Design Factors influencing Willingness to pay estimates in the Becker-Degroot-Marshark (BDM) mechanism and the non-hypothetical Choice Experiment: A case of Biofortified maize in Zambia

Priscilla Hamukwala, Adewale Oparinde, Hans Peter Binswanger-Mkhize and Johann Kirsten[1]

[Original submitted May 2017, Revision received October 2017, Accepted January 2018]

Abstract

Two of the experimental methods used to estimate willingness to pay (WTP) for a non-market good, the Becker-DeGroot-Marschak (BDM) mechanism and the non-hypothetical choice experiment (nHCE) often lead to significantly different WTP estimates, complicating the choice between the methods. In Zambia the same group of researchers used both techniques to evaluate WTP for orange maize, which provides more vitamin A than other varieties. This provided an opportunity to analyze the sources of the difference. In the BDM experiment, one group of respondents was provided with more training opportunities than the other, and made higher bids. Accounting for lexicographic behavior in the nHCE reduced the estimated WTP. These two design factors together resulted in a decrease in the WTP difference for orange maize (1279-632 ZMK) although the difference remains statistically significant. More training was also shown to eliminate the effects of different orders in which maize varieties were presented.

Key words: Valuation techniques, Becker-DeGroot-Marschak, non-hypothetical choice experiment, Willingness-to-pay, Design factors, Biofortified foods, Maize, Zambia

JEL Classifications: C35, C93, D12, Q13

[1] Priscilla Hamukwala (contact author: phamukwala@unza.zm) is in the Department of Agricultural Economics, Extension and Rural Development, University of Pretoria, South Africa, as was Hans Peter Binswanger-Mkhize prior to his death in 2017. Adewale Oparinde is with International Food Policy Research Institute, Harvest Plus, Washington, USA. Johann Kirsten is with the Bureau for Economic Research at Stellenbosch University, South Africa. The authors would like to thank the Editor in chief David Harvey, and the two anonymous reviewers for their elaborate and helpful comments and suggestions that have led to an improved manuscript. The views expressed in this paper are solely those of the authors.

Introduction

We compare the performance of the Becker-DeGroot-Marschak (BDM) mechanism and the non-hypothetical choice experiment (nHCE) under a field setting in Zambia. We use data from a study of rural Zambian consumers' valuation of a new maize variety that has been enriched with vitamin A through biofortification[2]. By focusing on maize - a staple food crop - we expect that dietary habits will not be affected, thereby reducing the chances of non-adoption, and ultimately increasing chances of alleviating the micronutrient deficiency problem. However, the biofortification process may change some product intrinsic attributes. For example, an increase in vitamin A gives maize the unfamiliar orange colour when the most preferred maize in Zambia is white (Meenakshi, et al., 2012). With historical evidence of yellow maize rejection in Zambia and in the region, it leaves a question of whether this problem also applies to biofortified maize which is likely to be orange.

Nonmarket valuation techniques are required when there is no market data. These include stated preference techniques and incentive compatible economic experiments. Stated preference techniques use consumers' WTP expressed for a hypothetical good or service in which their statements or choices are non-binding. There is a general consensus among researchers that these methods often lead to hypothetical bias. These tools nevertheless remain useful for estimating demand for new products or public goods for which revealed preference data do not exist. Researchers have proposed various methods of reducing hypothetical bias, such as the use of "cheap talk" scripts (Cummings & Taylor, 1999) or combining hypothetical with non-hypothetical experiments (Chowdhury, et al., 2011; Lusk & Shogren, 2007). Examples of stated preference techniques include; conjoint analysis, stated choice experiments and contingent valuation methods.

Incentive compatible economic experiments, on the other hand, use a hypothetical market with subjects facing real budget constraints and products, thereby providing respondents an incentive to reveal their true preferences. Since preferences are known to be revealed in an actual payment setting, they could be revealed in these experiments even when the market is hypothetical (Zawojska & Czajkowski, 2015). The BDM and the nHCE are common examples of such methods, and are becoming more common in the developing world. Some of the studies that have

[2] A process in which micronutrient contents of crops are increased through biotechnology or conventional plant breeding (De Groote, et al., 2014)

used the BDM to study acceptance of new agricultural technologies in Africa include willingness to pay for fortified yellow maize in Kenya (Morawetz, et al., 2011); willingness to pay for pro-vitamin A orange maize in Mozambique (Stevens & Winter-Nelson, 2008); consumer acceptance of pro-vitamin A orange cassava in Nigeria (Oparinde, et al., 2014); consumer acceptance of bean varieties biofortified with iron in Rwanda (Oparinde, et al., 2015). The non-hypothetical choice experiment studies examples include willingness to pay for orange maize in Zambia (Meenakshi, et al., 2012); and willingness to pay for orange fleshed sweet potatoes in Uganda (Chowdhury, et al., 2011).

