).

It is worth noting the stream of literature on the comparison of auctions and real choice experiments (RCE). In RCE, participants are presented with combinations of products at different price levels and are asked to choose the one they prefer most. Posted price could be viewed as a special case of RCE when there is only one choice combination. Most studies on this topic found that empirical WTP from RCE are significantly higher than WTP from auctions (Lusk and Schroeder 2006; Gracia, Loureiro, and Nayga 2011), but other studies found contradicting results (Su et al. 2011), or no significant difference (Banerji et al. 2018). Corrigan et al. (2009) found that RCE provided more stable WTP estimates compared with auctions. RCE and posted prices are similar in the way that consumers make decisions rather than submitting bids. However, they differ significantly since participants in posted price make separate

decisions for each product, while in RCE, participants make choices based on various product-price combinations.

Besides using auction and posted price mechanisms to elicit valuations for a product or environmental service, economists are often interested in the marginal effects of explanatory variables estimated in different applications. For example, many researchers and policymakers are interested in the WTP premium for specific environmental attributes in a product, such as the location (Wu *et al.* 2015) and growing methods (Loureiro, McCluskey, and Mittelhammer 2002). Surprisingly, little attention has been paid to comparing the marginal effect estimates between these two elicitation mechanisms. We fill this gap by including a comparison of the marginal effect estimates.

### *Potential explanations of the WTP difference*

In the existing literature, several possible reasons on what might have caused the difference in WTP estimates of auction and posted price have been investigated. These candidate explanations include the anchoring effect, the asymmetric inconsistent preferences effect, and the lack of familiarity with auction formats.

The anchoring effect (also known as starting-point bias) occurs when respondents' valuations are influenced by and biased toward the posted offer in posted price questions (Tversky and Kahneman 1974; Herriges and Shogren 1996). This anchoring effect could influence both the decisions in the posted price setting and the subsequent auction bids (Ariely, Loewenstein, and Prelec 2003). While Frykblom and Shogren (2000) did not observe any anchoring effect in posted price decisions and Kriström (1993) as well as Berry, Fischer, and Guiteras (2020) observed no anchoring effect in the auction bids, Green *et al.* (1998) found strong evidence of anchoring on both tasks.

The asymmetric inconsistent preferences effect originates from the “yea-saying” effect in the contingent-valuation literature that describes a tendency for some respondents in hypothetical choice settings to choose affirmatively in a posted price setting regardless of their true preferences (Couch and Keniston 1960; Ready, Buzby, and Hu 1996). Therefore, it leads to an overestimation of overall WTP in the posted price setting. For instance, Kanninen (1995) concluded that 20 percent of respondents in the sample were “yea-sayers”. Ready, Buzby, and Hu (1996) found similar evidence with 20–22 percent of the respondents being “yea-sayers” in a split sample contingent-valuation study for food safety improvements. However, as Frykblom and Shogren (2000) noted, “nay-saying” has received little attention and seems to have been generally ignored in the contingent-valuation literature, while this effect would lower WTP from posted price settings. In the posted price setting with real economic incentives, it is possible that similar effects might still be present. If these effects resulted in differences in WTP estimates between posted price and auction, we could treat the auction bids as the “undisturbed preferences” and test whether the participants deviated significantly to one side from the bids. For example, one inconsistency resulting from “yea-saying” would be when the auction bid is lower than the posted price offer, but the participant accepted the posted price; the inconsistency resulting from “nay-saying” would be when the auction bid is higher than the posted price offer, but the participant rejected the price. These two inconsistent preferences would cause WTP discrepancies between posted price and auction if their effects were asymmetric.

A lack of familiarity with auction is another potential explanation. Plott (1996) in the discovered preference hypothesis (DPH) casts economic decision-making as a

process of discovery that assumes that participants have stable underlying preferences that are consistent with expected utility maximization. If there is appropriate feedback, decision-making converges to expected utility behavior in a series of three steps, starting with myopic “impulsive” behavior and gradually advancing to behavior that is more systematic as the decision-maker gains additional information through familiarization and feedback. As the NOAA panel pointed out, open-ended questions typically lack realism and are sensitive to trivial characteristics of the scenario presented. In contrast, posted price questions better approximate an actual purchasing environment and are easier for respondents to answer accurately (Arrow et al. 1993). Although one cannot claim that either posted price or auction reveals the “correct” WTP, posted price is obviously the format that is more familiar, easier to understand, and similar to a real-world purchasing decision. Familiarity with auctions is a form of institutional information and choice framing, and many consumers may not be familiar with auction formats because they do not routinely participate in any form of auction. In that case, we would expect to see an experience effect as an auction’s rounds progress. Empirically, Corrigan et al. (2009) observed such a round effect in uniform-price Vickrey auctions, suggesting that more effort must be taken to familiarize participants in auctions. Berry, Fischer, and Guiteras (2020) also mentioned that misunderstanding or uncertainty of the BDM mechanism in a field experiment may contribute to a WTP estimate discrepancy from the posted price.

Beside the explanations offered in the literature, it is also possible that participant engagement plays a role in repeated auction experiments. Roosen, Marette, and Blanchemanche (2010) explored how BDM compares with a discrete choice mechanism that evaluates WTP by measuring the propensity of substitution between two goods and found that differences in WTP disappear when considering only engaged bidders with non-zero bids. Even though the discrete choice mechanism is more similar to RCEs than posted price since participants are making a series of choices between two goods with different price vectors, the finding demonstrates that bidder engagement may impact the elicited WTP in auctions. Similarly, Berry, Fischer, and Guiteras (2020) found in a field experiment with BDM that participants with less at stake may not take the auction seriously.

### *Contribution to the literature*

Our study contributes to the literature in several ways. First, we carefully design an experiment that avoids many drawbacks of existing ones in the literature. Instead of using the student population, we made efforts to target on primary shoppers of households since it has been shown that experience with the good can reduce market anomalies (List 2003). Compared with the literature, our experiment includes more extensive training, including written instructions, oral presentations, and two training rounds to give participants better understanding on their tasks. In a setting with unfamiliar tasks, extensive training is necessary because even if subjects are told it is in their best interest to bid their “true value”, subtle misconceptions about how the elicitation mechanism works might trigger subjects to default to the strategies associated with familiar auctions (Plott and Zeiler 2005). Moreover, we argue that for our purposes, running an experiment in a more controlled environment in terms of information and feedback introduces less noise into participants’ decision-making process compared with a field setting (Plott 1996). Second, we test if discrepancies exist using both a within-subject and a between-subject design. Compared with the literature where only one kind of comparison is used, combining both within- and between-subject design adds robustness to our results. Third, we

**Table 1.** Sample demographic characteristics

Variable definition	Mean	St. Dev.	U.S. average
Gender			
1 = female; 0 = male	0.77	0.42	0.51
White	0.74	0.44	0.76
Age (years)	41.93	14.27	37.9
Years of education	16.39	2.85	NA
Household yearly income	\$76,086	\$48,373	\$68,703
Primary shoppers	0.77	0.42	NA

introduce more flexible price vectors into the posted price section. Since prices vary randomly for each posted price question, we control for the possibility that consumers treat the price offers as a quality signal and, therefore, alleviates valuation being anchored to the price offers. Using flexible price vectors also avoids picking inappropriate price offers in the situation where it is difficult to form fixed price points or appropriate widths between each price point. Fourth, we explicitly test for several possible explanations for the discrepancy and provide our own explanation. Lastly, few studies have compared important findings generated by auctions versus posted price mechanisms other than the mean WTP. However, auctions are not mainly used to measure average WTP for products. Rather, they are often adopted to measure the relative WTP for product attributes, information and policy treatment effects, and heterogeneous demographic responses. Therefore, we further extend the research question to comparing the sign and statistical significance of coefficient estimates.