Although incentive compatible economic experiments are often considered superior to stated preference methods, in practice these experiments can still mask several background variables that may influence participants' WTP (Voelckner, 2006). Experimental literature has attributed this in part to experimental design contexts and information cues. For example, Berry et al. (2011) noted that if an individual's response to the WTP question is believed to affect the actual price of the product, they have an incentive to respond strategically. They found, using a BDM experiment, that consumer's WTP estimates for water filters were lower if respondents knew that their bid would influence its future pricing.

Further, there are some theoretical debates of whether these methods are truly incentive compatible. For example, it has been argued that the BDM may not always be incentive compatible beyond the expected utility context. Karni & Safra (1987) demonstrate that the BDM incentives do not fully explain consumer preferences over lotteries (or when unsure of the value of the good) for individuals who are not maximising their expected utility. Horowitz (2006) demonstrates that even when individuals are sure of the good's value, the BDM is still not incentive compatible if they are not maximizing their expected utility. Carson & Plott (2014) also argue that the BDM is not consistent empirically and attribute this to its complexity such that some inexperienced subjects may fail to recognize its incentive structure.

The nHCE is relatively simpler and more demand revealing, given that consumer choices are similar to those made in a real market (Louviere & Woodworth, 1983). However, it has not been without criticisms. One of the concerns is the possible violation of the independence axiom of the expected utility theory. Holt (1986) indicated that this usually happens when there are several choice-problems, and real incentives involve some random selection of only one of them as binding. The author demonstrated that a choice problem presented with other choice problems may not yield the same response as it would if it were the only problem being faced. This implies that incentive compatibility does not hold since a decision in one choice problem may be influenced by other choice scenarios within the experiment.

These theoretical concerns are further supported by divergent empirical WTP estimates in the studies that have compared these methods when addressing the same research question in similar contexts, e.g., Banerji et al. (2013); Gracia, et al. (2011) and Lusk & Schroeder (2006). Theoretically, incentive compatible methods are expected to yield similar results but the nHCE in these studies exhibited consistently higher WTP estimates over the BDM. Market researchers have not reached a consensus as to which method gives better estimates. Voelckner (2006) noted that there is no simple answer to this question because consumer's true WTP is latent, implying that each valuation technique only represents the attempt to come close to the true WTP. Thus, observed estimates being similar to theoretical predictions or valuation techniques yielding comparable estimates, provides some assurance that the WTP estimates are valid and reflect the true market demand. However, if the WTP estimates significantly differ by elicitation methods, systematic differences can be observed and sources of differences can be identified (Banerji, et al., 2013).

Given these concerns, we explore the effects of four of the design issues discussed in literature affecting the BDM and the nHCE in the acceptance of biofortified maize in Zambia. First, we look at whether subjects exhibit lexicographic preferences in the nHCE, which could potentially affect WTP. In a choice experiment with repeated choices, individuals are expected to consider all the attributes and trade off at each attribute level. However, some individuals may exhibit lexicographic behavior or consistently choosing the same option across choice sets, thereby biasing WTP estimates. This behaviour could be due to complicated choice-tasks or simply a reflection of the individual's strong preference of an attribute that is not altered within the range of prices offered in the experimental rounds (Killi, et al., 2007).

Second, we investigate how use of repeated auctions prior the main BDM experiment may affect subjects' WTP. According to the "Preference learning hypothesis" (Plott, 1996), preferences are learned through experience such that the bidding behaviour in the short-run may be random. With successive bidding rounds however, preferences are discovered and bids become more rational. Thirdly, we are interested in whether experimenters influence the subject's bidding behavior. Literature attributes this to the subject's unfamiliarity of the experimental procedures that may require translating of instructions into local language or rephrasing (Whitting, 2002; Morawetz, et al., 2011).