### Experimental design

We designed a homegrown-value artefactual field experiment in which we offered adult subjects the opportunity to purchase honey presented in a variety of jars. This research was conducted in an experimental economics laboratory at a large university in the Northeastern United States. We recruited 115 adult participants through various sources that included the university's online newspaper, local community meetings, emails to staff members, and the laboratory's Web site. We endeavored to recruit adult consumers rather than students so that the sample would better represent the community as a whole and to ensure that participants were experienced buyers (List 2003; Chang, Lusk, and Norwood 2009; Gracia, Loureiro, and Nayga 2011).

Table 1 describes the socio-demographic characteristics of the participants as well as a comparison with the U.S. Census. The average participant age was about 42 years. Most of the participants were female and it was expected and desirable, as we targeted primary shoppers in the households. About 74 percent of the participants identified themselves as white, which is close to the national average of 76 percent. The average household income was between \$70,000 and \$80,000 and the average number of years of education was 16. The relatively high education level and income among the participants likely reflects the population of a university town.

As with most experimental studies, sample representativeness is a concern. The participants in our sample may not be perfectly representative of the honey consumer

population. Meanwhile, there may be some sample selection bias as we likely oversample people who are interested in participating in research studies and are interested in scientific information.

Fifteen 1-h sessions were held, with participants receiving \$20 in cash and/or products for the session (\$5 showup fee and \$15 to be spent during the experiment). The participants were informed that they could keep any portion of the money that they did not spend and that they would be given the opportunity to purchase a jar of honey during the session. The participants received the money and products purchased at the end of the session.

At the beginning of the experiment, the administrator randomly assigned the participants to computer terminals equipped with privacy screens to ensure confidentiality. The participants were asked to read information about the experiments once they were seated. A presentation then was given to explain the steps involved and how to use the program. No communication among the participants was permitted, but the participants were welcomed to ask questions to the administrator at any time. Data was collected through the use of Excel files that were programmed with Visual Basic with Applications and stored in an Access database.

The experiment involved investigating the effects of labeling and packaging on consumers' WTP for honey products. Specifically, we tested WTP for honey of three origins (local, domestic, and international) that were each distributed to five types of jars that had different shapes but the same volume (12 ounces), making 15 jar/origin combinations. In the auction, participants bid on all 15 honey products. In the posted price rounds, they made purchase decisions for the five jars of U.S. honey only. Therefore, each participant made 20 honey-purchasing decisions in total. In this article, we limit our comparison of WTP estimates with purchases of U.S. honey because it is the most commonly sold item in grocery stores and is the most familiar with the general public. A set of labeled jars (Jar 1, Jar 2, ..., Jar 5) of honey was placed on the administrator's desk and on the desk of each participant throughout the experiment, and the participants were encouraged to examine the items, but not open the jars. Since the three types of honey (USA, international, and local) were indistinguishable in terms of appearance, we just displayed the U.S. honey due to desk space constraints. The sequence of the posted price experiment and the auction experiment was randomly determined before the session, and the order in which the products were presented was also randomized.

To address the concern of demand reduction, at the end of each session, only one of the 20 decisions made by the participants (15 in the auction and five in the posted prices) was selected at random to determine which product would be binding and used to calculate cash earnings (List and Lucking-Reiley 2000; Lusk, Feldkamp, and Schroeder 2004; Messer et al. 2010). This binding decision was selected by having a volunteer draw a labeled ball from a cage containing 20 balls, each representing one decision. In order to reinforce the understanding of this concept, demonstrations of how the ball would be drawn to determine the binding round were shown to the participants prior to them making any decisions. It was also emphasized that no decision was affected by prior or subsequent decisions. As explaining the dominant strategy to participants in homegrown-value experiments is regarded as "best practice" and is widely used, we also informed the participants that it was in their best interest to bid as close to the worth of the item to them as possible (Rutström 1998; Lusk, Feldkamp, and Schroeder 2004).

In the posted price experiment, the question to the participants was, "Are you willing to purchase Jar Y of U.S. honey at \$A?" The price of each product varied randomly for

each decision and was distributed uniformly between \$0 and \$15. The participants were informed that clicking “yes” was a decision to purchase the jar of honey at the posted price; clicking “no” meant they would not receive Jar Y nor pay the price.

In the auction, a number representing the participant’s bid for the item was shown on the screen in front of each participant. Once the auction started, this bid increased incrementally at a speed of about \$0.10 per second from \$0 to \$15.<sup>2</sup> The participants were asked to click the “withdraw from auction” button when they saw the bid representing the maximum amount they were willing to pay for the product displayed on the screen. When they indicated a desire to withdraw from the auction, a second box appeared that asked them to confirm the number on their screen as their bid. The participants could choose to restart the auction round (incremental ascending increases in the number) from \$0 and bid again or could confirm the bid and submit it. The auction stopped either when all participants’ bids were confirmed or when the bid reached the preset upper limit of \$15. The bids by each participant were stored in a database and the auction then proceeded to a new bidding decision.

To help the participants better understand the bidding procedure, two practice rounds were held first. The participants were given \$3 in the practice rounds and were asked to submit bids on a pencil and a ballpoint pen. In the practice auction, the winner and the second highest bidder were announced after each round. It was emphasized to the participants that the winner pays only the amount of the second highest bid so it was in their best interest to focus on determining their own value for the item and to bid as closely to that as possible.

After the practice rounds, the participants were asked to submit bids on different jars of honey following the same procedure. This research followed the proposed “best practice” in Harrison, Harstad, and Rutström (2004) to clearly train and inform the subjects that their dominant strategy is to bid their true value. At the beginning of each new purchasing decision, the participants were provided with the list of items already auctioned and the bids they submitted for each. After each decision, no feedback was given to the participants with regard to the winner or the winning price as a means of reducing market feedback (Corrigan *et al.* 2011). At the end of the session, the participants were asked to fill out a survey about their demographics background and consumer behavior.

The only announcement was the winner of the binding round at the end of the experiment. This was done by having a volunteer draw a ball to determine which of the 20 purchase decisions was binding. Each participant’s screen then displayed a chart showing their decisions and products. Based on this binding decision, the computer program calculated each participant’s earnings and products purchased (if any) and displayed them on that person’s screen to assist them in filling out receipts.

### Model and testable hypotheses

In this section, we describe the model and the hypotheses that we will be testing in the experiment. We proceed by first verifying if WTP estimate difference exists between the two value elicitation methods. We then test if the observed difference (if any) is a result of the interinfluence between the posted price and the auction parts. Next, we examine two other behavioral factors that may result in WTP estimate differences. We conclude by comparing the marginal effects offered by the two elicitation methods.

<sup>2</sup>Since participants started the program by themselves, the participants’ bids were not synchronized, making it impossible for other participants to know whether they stopped the program on a low or high bid.

**Table 2.** Hypotheses

Question	Hypothesis	Result
1. Is there a difference in response between the posted price mechanism and auction?	$H_0: WTP_{\text{Posted\_Price}} = WTP_{\text{Auction}}$	Reject—There is a difference between measured WTP
2. Is the difference due to the two tasks influencing each other?		
2.1. Is there a difference only comparing the first task completed?	$H_0: WTP_{\text{Posted\_Price} \text{Posted\_Price\_First}} = WTP_{\text{Auction} \text{Auction\_First}}$	Reject—There is a difference even for the first tasks completed
2.2. Is this difference due to the anchoring effect?	$H_0: \beta_p = 0$ (Posted Price First) $H_0: \beta_p = 0$ (Auction First)	Fail to reject—No evidence of anchoring
3. Is the difference due to behavioral factors?		
3.1. Is this difference due to asymmetric inconsistent preferences?	$H_0: \Pr(\text{Accept}=1   \text{ShouldAccept}=0) = \Pr(\text{Accept}=0   \text{ShouldAccept}=1)$	Fail to reject—No evidence of asymmetric inconsistent preferences
3.2. Is this difference due to a lack of familiarity with an auction setting?	$H_0: \beta_{\text{Auction, RoundNumber}} = 0$	Reject—There is evidence that the difference decreases with learning
4. Are the marginal effects comparable?	$H_0$ : The signs and significance levels are similar	The signs are similar, and significance levels are higher in auction

**Comparison of posted price and auction WTP elicitation**

The series of hypotheses tested in this research are summarized in Table 2. The first hypothesis is that the WTP estimates from the posted price mechanism equal to those from the auction.

$$H_0 : WTP_{\text{Posted\_Price}} = WTP_{\text{Auction}}$$

where  $WTP_{\text{Posted\_Price}}$  denotes WTP estimates obtained from the posted price questions and  $WTP_{\text{Auction}}$  denotes WTP estimates derived from the experimental auctions.<sup>3</sup>

The posted price generates binary responses, while the auction generates continuous bids. To make the two types of data comparable, the auction data can be transformed to simulated binary responses, or average WTP point estimates can be inferred from posted price responses. For consistency with the literature, we follow the procedure documented in Green et al. (1998) and Frykblom and Shogren (2000) where the auction data is transformed into synthetic binary responses and compared with the actual responses. Let  $b_{ij}$  denote the bid that participant  $i$  submit for good  $j$ , and  $p_{ij}$  denote the posted price offer of participant  $i$  for good  $j$ , and  $\delta_{ij}$  denote whether participant  $i$

<sup>3</sup>We refer WTP as a general term for inferred consumer behavior (in this case WTP for a product) instead of just the bids from auction. In this paper, it refers to the mean WTP inferred from different mechanisms.

responded yes in the posted price section for good  $j$ . Since each participant responded in both the auction and the posted price formats, we can compare their auction responses,  $b_{ij}$ , with the binary response that would be consistent with the prices they see in the posted price section,  $p_{ij}$ , for the same good. We generate a synthetic posted price choice response variable  $\delta'_{ij}$ , where  $\delta'_{ij} = 1$  if  $b_{ij} \geq p_{ij}$ ;  $\delta'_{ij} = 0$  if  $b_{ij} < p_{ij}$ . Here,  $\delta'_{ij}$  can be interpreted as when facing the price offers, what participants' response would be based on the bids they indicated. Theoretically, if the null hypothesis holds, we will not observe a significant difference in the WTP inferred from  $\delta$  and  $\delta'$ .

To test if  $\delta$  and  $\delta'$  significantly differ from each other, we perform both parametric and nonparametric tests. The advantage of a nonparametric test is that no distributional assumption is placed on the variables. Since  $\delta$  and  $\delta'$  are binary variables and since these are considered as paired observations, we use McNemar's nonparametric test (McNemar 1947).

Since nonparametric tests generally have lower power than parametric tests, we also do a parametric test assuming a normal/logistic distribution on the underlying WTP (Frykblom and Shogren 2000). Formally, we assume that a consumer's WTP,  $w$ , follows some probability distribution with  $\mu$  as the location parameter and  $\sigma$  as the scale parameter. We denote  $F$  as the cumulative distribution function and  $S$  as the survival (or duration) function. Therefore, for a given posted price offer  $p$ ,  $F(p) = \text{Prob}(w \leq p)$ ,  $S(p) = 1 - F(p) = \text{Prob}(w > p)$ ,  $f(p) = dF(p)/dp$ . So, the survival function, in this case, represents the probability that a "yes" response in the posted price format will continue above a given price. We estimate  $\mu$  and  $\sigma$  by maximizing the log-likelihood function  $L$ , which is written as

$$L = \sum (\delta_{ij} \log(S(p_{ij})) + (1 - \delta_{ij}) \log(F(p_{ij}))) \quad (1)$$

where  $\delta_{ij}$  is equal to 1 if participant  $i$  accepted posted price offer for the  $j^{\text{th}}$  object ( $p_{ij}$ ) and equal to 0 otherwise. The survival function for normal distribution is

$$S(p_{ij}) = 1 - F(p_{ij}) = 1 - \Phi\left(\frac{p_{ij} - \mu}{\sigma}\right) \quad (2)$$

For logistic distribution, the corresponding function is

$$S(p_{ij}) = 1 - F(p_{ij}) = \frac{1}{1 + \exp\left(\frac{p_{ij} - \mu}{\sigma}\right)} \quad (3)$$

The estimated mean for both distributions is  $\mu$ , the estimated variance for normal distribution is  $\sigma^2$ , while for the logistic distribution, it is  $\sigma^2 \pi^2/3$ .

From this maximum likelihood estimation, we would be able to infer the distribution of WTP that generated the posted price responses. To test if the estimated mean from the two samples is different, we follow the same test as Frykblom and Shogren (2000), which is recommended by Kimentia (1986):

$$Z_{w_1 - w_2} = (w_1 - w_2) / \sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)} \quad (4)$$

where  $Z$  is an approximately standard normal variable,  $w_k$  is the estimated mean in offer format  $k$ ,  $s_i^2$  is the estimated sample variance, and  $n_i$  is the sample size.

**Possible task interinfluences**

Since our experiment consists of different tasks within a subject, we address the most common problem for a within-subject design—the tasks potentially influencing each other. We do this in two ways: first, we test if the difference still exists if we only utilize the first task that each participant completed; second, we test for the anchoring effect to see if the bids in the auction are influenced by posted price offers that were presented to the participant.