The fourth issue is whether order-effects influence subject's bidding behaviour, which may result from inadequate learning of the experimental procedure or participant fatigue (Lusk & Shogren, 2007; Morawetz, et al., 2011). Also, in sensory testing, O'Mahony, (1986) finds that the order in which one taste samples may affect the sensory scores. He attributes this to the inability of the subject to detect the differences between stimuli or "adaptation". This comes from reduced sensory acuity as one gets used to the taste of the product, especially when tasting a strong sample before a weak one. He further notes that there's almost no effect when the stimuli have different taste qualities.

On this basis, we aim to determine whether WTP estimates elicited from the BDM experiments and the nHCE are equal. If not, we examine how design effects associated with each method affect the WTP disparities. In particular, we investigate the extent to which design factors (ordering effects, experimenter's effect, use of repeated auctions and lexicographic behaviour) affect the difference between the two methods.

We make two main contributions to literature. First, we compare the performance of two incentive compatible valuation techniques in a development country setting. Second, we explore experimental design features that can potentially bias the estimates of the BDM and the nHCE and possibly explain the disparity in empirical estimates of the two valuation techniques. We expect that the inclusion of these design features will narrow the WTP gap between the two mechanisms, thereby resulting in WTP estimates that are more comparable.

The paper is organized as follows: In the following section we describe the methodology including the features of the nHCE, the BDM experiment and the estimation procedure. This is followed by the results and discussion, and lastly the conclusion with the implications of our findings.

Data Description

The data from the BDM and the nHCE experiments are part of the larger survey designed to determine consumer's acceptance of vitamin A biofortified-maize in Zambia reported in Meenakshi et al. (2012). Before the experiments, the participants evaluated the sensory attributes (appearance, aroma, taste, texture and overall liking) of *nsima* made from three maize varieties: biofortified maize; conventional white maize; yellow maize (which is considered inferior in the market) using likert scales of 1 to 5 (1=dislike very much, 5= like very much). This was done to familiarise participants with all the maize varieties, thereby ensuring that consumer's choices and bids were fully informed. In both experiments, no nutrition information was provided about the biofortified maize variety. The BDM experiment consisted of the biofortified maize grain only, while the nHCE consisted of all the three maize varieties. In the rest of the study, we refer to the biofortified maize as orange maize, conventional white maize as white maize, and the inferior yellow maize as yellow maize.

The nHCE was labelled with maize variety kernel colour (orange, yellow and white) at varying price levels. The colour encompassed all sensory attributes of each of the three maize varieties. Sixteen choice sets were generated by fractional factorial design using a statistical package for social sciences (SPSS) software, subject to orthogonality and balance level properties. An opt-out option was included in each choice sets (see Table1). Subjects were asked the option they preferred in each scenario involving the purchase of 2.5kg grain of each of the three maize varieties.

The BDM experiment had two treatment arms in which participants were either subjected to one or ten bidding-rounds prior the main BDM bid. They were then asked to state the maximum amount they would be WTP for a 2.5kg orange maize grain in each of the rounds. In both experiments the prices were varied depicting 30-50% discounts and premiums of the median price of the conventional white maize varieties that prevailed in the study area (i.e., posted prices for the nHCE and random drawn price distribution for the BDM).

Respondents in both experiments were given a participation fee of 2000 ZMK each (equivalent to 50 US cents) to use in their purchasing tasks for them not to run out of pocket money (which they could also keep if they desired not to make any purchases). Once the auctions and choices were over, one of the choice scenarios in the choice experiment and one of the bidding rounds were randomly drawn and executed (i.e., subjects were rewarded according to the decision each had made on the randomly selected bid or choice). In the BDM experiment, this meant making a purchase of 2.5kg of orange maize grain if the randomly selected BDM round is where the respondent's submitted bid was higher than the randomly drawn price, or not making a purchase, if the converse was true.

Similarly, for the choice experiment, if the randomly drawn scenario is where the respondent chose a given variety, they would purchase a 2.5kg grain of that variety at its given price. If on the other hand they chose the opt-out option, they would not make a purchase. This therefore made both experiments incentive compatible as it was in the individual's best interest to make a truthful bid or choice in each round to avoid missing buying a product they real wanted or buying it when they really did not want it.