*Testing for a difference using first task only*

The intuition is to make comparisons only from data of the first task that a participant did. Specifically, we estimate WTP only from participants who went through the auction first and compare with posted price WTP estimates from participants who did the posted price first. In this way, we are actually making a between-subject comparison. The procedure used for this test is similar to the one described earlier where we transfer auction data into yes/no responses and compare it with the posted price data. One issue in generating the synthetic yes/no responses is that there does not exist a corresponding relationship between the auction-first group’s bids and the posted price-first group’s price offers. Therefore, we use a complete combinatorial approach similar to the one suggested in Poe, Giraud, and Loomis (2005). For every auction bid (suppose  $n_1$  total observations), we generate a yes/no response according to every posted price offering (suppose  $n_2$  total observations), resulting in  $n_1 * n_2$  pairs of observations on bids ( $b$ ), synthetic yes/no ( $\delta'$ ), price offer ( $p$ ), and real yes/no response ( $\delta$ ). Next, we compare  $\delta$  with  $\delta'$  following the procedure discussed before.

*Testing for the anchoring effect*

We randomized the posted price offered for each decision in the experiment to control for a possible anchoring of participants’ valuation of each item to the posted price. However, posted price offers might still be affecting consumers’ value formation process in two ways. First, the WTP estimates from posted price could increase if the participant saw a higher posted price offer for the item (Frykblom and Shogren 2000). Second, the WTP estimates from the auction could be affected by the posted price offers if the subject participated in posted price first (Kriström 1993). We assume that if the underlying valuation of the product is changed by the posted price offer, it is likely reflected in both the posted price part and the auction part, meaning that the presence of the two presentations of anchoring effects is positively related. The design allows us to test for the second type of anchoring effect by a Tobit model that includes posted price offers as an independent variable. Since bids were limited to a range of \$0–\$15, a two-limit random-effects Tobit model was appropriate to analyze WTP.<sup>4</sup> The dependent variable is defined based on a latent variable  $y_{ijk}^*$  that cannot always be observed and is specified as

$$y_{ij} = \begin{cases} y_{ij}^* & \text{if } 0 < y_{ij}^* < 15 \\ 0 & \text{if } y_{ij}^* \leq 0 \\ 15 & \text{if } y_{ij}^* \geq 15 \end{cases} \tag{5}$$

<sup>4</sup>An OLS model without censoring gives very similar estimates.

For subject  $i$  and item  $j$ ,  $y_{ij}^*$  is limited to a value between 0 and 15 and depends linearly on  $X_{ij}$  via a parameter (vector),  $\beta$ . The following random-effects Tobit model was used:

$$\begin{aligned} y_{ij}^* &= \alpha + \beta X_{ij} + U_i + u_{ij} \\ &= \alpha + \beta_p \text{Posted\_Price}_{ij} + \beta_2 \text{Jar type 2}_{ij} + \beta_3 \text{Jar type 3}_{ij} + \beta_4 \text{Jar type 4}_{ij} \\ &\quad + \beta_5 \text{Jar type 5}_{ij} + U_i + u_{ij} \end{aligned} \quad (6)$$

where  $\alpha$  is the average bid for the entire population,  $U_i$  represents the individual random effects, and  $u_{ij}$  is the error term for individual  $i$  for product  $j$ . We also include a specification with bootstrap standard errors. The variables Jar type 2 through Jar type 5 are dummy variables indicating which item was auctioned.

Under the null hypothesis that there is no anchoring effect, we would expect that  $\beta_p = 0$  when the posted price section is before the auction. Meanwhile, when the posted price section is after the auction, it is expected that an anchoring effect should not be observed. We test these two hypotheses separately.<sup>5</sup> Specifically, as Hypothesis 2 in Table 1, we test:

- $H_0: \beta_p = 0$  when posted price is before auction  
 $H_0: \beta_p = 0$  when posted price is after auction

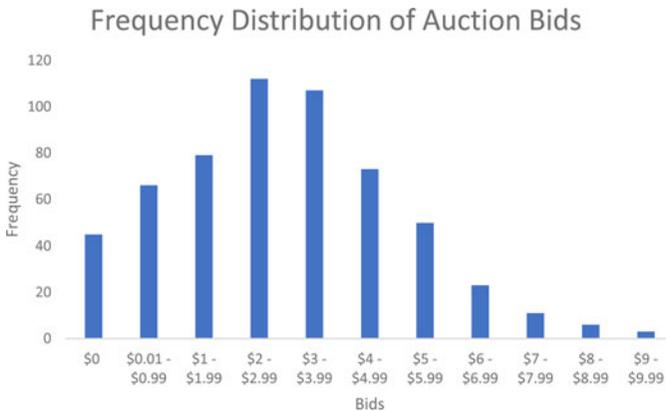
### Testing for behavioral factors

After testing for potential interinfluences between the tasks, we investigate behavioral factors that may result in WTP estimate differences between the two methods. As explained previously, asymmetric inconsistent preferences and the fact that participants are more unfamiliar with auctions may both lead to discrepancies in WTP estimates.

#### *Asymmetric inconsistent preferences hypothesis*

If asymmetric inconsistent preferences were the cause of the WTP discrepancy, we would expect to observe a difference in the following two inconsistencies: (1) when the bid is higher than the price and (2) when the bid is lower than the price. When a participant answers “yes” to a posted price question even though the price is higher than their bid, we define it as “affirmative inconsistent preference”. In contrast, when a participant answers “no” to a posted price question even when the price is lower than their bid, we define it as “negative inconsistent preference”. Affirmative inconsistent preference can be denoted as: WTP in posted price offer ( $p$ ) > bid in auction ( $b$ ), meaning when  $\delta' = 0$ ,  $\delta = 1$ . Negative inconsistent preference can be denoted as: WTP in posted price offer ( $p$ ) < bid in auction ( $b$ ), meaning when  $\delta' = 1$ ,  $\delta = 0$ . If the inconsistent preferences cause the WTP discrepancies, we would expect that one inconsistency would be more prevalent than the other. We test whether the probability of an affirmative inconsistent preference is larger than the probability of a negative inconsistent preference—specifically, whether  $\Pr(\delta = 1 \mid \delta' = 0) > \Pr(\delta = 0 \mid \delta' = 1)$ . If this hypothesis is rejected, it means that participants are not more likely to have affirmative inconsistent behavior than negative inconsistent behavior, and asymmetric inconsistent preference does not explain any discrepancy.

<sup>5</sup>Using one regression to include both orders with an interaction term would generate the same conclusion.



**Figure 1.** Frequency distribution of bid amounts in the auction

### *Familiarity hypothesis*

As compared with answering a posted price question, auction is a mechanism that is relatively unfamiliar with most participants. Even if participants do not receive direct feedback after each round, all information available to a participant may evolve due to additional opportunities for introspection, belief reinforcement, learning, and similar mechanisms. In that case, we would expect to see an experience effect as an auction's rounds progress. We test if roundnumber (the number of bidding decisions a participant has made) has an effect on the bids. Under the null,  $\beta_{\text{Auction, RoundNumber}}$  would be significantly different from 0. Specifically,

$$H_0: \beta_{\text{Auction, RoundNumber}} = 0$$

### *Comparing marginal effects in the two methods*

Despite any WTP estimate differences that may exist and the reasons that may cause the differences, in practical research, we are often not only interested in the absolute WTP estimates of a homegrown good, but also care about the marginal effects, or the ability that the estimation method is able to provide relative comparison conclusions on the effects of some particular attributes. When the research question is not about estimating absolute WTP values but instead about testing the marginal effects of attributes, it is important to learn if the two mechanisms provide similar results. We compare the marginal effect estimates on jar attributes from the two parts in terms of the signs and significance levels of the coefficients.