Table 1 Choice sets for the Choice Experiment at varying price levels (in Zambian Kwacha)

Choice set	White-Maize	Yellow-Maize	Orange-Maize	Opt-out option
1	1200	1200	2000	None of these
2	1200	800	1200	None of these
-	-	-	-	-
-	-	-	-	-
16	1200	1500	1500	None of these

Data was collected from areas of Zambia that were likely to benefit more from vitamin A biofortified maize, thus two districts, each of Southern and Central provinces were purposively selected using the Zambian census 2010 data. As indicated by Meenakshi et al (2012), these were districts with highest levels of maize production and consumption, as well as poverty. The experiments were conducted at a central location in each district where participants were recruited from nearby households and or/ villages and assigned randomly to each experiment as they came. Relative to the nHCE, the BDM experiment was assigned more participants because of the repeated bidding treatment. This was particularly important to maintain comparable sample size to the nHCE in case some participants refuse to participate in repeated bidding. Fortunately, this was not the case. There were no significant differences of individual responses under both experiments by region, hence regional data were merged. The final sample used in the analysis after data cleaning from the nHCE is 107 respondents, while that for the BDM is 145.

Table 2 shows summary statistics under both experiments, indicating that individuals were similar in most characteristics except in age and land area cultivated. Participants in the nHCE were about 3 years older and owned 2 hectares less land than those in the BDM ($p<0.01$). Most participants were married and male in both groups and had on average slightly less than 9 years of formal education. The average household size was 9 members while 8 members shared the same meal. Common household owned assets include livestock, radio, farm implements and cell phones and from these we constructed an asset index using principal component analysis following Filmer and Prichett, (2001).

Table 2 Summary Statistics of Participants (Mean)

	nHCE	**BDM**	**Difference**	**(P-value)**
Age (in years)	42.44	38.57	3.87	0.02
Gender (1=male, zero otherwise)	0.63	0.57	0.05	0.39
Marital status (1=married, zero otherwise)	0.82	0.87	-0.05	0.31
Education (in years)	8.37	8.25	0.12	0.73
Household size	8.59	8.69	-0.10	0.87
No. of people sharing meals	7.68	8.19	-0.51	0.27
Land area cultivated (ha)	2.85	4.96	-2.11	0.01
Income (US$ from 3 major sources)	933.52	1470.57	462.95	0.28
Asset-index	0.59	0.53	0.06	0.13

Model Specifications

Choice Experiment

The basis for modelling the choice experiment is Lancaster, (1966), where consumers derive utility (U_{ij}) of products from their attributes (V_{ij}), and from the random utility theory (McFadden, 1974) which divides utility into a systematic and a random component as follows:

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

Where;

U_{ij} is the utility of individual i for choice j

V_{ij} is the explanatory part of the utility function

ε_{ij} is the random component of utility for choice j

Given the latent nature of utility, consumer's decisions are analyzed from the probabilistic theory, with the probability that consumer i will choose option j given by;

$$\pi_{ij} = \Pr\{Y_i = j\} = \Pr\{\max(U_{i1,} U_{i2,} \dots, U_{iJ}) = U_{ij}\} \tag{2}$$

Assumptions of the error term distribution leads to different models. The conditional logit model is appropriate if the error term is independently and identically distributed (i.i.d) across individuals and alternatives (McFadden, 1974). Thus, the probability of individual i choosing option j in choice t is:

$$P_{ijt} = \frac{e^{v_{ijt}}}{\sum_{k=1}^{j} e^{v_{ijt}}} = \frac{e^{X_{ijt}\beta}}{\sum_{k=1}^{j} e^{X_{ijt}\beta}} \tag{3}$$

Equation (3) implies that all individuals have the same preferences ($\beta_i=\beta_{k,}$ $\forall$ k). The mixed logit is an alternative model which relaxes this assumption by allowing for random parameters $d\ (\beta)$ with a distribution of $f(\beta)$, implying that parameters are different across individuals (Lusk & Schroeder, 2006). Thus, the probability of individual i choosing alternative j in choice t is;

$$P_{ijt} = \int \frac{e^{v_{ijt}}}{\sum_{k=1}^{j} e^{v_{ijt}}} f(\beta) d\ (\beta) \tag{4}$$

Empirically, V_{ij} is a function of product attributes and demographics (equation 5). The final variables included in the model, subject to multi-collinearity specification tests, are as shown in equation (5).

$$V_{ijt} = \beta_{1jt} + \beta_2 Price_{jt} + \beta_{3j} * Gender_i + \beta_{4j} * Education_i + \beta_5\ Age_i * price_{jt} +$$
$$\beta_6\ Assets_i * price_{jt} + \beta_{7j} * Age_i + \beta_{8j} * Assets_i \tag{5}$$

Where:

V_{ijt} is consumer i's choice of maize variety j in choice set t

$\beta_{1jt:}$ is the alternative specific constant or dummy variable for a maize variety j (attributes of each maize variety not captured by the model) which is compared to the opt-out option.