## **Results**

### *Descriptive statistics on bids and posted price*

A histogram on the frequency distribution of the bids is displayed in [Figure 1](#). The majority of bids was in the \$2.00–\$4.00 range. The mean of the bids was \$2.91 and the standard deviation was 1.97. [Figure 2](#) shows the frequency distributions of posted prices conditioning on whether the posted price was accepted or declined. As expected, the number of “Yes” choices decreases as prices go up, and the number



Figure 2. Frequency distribution of accepted and declined posted price offers

of declines increase as prices increase. In general, we do not observe fat tails in the distributions.

In total, there were 45 zero bids in the auction. Out of the 115 participants, four people (3.5 percent of the total participants) bid zero for all five auctions of honey. This seems to be a reasonable proportion of people who would not be interested in purchasing honey at any price. Of these four participants, three also declined the honey in all the posted price questions. So, their behavior appears to be generally consistent.

*Hypothesis 1: test for response difference,  $H_0: WTP_{posted\_Price} = WTP_{Auction}$*

As shown in Panel A of Table 3, for a within-subject comparison, the average of the actual binary response ( $\delta$ ) in posted price is 0.2904; the average of the generated synthetic binary response ( $\delta'$ ) based on bids in the auction and posted price offer is 0.1652. Since McNemar's  $\chi^2$  test statistic equals 150.45 and the corresponding p-value is less than 0.0001, we reject the null hypothesis.<sup>6</sup> This indicates that, compared with their actual behavior in posted price questions, participants are less likely to accept the same offer based on response inferred from their bids in auction.

As discussed earlier, we also do parametric tests assuming that the underlying WTP distribution is either normal or logistic (Panel B of Table 3). With a normal distribution, the estimated WTP from auction bids has a mean of 2.4889, while the estimated mean of WTP from posted price is 4.0587. A Z test rejects the null hypothesis that the two WTP means are equal. With a logistic distribution assumption, the results are similar. The estimated mean of WTP is 2.4579 for auction bids and 4.0570 for posted price. The Z test also rejects the null at 1 percent level. The results suggest that WTP estimate

<sup>6</sup>To our knowledge, these tests do not explicitly account for the panel nature of the data. To test the robustness of these results, we completed an additional test using only the responses to the first question of posted price and the corresponding auction bids. Thus, each participant has only one observation and the overall result still holds.

**Table 3.** Within-subject comparison of estimated WTP from posted price and auction

Panel A: nonparametric	Mean( $\delta$ ) (St. Dev.)	Mean( $\delta'$ ) (St. Dev.)	McNemar (p-value)
	0.2904 (0.4544)	0.1652 (0.3717)	150.45 (<0.0001)
Panel B: parametric assumption	WTP <sub>Auction</sub> (St. Dev.)	WTP <sub>PP</sub> (St. Dev.)	Z (p-value)
Normal	2.4889 (2.0898)	4.0587 (4.5021)	7.5838 (<0.0001)
Logistic	2.4579 (1.1698)	4.0570 (2.5562)	7.5199 (<0.0001)

from the auction is approximately 39 percent lower than that from posted price. The above results demonstrate that the WTP inferred from auction is significantly lower than that from posted price.

*Test for task interinfluences*

To address potential concerns that a within-subject design involving two tasks might influence each other, we test for the discrepancy using first task only and then test for anchoring effects between the tasks.

**Hypothesis 2.1: test for discrepancy using first tasks only,  $H_0: WTP_{Posted\_Price|Posted\_Price\_First} = WTP_{Auction|Auction\_First}$**  We conducted a between-subject comparison using data only from the first task each participant completes. In other words, we generate WTP estimates for auction from participants who did auction first and then generate WTP estimates for posted price from those participants who did posted price first. Since there is no one-to-one corresponding relationship between the bids and posted price offers, we do a complete combinatorial procedure (Poe, Giraud, and Loomis 2005) on bids and price offers to generate a synthetic binary response ( $\delta'$ ) and compare it with the corresponding actual binary response ( $\delta$ ). Again, a McNemar’s test rejects the null that the two series are equal, suggesting that participants are less likely to accept the same posted price offer based on response inferred from the auction (Table 4).

In a similar fashion, we conducted parametric tests assuming either normal or logistic distribution on the underlying WTP. Under normal distribution assumption, the estimated mean WTP is 3.171 for auction and 4.647 for posted price. Under the assumption of logistic distribution, the estimated mean WTP is 3.137 for auction

**Table 4.** Between-subject comparison of estimated WTP from posted price and auction

Panel A: nonparametric	Mean ( $\delta$ ) (St. Dev.)	Mean ( $\delta'$ ) (St. Dev.)	McNemar (p-value)
	0.3396 (0.4736)	0.2192 (0.4137)	14272 (<0.0001)
Panel B: parametric assumption	WTP <sub>Auction</sub> (St. Dev.)	WTP <sub>PP</sub> (St. Dev.)	Z (p-value)
Normal	3.1710 (2.1021)	4.6466 (4.9929)	78.0702 (<0.0001)
Logistic	3.1369 (1.1694)	4.6026 (2.8807)	74.4968 (<0.0001)

**Table 5.** Test for anchoring when posted price is before auction

	Random-effects Tobit			Random-effects Tobit with bootstrap St. Err.		
	Marginal effect	St. Err.	p> z	Marginal effect	St. Err.	p> z
Price	-0.0004	0.001	0.971	-0.0004	0.009	0.966
Jar type 2	0.459	0.117	0.000	0.458	0.087	0.000
Jar type 3	-0.044	0.117	0.711	-0.044	0.116	0.707
Jar type 4	0.110	0.117	0.347	0.110	0.109	0.314
Jar type 5	0.342	0.117	0.003	0.342	0.160	0.033
_cons	2.280	0.305	0.000	2.280	0.318	0.000
Wald $\chi^2$	28.33			43.33		
Prob > $\chi^2$	0.000			0.000		
Log-likelihood	-322.606			-322.606		
Number of observations	265			265		
Left-censored observations	31			31		
Uncensored observations	234			234		
Right-censored observations	0			0		

and 4.603 for posted price. In both cases, Z test rejects the null that the two estimated means are equal. This indicates that estimated mean WTP from auction is approximately 32 percent lower than that from posted price.