$Price_{jt}$: is the price of the maize variety j in choice set t

$Gender_{i:}$ is the gender of the i^th^ consumer choosing variety j

$Education_{i:}$ is the number of years of formal education for the i^th^ consumer choosing variety j

$Assets_{i:}$ is the asset index for the i^th^ consumer choosing a maize variety j

BDM Experiment

The BDM (Becker , et al., 1964) is a valuation technique in which subjects are asked to indicate the highest price they would be WTP for a fixed quantity of a good which is then compared to a randomly drawn selling price. It is not an auction *per se* since an individual makes decisions independent of other participants by bidding against a random price instead of other individuals as is the case in a typical auction. It is however, theoretically equivalent to a second price auction in which there are at least two bidders, with the highest bidder winning and paying the price of the second highest bidder (Lusk, et al., 2007).

To explain respondents' optimal bidding behavior in the BDM, the analysis derived by Lusk & Shogren, (2007) is followed. Let *Vi* represent the value that subject *i* places on a good. An individual purchases the good if the submitted bid is higher than or equal to the randomly drawn price. The utility (U_i) that individual *i* derives from their maximum bid is from the difference between the randomly drawn price (P) and the value (V) they place on the good (equation 6):

$$U_i = (V_i - P) \tag{6}$$

If the subject does not win by bidding lower than p, they receive and pay nothing, and their monetary value for the good is normalized to zero. Since the bidder does not know the winning price, their expected price can be assumed to be from a random distribution with a cumulative density function G(p) and a probability density function g(p). The dominant strategy for the bidder is to bid one's true value of a good, i.e., submitting a bid that will maximize their expected utility (equation 7):

$$E[U_i] = \int_{pi}^{bi} U_i\,(V_i - p)\; d\; G_i\,(p) + \int_{pi}^{bi} U_i\,(0) \;= \int_{pi}^{bi} U_i\,(V_i - p)\; d\; g_i\,(p)dp + \int_{pi}^{bi} U_i\,(0) \tag{7}$$

The first integral is taken over all price levels less than the bid (winning range) while the second is taken over levels greater than the bid (losing range). Normalizing U_i (0) = 0, the optimal bid is obtained by taking a derivative with respect to b_i and setting it equal to zero. The optimal bid (b^*_i) is one which is equal to the randomly drawn price (p) as shown below:

$$\frac{\partial E[U_i]}{\partial b_i} = U_i(v_i - p)f_i(b_i) = 0 \text{ when } b_i = V_i \tag{8}$$

The model choice depends on the data distribution and our data show censoring of 21 percent at 2000ZMK. Incidentally this is the same amount given as the participation fee (money given for taking part in the experiment which could also be used in their purchasing tasks). It is likely that participants were willing to submit bids higher than 2000ZMK but had only the participation fee at their disposal, or they saw the 2000ZMK as the maximum value of orange maize. This requires the use of a right-censored model, suggesting that true WTP for such respondents was at least 2000ZMK. The parametric Tobit model is used and compared to the censored least absolute deviation (CLAD) and symmetrically censored least squares (SCLS) estimators which are both semi-parametric (Powell, 1986; 1984). Each of these models is discussed below:

Tobit

The Tobit model (Tobin, (1958)) takes the linear form:

$$Y_i^* = X_i'\beta + \varepsilon_i \quad i = 1, \ldots, n \qquad \text{where;} \qquad (\varepsilon \sim N(0, \sigma^2)) \qquad (9)$$

Where;

Y_i^* Is the latent dependent variable;

X_i is a vector of independent variables;

ε_i is the error term.