*Hypothesis 2.2: the anchoring effect.* We perform a test similar to that of Kriström (1993), where we examine if the respondents' auction bids are anchored to the posted price offers when they participated in the posted price setting first. To test  $H_0: \beta_p = 0$  when posted price were conducted first, we regressed the bids on posted price offers from sessions in which the posted price mechanism was conducted first (Table 5). The left panel of Table 5 reports a random-effects Tobit model, while the right panel reports the same model with bootstrapped standard errors included. As shown in both panels of Table 5,  $H_0: \beta_p = 0$  cannot be rejected, meaning we find no evidence that posted price in early rounds influenced subsequent bids in the auction.

Similarly, to test  $H_0: \beta_p = 0$  when auctions were held first, we regressed the bids on posted price offer from sessions in which the posted price mechanism was conducted first. Again, the left panel of Table 6 reports a random-effects Tobit model, while the right panel includes bootstrap standard errors. As both panels demonstrate, posted price offers do not have an effect on bids when auction was conducted first. Therefore, no anchoring effect is observed.

In sum, we conclude that WTP estimates from auction significantly differ from WTP estimates from posted price and it is not likely a result of the two tasks influencing each other but rather due to behavioral reasons.

**Table 6.** Test for anchoring when posted price is after auction

	Random-effects Tobit			Random-effects Tobit with bootstrap St. Err.		
	Marginal effect	St. Err.	p> z	Marginal effect	St. Err.	p> z
Price	0.004	0.012	0.714	0.004	0.009	0.633
Jar type 2	0.303	0.136	0.026	0.303	0.105	0.004
Jar type 3	0.115	0.136	0.397	0.115	0.095	0.226
Jar type 4	0.246	0.137	0.073	0.246	0.151	0.104
Jar type 5	0.361	0.136	0.008	0.361	0.143	0.011
_cons	2.926	0.300	0.000	2.926	0.300	0.000
Wald $\chi^2$	9.39			18.08		
Prob > $\chi^2$	0.094			0.003		
Log-likelihood	-452.642			-452.642		
Number of observations	310			310		
Left-censored observations	14			14		
Uncensored observations	296			296		
Right-censored observations	0			0		

*Tests for behavioral factors*

*Hypothesis 3.1: the asymmetric inconsistent preference effect.* We test the hypothesis that the proportion of affirmative inconsistency is greater than the proportion of negative inconsistency. Of the 480 times when WTP estimated from the posted price setting was higher than WTP estimated from the auctions, affirmative inconsistency occurred 89 times. Of the 95 times when WTP under posted prices was lower than under auctions, negative inconsistency happened 17 times. A proportion test of equality does not reject the null hypothesis that the two proportions are equal (a p-value of 0.88). Thus, the proportion of affirmative inconsistency is not significantly greater than the proportion of negative inconsistency. Therefore, this hypothesis is rejected and asymmetric inconsistent preferences should not be driving the differences in WTP.

*Hypothesis 3.2: a lack of familiarity with auction settings.* Participants’ institutional information might be affected by their lack of familiarity with auction formats. We test if roundnumber (the number of bidding decisions a participant has made) has an effect on the bids by specifically testing whether  $\beta_{\text{Auction, RoundNumber}} = 0$  holds. In order to gain more insights from the data, the regression in this part involves auction bids for all of the honey products (local, USA, and international). We test this hypothesis with a Tobit model, adding a set of experimental controls. The experiment controls include three information treatments, origin-information interactions, survey variables

**Table 7.** The effect of round number

	WTP bid amount—random-effects Tobit			Likelihood of zero WTP— random-effects Logit
	All bidders	On-margin bidders	Off-margin bidders	All bidders
RoundNumber	−0.0393***	0.0237**	−0.0776***	0.197***
Experimental controls	X	X	X	X
On-margin bidder	X	X		X
Off-margin bidder	X		X	X
Jar type 2	0.421***	0.411***	0.441***	−1.243***
Jar type 3	0.180*	−0.0424	0.130	−0.642
Jar type 4	0.328***	0.202*	0.299***	−0.797**
Jar type 5	0.510***	0.537***	0.406***	−1.228***
_cons	2.004***	1.658***	2.257***	−6.194***
Wald $\chi^2$	942.25	772.78	650.34	120.57
Prob > $\chi^2$	0.000	0.000	0.000	0.000
Log-likelihood	−2,812.153	−1,165.610	−1,254.617	−303.712
Number of observations	1,725	773	952	1,725

Notes: \*\*\*, \*\*, and \* represent significance at the 1 percent, 5 percent, and 10 percent levels, respectively. Estimates include subject random effects. Experimental controls include several order effects and information treatments, and details can be found in Wu *et al.* (2015).

on consumer attitude toward honey, and other socio-demographic variables.<sup>7</sup> As shown in Table 7 column 1,  $\beta_{\text{Auction, RoundNumber}}$  is significantly different from zero with a coefficient estimate of  $-0.039$ . Besides, in a logit model examining the probability that a participant submits a zero bid, we find that a zero bid is more likely to appear as the auction progresses (as shown in column 4, Table 7). Thus, the null hypothesis that  $\beta_{\text{Auction, RoundNumber}} = 0$  is rejected. As the auction rounds progress, participants tend to adjust their behavior based on information gathered through the process.

The underlying reason for the change of WTP in the auctions in successive rounds is not obvious, especially since there was no feedback regarding the price and winners. Meanwhile, it is possible that some participants lost interest and stopped bidding. Thus, we considered if off-margin and on-margin bidders behaved differently. Given the size of the bids, it is reasonable to define “on-margin” bidders as those whose bids are less than \$1 below the second highest bid and the rest as “off-margin” bidders:<sup>8</sup>

On-margin: bid > second highest bid − \$1;  
Off-margin: bid ≤ second highest bid − \$1.

<sup>7</sup>For a detailed list of the control variables used, see supplemental materials or reference Wu *et al.* (2015). That paper is a valuation study that looks at consumer valuation for various honey characteristics.

<sup>8</sup>This research followed Shogren *et al.* (2001) by setting the cutoff at \$1. In that study, approximately 55–69 percent of the bidders were categorized as off-margin. At the \$1.00 cutoff, approximately 55% of the bidders in this study were categorized as off-margin.

Column 2 and 3 in [Table 7](#) show Tobit regression results for on-margin and off-margin bidders, respectively, with experimental controls. The results are significant and robust to inclusions of demographic and attitude controls. In sum, bids by on-margin bidders increase \$0.024 each round and bids by off-margin bidders decrease \$0.078 each round. Therefore, on-margin bidders seem to show a gradually increasing pattern in their revealed WTP. Combined with the result that WTP elicited from the posted price section is higher than auctions and participants are generally more familiar with posted prices, this result adds some evidence to the DPH. With the more engaged responses, participants are gradually advancing to behavioral responses that are more systematic as the decision-maker gains additional information through familiarization and feedback. Even though the gap still exists between posted price and auction even with this gradual increase, it shows that interested participants are realizing and adjusting their bids to decrease this gap as the auction progresses. Note that even though we did 15 rounds of the auction, even by the end, the auction mechanism is still a relatively unfamiliar task. Therefore, even though the increase in bids of on-margin bidders may not explain the entire discrepancy, it could be a plausible explanation for this discrepancy. This reiterates the importance to have appropriate and sufficient training before an auction to familiarize participants with the mechanism. For off-margin bidders, the result may be of similar reason as pointed out by Knetsch et al. (2001) and Shogren et al. (2001)—second price auctions may not work very well for off-margin bidders as they may lose interest if they know they cannot win. In our English auction, it is reflected as shorter waiting times, and hence, decreasing bids.<sup>9</sup>

#### *Hypothesis 4: marginal effects between two methods*

In this part, we compare the marginal effect estimations in the two elicitation mechanisms. Since the auction bids implement a random-effects Tobit model and the posted price binary responses use a random-effects Logit model, the magnitudes of the variables are not directly comparable. However, the signs and significance levels of the attribute coefficients should be comparable. As demonstrated in [Table 8](#), we examine the sign estimates and significance levels for the jar attributes in the posted price and auction parts. A positive sign in the coefficient would indicate a WTP premium for that attribute and a negative sign indicates the opposite. Significance levels indicate the ability to detect a preference for attributes.

The first three columns show results for auction bids. Compared with the baseline Jar1, participants are willing to pay more for honey packaged in Jar2, Jar4, and Jar5. As shown in the last three columns for the posted price part, with the same number of observations, we can only demonstrate that the participants are significantly willing to pay more for Jar2, while the rest of the jar attributes are insignificant. However, if we focus on the sign estimates in the posted price part, the result suggests that the participants are willing to pay more for Jar4 and Jar5 than for Jar1, which is consistent with the auction. The coefficient is negative but insignificant for Jar2, which is also insignificant in the auction part. Therefore, even though we obtain less significance in the posted price part, the sign estimates mostly agree with the auction part. The above analysis suggests that posted price and auction generate similar qualitative

<sup>9</sup>It is possible that this behavior of off-margin bidders led to a decrease in the overall WTP elicited from the auction, even though previous studies have empirically shown that this auction format was able to retain the dominant strategy better than an ordinary English format. It is worth using other auction formats to test for the same effects in future studies.

**Table 8.** Marginal effect estimation comparison in posted price and auction

	Auction			Posted price		
	Coefficient	Sd. Err.	p> z	Coefficient	St. Err.	p> z
Jar type 2	0.374	0.092	0.000	0.710	0.313	0.023
Jar type 3	0.043	0.092	0.639	-0.054	0.329	0.869
Jar type 4	0.183	0.092	0.048	0.351	0.318	0.270
Jar type 5	0.355	0.092	0.000	0.303	0.319	0.342
_cons	2.643	0.198	0.000	-1.311	0.253	0.000
Wald $\chi^2$	28.06			7.79		
Prob > $\chi^2$	0.000			0.099		
Log-likelihood	-786.041			-335.720		
Number of observations	575			575		
Left-censored observations	45			0		
Uncensored observations	530			575		
Right-censored observations	0			0		

marginal effects for attributes, but auctions are more efficient in revealing these underlying preferences.

## Conclusion

Experimental auctions are a popular instrument for measuring consumer WTP for various attributes of a commodity or environmental service. A key attractive feature of auction mechanisms is that they provide point estimates of WTP. However, posted price formats are how most consumer choices are made. Inferring consumers' WTP for a posted price market from auction bids can be problematic, since consumers generally may have relatively limited experience with auctions and may not behave in a consistent manner in both mechanisms. Therefore, some attention has been paid to comparing the estimated mean WTPs using these two mechanisms. On the other hand, the comparison of other important aspects, such as the estimation of how WTP varies with certain product attributes, has not been thoroughly examined in the literature. In this research, we test the mean WTP differs in the two elicitation methods and further offer explanations of such a discrepancy using an artefactual field experiment. Moreover, we compare the signs and significance levels of marginal effects for different product characteristics.

First, in our auction, the estimates of WTP from bids are significantly less than estimates of WTP for the same product via the posted price mechanism. We conducted both within-subjects and between-subjects tests, and the results are robust. We test several potential explanations related to information and framing effects. The differences in WTP do not appear to be due to either an anchoring effect or asymmetric inconsistent preferences. The results do suggest that the reason for the difference in auctions is research participants' lack of familiarity with auctions. Second, we run regressions to test the marginal effects of different product attributes

on WTP. The signs of coefficients are consistent in the auction and posted price mechanisms. Third, we find that the significance level is much higher using auctions for each confident.

Our research sheds light on which economic evaluation elicitation format, namely auctions and posted price mechanisms, is more suitable under different circumstances. We show that a WTP estimate difference does exist between the two mechanisms. Participants do demonstrate an adaption process in the auction format. Meanwhile, the posted price mechanism is more familiar with the general public and participants may focus more on the task itself. This is particularly true in a field setting where researchers usually recruit participants from busy marketplaces, where attention and time allocated to experiments are generally limited. However, we show that both methods elicit similar signs for the marginal effects of specific product attributes. Thus, using either auctions or posted price mechanisms can provide credible prediction on the marginal effects of important product characteristics. But auctions have a clear advantage over posted price in terms of statistical power, which indicates that a larger sample size is required for a posted price mechanism to reveal the preference for specific product attributes. Therefore, it is reasonable to consider using posted price when one's goal is to understand absolute WTP values and to use auctions when one is interested in relative WTP comparisons associated with different attributes.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/age.2021.6>.

**Funding statement.** This work was supported by the Northeast Center for Risk Management Education (grant no. RME-M7604495).

**Competing interest.** None declared.

**Data availability statement.** Replication materials are available in the supplementary materials section of the article.

## References

- Ariely, D., G. Loewenstein, and D. Prelec. 2003. "Coherent Arbitrariness: Stable Demand Curves without Stable Preferences." *The Quarterly Journal of Economics* 118(1): 73–106.
- Arrow, K., R. Solow, P.R. Portney, E.E. Leamer, R. Radner, and H. Schuman. 1993. "Report of the NOAA panel on contingent valuation." *Federal Register* 58(10): 4601–4614.
- Banerji, A., S. Chowdhury, H. Groote, J.V. Meenakshi, J. Halegoah, and M. Ewool. 2018. "Eliciting Willingness to Pay through Multiple Experimental Procedures: Evidence from Lab in the Field in Rural Ghana." *Canadian Journal of Agricultural Economics/Revue canadienne d'agroéconomie* 66(2): 231–254.
- Becker, G.M., M.H. DeGroot, and J. Marschak. 1964. "Measuring Utility by a Single-response Sequential Method." *Behavioral Science* 9(3): 226–232.
- Bernard, J.C. 2006. "Finding and Retaining the Dominant Strategy: The Second-Price, English and 'Sealed Offer' English Auctions." *Journal of Economic Behavior and Organization* 60(4): 553–561.
- Berry, J., G. Fischer, and R. Guiteras. 2020. "Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana." *Journal of Political Economy* 128(4): 1436–1473.
- 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.
- Corrigan, J.R., D.P.T. Depositario, R.M. Nayga Jr, X. Wu, and T.P. Laude. 2009. "Comparing Open-Ended Choice Experiments and Experimental Auctions: An Application to Golden Rice." *American Journal of Agricultural Economics* 91(3): 837–853.
- Corrigan, J.R., A.C. Drichoutis, J.L. Lusk, R.M. Nayga, and M.C. Rousu. 2011. "Repeated Rounds with Price Feedback in Experimental Auction Valuation: An Adversarial Collaboration." *American Journal of Agricultural Economics* 94(1): 97–115.