With right censoring, the dependent variable Y_i^* is only observed when it is less than some scalar c_i (2000 in our case) as shown below:

$$Y_i^* = min\{Y_i^*, c_i\} = \min\{X_i'\beta + \varepsilon_i, c_i\} = \min\ \{X_i'\beta + \varepsilon_i, 2000\} \qquad (10)$$

Symmetrically censored least squares (SCLS)

The SCLS estimator (equation 11) proposed by Powell, (1986) is an alternative to the Tobit as it relaxes the homoscedastic assumption. It is based on the assumption of symmetrically and independently distributed error term with the true dependent variable (y^*) following the same distribution. When the observed part of the dependent variable (y_i) is asymmetric, it can be restored to symmetric through symmetrically censoring it. Estimation is done using least squares of only symmetrically trimmed data. The symmetric assumption is less restrictive than the parametric assumption thereby providing consistent estimates when parametric assumptions fail to hold. It is however stronger than the zero median assumption of the CLAD model.

$$\hat{\beta}_T = \left[\sum_{t=1}^{T} 1\left(X'_t\hat{\beta}_T > 0.X_tX'_t\right)\right]^{-1}.\sum_{t=1}^{T} 1\left(X'_t\hat{\beta}_T > 0\right).min\left(Y_t, 2X'_t\hat{\beta}_T\right).X_t \quad (11)$$

Where; β is the parameter to be estimated;

X is a vector of independent variables;

T is the sample size;

Y_t is the dependent variable.

Censored Least Absolute Deviation (CLAD) Estimator

The CLAD estimator is another alternative when parametric conditions fail. It uses median regression and since censoring only affects the mean and not the median (if < 50%), it is consistent when distribution assumptions are violated (Powell, 1984). The CLAD estimator assumes that $med(\mathcal{E}_i/X_i, c_i) = 0$. It is estimated as:

$$Y_i^* = med\{Y_i^*, c_i\} = \text{med}\ \{X_i'\beta + \mathcal{E}_i, c_i\} = \text{med}\ \{X_i'\beta + \mathcal{E}_i, 2000\} \quad (12)$$

Comparison of the Bids and Choice data

The estimates from the nHCE and the BDM experiments are not directly comparable. Respondents' direct bids from the BDM are directly interpreted as their WTP as below (Lusk & Shogren, 2007).

$$WTP^* = BID_i = \beta X_i + \mathcal{E}_i \quad (13)$$

Where:

WTP^* is individual*'s* willingness to pay;

BID_i is the individual's bid;

X_i is a vector of explanatory variables;

$\mathcal{E}_i$ is an error term.

In the nHCE, participants did not bid directly on how they valued each maize variety but chose one maize variety at different prices. Therefore, regression estimates do not directly reflect consumer's WTP. Instead, WTP per attribute is given by dividing the attribute parameter by the negative value of the parameter for the price attribute $\left(-\frac{\alpha_i}{\beta}\right)$ after running the regression (Lusk & Schroeder, 2006). Mean WTP for the product will be a sum of these which can also be obtained by determining the price of the maize variety j that will equate the systematic component (in equation 5) equal to zero, while holding individual characteristics at sample averages (Hole & Kolstad, 2012).

The computation of the mean WTP for maize variety j therefore is translated to equation (14) which is simply the sum of the parameter attributes (i.e., constant, gender, education, age and assets) evaluated at their sample means divided by the parameters of the negative price attribute (i.e., parameters of the price attribute include the price coefficient β_2; age interacted with price β_5 and assets interacted with assets β_6. The variables and coefficients are as defined in equation (5). Data analyses for both BDM and Choice data were done using Stata 12 software.

$$\overline{WTP_j} = \overline{Price_j} = -\left[\left(\frac{\beta_{1j}+\beta_{3j}*\overline{Gender_i}+\beta_{4j}*\overline{Education_i}+\beta_{7j}*\overline{Age_i}+\beta_{3j}*\overline{Assets_i}}{\beta_2+\beta_5\overline{Age}_i+\beta_6\overline{Assets_i}}\right)\right] \tag{14}$$

Results and Discussion

Factors influencing bidding behaviour in the BDM experiment

WTP estimates in the BDM experiment, using the Tobit, CLAD and SCLS models, are shown in Table 3. The response variable in each is the bid submitted to purchase 2.5kg of orange maize grain by each individual. A comparison of results across the models revealed that the parameter estimates were similar in signs except on the household size variable in the CLAD model. The SCLS model produced results similar to the Tobit model in the coefficient signs and significance levels (except gender). Gender is the only variable that consistently gave a positive and significant coefficient across the three specifications. It is therefore clear that men were willing to pay more for orange maize than women.