- Couch, A., and K. Keniston.** 1960. "Yeasayers and Naysayers: Agreeing Response Set as a Personality Variable." *The Journal of Abnormal and Social Psychology* 60(2): 151.
- Dillaway, R., K.D. Messer, J.C. Bernard, and H.M. Kaiser.** 2011. "Do Consumer Responses to Media Food Safety Information Last?" *Applied Economic Perspectives and Policy* 33(3): 363–383.
- Frykblom, P., and J.F. Shogren.** 2000. "An Experimental Testing of Anchoring Effects in Discrete Choice Questions." *Environmental and Resource Economics* 16(3): 329–341.
- Gracia, A., M.L. Loureiro, and R.M. Nayga.** 2011. "Are Valuations from Nonhypothetical Choice Experiments Different from those of Experimental Auctions?" *American Journal of Agricultural Economics* 93(5): 1358–1373.
- Green, D., K.E. Jacowitz, D. Kahneman, and D. McFadden.** 1998. "Referendum Contingent Valuation, Anchoring, and Willingness to Pay for Public Goods." *Resource and Energy Economics* 20(2): 85–116.
- Harrison, G.W., R.M. Harstad, and E.E. Rutström.** 2004. "Experimental Methods and Elicitation of Values." *Experimental Economics* 7(2): 123–140.
- Herriges, J.A., and J.F. Shogren.** 1996. "Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning." *Journal of Environmental Economics and Management* 30(1): 112–131.
- Kanninen, B.J.** 1995. "Bias in Discrete Response Contingent Valuation." *Journal of Environmental Economics and Management* 28(1): 114–125.
- Kimanta, J.** 1986. *Elements of Econometrics*. Ann Arbor, MI: University of Michigan Press.
- Knetsch, J.L., F.F. Tang, and R.H. Thaler.** 2001. "The Endowment Effect and Repeated Market Trials: Is the Vickrey Auction Demand Revealing?" *Experimental Economics* 4(3): 257–269.
- Krström, B.** 1993. "Comparing Continuous and Discrete Contingent Valuation Questions." *Environmental and Resource Economics* 3(1): 63–71.
- Li, T., M. Kecinski, and K.D. Messer.** 2017. "Heterogeneous Responses and Demographic Differences for Oyster Consumption: Evidence from Field Experiments." *Agricultural and Resource Economics Review* 46(2): 296–314.
- Li, T., and J.J. McCluskey.** 2017. "Consumer Preferences for Second-Generation Bioethanol." *Energy Economics* 61: 1–7.
- List, J.A.** 2003. "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics* 118(1): 41–72.
- List, J.A., and C.A. Gallet.** 2001. "What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics* 20(3): 241–254.
- List, J.A., and D. Lucking-Reiley.** 2000. "Demand Reduction in Multiunit Auctions: Evidence from a SportsCard Field Experiment." *The American Economic Review* 90(4): 961–972.
- 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): 109–123.
- Loureiro, M.L., J.J. McCluskey, and R.C. Mittelhammer.** 2002. "Will Consumers Pay a Premium for Eco-Labeled Apples?" *Journal of Consumer Affairs* 36(2): 203–219.
- Lucking-Reiley, D.** 1999. "Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet." *American Economic Review* 89(5): 1063–1080.
- Lusk, J.L., T. Feldkamp, and T.C. Schroeder.** 2004. "Experimental Auction Procedure: Impact on Valuation of Quality Differentiated Goods." *American Journal of Agricultural Economics* 86(2): 389–405.
- Lusk, J.L., and T.C. Schroeder.** 2006. "Auction Bids and Shopping Choices." *Advances in Economic Analysis & Policy* 6(1): 4.
- Lusk, J.L., and J.F. Shogren.** 2007. *Experimental Auctions: Methods and Applications in Economic and Marketing Research*. Cambridge, UK: Cambridge University Press.
- McNemar, Q.** 1947. "Note on the Sampling Error of the Difference between Correlated Proportions Or Percentages." *Psychometrika* 12(2): 153–157.
- Messer, K.D., G.L. Poe, D. Rondeau, W.D. Schulze, and C.A. Vossler.** 2010. "Social Preferences and Voting: An Exploration using a Novel Preference Revealing Mechanism." *Journal of Public Economics* 94(3): 308–317.
- Plott, C.R.** 1996. "Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis." In: K.J. Arrow, E. Colombatto, M. Perlman, C. Schmidt (eds), *The rational foundations of economic behavior*. IEA Conference Vol. 114 (pp. 225–250). London: McMillian.

- Plott, C.R., and K. Zeiler.** 2005. "The Willingness to Pay-Willingness to Accept Gap, the 'Endowment Effect, 'Subject Misconceptions, and Experimental Procedures for Eliciting Valuations." *American Economic Review* 95(3): 530–545.
- Poe, G.L., K.L. Giraud, and J.B. Loomis.** 2005. "Computational Methods for Measuring the Difference of Empirical Distributions." *American Journal of Agricultural Economics* 87(2): 353–365.
- Ready, R.C., J.C. Buzby, and D. Hu.** 1996. "Differences between Continuous and Discrete Contingent Value Estimates." *Land Economics* 72(3): 397–411.
- Roosen, J., S. Marette, and S. Blanchemanche.** 2010. "Value Elicitation using BDM and a Discrete Choice Mechanism." *Economics Bulletin* 30(2): 1554–1563.
- Rutström, E.E.** 1998. "Home-Grown Values and Incentive Compatible Auction Design." *International Journal of Game Theory* 27(3): 427–441.
- Shogren, J.F., M. Margolis, C. Koo, and J.A. List.** 2001. "A Random Nth-Price Auction." *Journal of Economic Behavior and Organization* 46(4): 409–421.
- Su, L., B.D. Adam, J.L. Lusk, and F. Arthur.** 2011. "A Comparison of Auction and Choice Experiment: An Application to Consumer Willingness to Pay for Rice with Improved Storage Management." *Agricultural and Applied Economics Association (AAEA) 2011 Annual Meeting* (No. 321-2016-10908).
- Tversky, A., and D. Kahneman.** 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124–1131.
- Venkatachalam, L.** 2004. "The Contingent Valuation Method: A Review." *Environmental Impact Assessment Review* 24(1): 89–124.
- Vickrey, W.** 1961. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *The Journal of Finance* 16(1): 8–37.
- Wu, S., J. Fooks, K.D. Messer, and D. Delaney.** 2015. "Consumer Demand for Local Honey." *Applied Economics* 47(14): 4377–4394.
- Yoo, S., and H. Yang.** 2001. "Application of Sample Selection Model to Double-Bounded Dichotomous Choice Contingent Valuation Studies." *Environmental and Resource Economics* 20(2): 147–163.