Table 3 WTP Estimates from the BDM Auction

	(Tobit)		**(CLAD)**		**(SCLS)**	
Bid	**β**	***Se***	**β**	***Se***	**β**	***Se***
Household size	2.95	(7.42)	-1.96	(13.38)	3.69	(4.37)
Age	10.88	(15.81)	3.93	(20.52)	5.38	(11.72)
Age-square	-0.16	(0.19)	-0.09	(0.21)	-0.091	(0.12)
Male	140.10*	(80.39)	145.22*	(122.23)	125.50**	(63.18)
Education	-11.53	(13.61)	-13.90	(18.78)	-10.06	(10.12)
Asset-index	8.75	(16.31)	20.78	(26.02)	4.31	(13.47)
Constant	1581.30***	(320.10)	1821.26***	(486.60)	1609.70***	(251.10)
Sigma constant	390.40***	(28.57)				
Normality test- $X^2(1)$	99.39***					
Homoscedastic test- $X^2(1)$	317.88***					
Root-mean-square-error			254.81		229.03	
Mean-prediction-error			-109.87		$-3.54\text{x}10^{-06}$	
N	138		135		138	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Specification Test for the Tobit & Model Selection

Under parametric assumptions, the Tobit model is more consistent and efficient than the semi-parametric estimators (Powell, 1986), and these assumptions are tested using the Lagrange multiplier following Cameroon & Trivedi, (2010). The results (Table 3) indicate that both the normality and the homoscedasticity tests are significant at p<0.01, thereby making the Tobit model inconsistent. As a result, semi-parametric estimators are considered. From the results of the mean-prediction-errors and root-mean-square errors, both of which require lower values, it is clear that inferences drawn from the SCLS estimator are superior to that of the CLAD estimator. The SCLS is therefore used for the rest of the analyses.

Factors Influencing Choice behaviour in the Choice Experiment (nHCE)

Table 4 shows results of the nHCE from both the conditional and mixed logit[3] along with model fit statistics from 107 respondents. Each individual provided 16 completed choice-tasks and each task had 4 possible outcomes resulting in 6848 observations (See variable descriptions at the end of Table 4). Both models include alternative specific constants of the three maize varieties and the opt-out option. The dependent variable is the subject's choice of a maize variety at varying price levels. It takes the value of one on the chosen alternative, zero otherwise. The Akaike information criteria (AIC), Bayes information criteria (BIC) and the log likelihoods were used to choose between the two models. A model that minimises the AIC and BIC scores, and has a higher log likelihood value is most preferred. Results of these indicate that the mixed logit outperforms the conditional logit, hence it is used to interpret results.

As shown in model 4, the white-maize constant is significant and positive ($p<0.01$), suggesting that an individual would rather have white maize, than none at all. The orange and yellow maize constants are not significant, indicating that respondents were indifferent to choosing them. This was expected in a population where white-maize varieties are preferred. The probability of choosing a given maize variety was negatively associated with price ($p<0.01$), and older respondents were more likely to be price sensitive ($p<0.1$) in the crop-variety choice.

Orange maize choices were positively influenced by age, education and assets, implying the older, more educated and wealthier respondents were more likely to select it ($p<0.01$). The choice of white maize was positively influenced by education ($p<0.01$) suggesting that individuals with more education were more likely to select white maize. Yellow maize choices were influenced by age ($p<0.05$) and asset-index ($p<0.01$), where the older, and poorer individuals were more likely to choose it.

[3] A taste of randomness of all parameters was done using the T-test for standard deviation and only three education variables were found to be random.

Table 4 Parameters Estimates from the non-hypothetical choice experiment

	(Conditional logit)		**(Mixed logit)**	
	β	***Se***	**β**	***Se***
Price^{-02}	-0.24***	(0.03)	-0.28***	(0.06)
Gender*white	-0.05	(0.39)	0.86	(0.89)
Gender*yellow	0.39	(0.40)	0.99	(0.94)
Gender*orange	0.16	(0.39)	1.14	(0.98)
Education*white	0.04	(0.05)	0.44***	(0.16)
Education*yellow	0.02	(0.05)	0.16	(0.17)
Education*orange	-0.03	(0.05)	0.39**	(0.18)
Age*price^{-04}	0.19***	(0.06)	-0.25*	(0.15)
Asset*price-^{03}	-0.02	(0.08)	-0.30	(0.19)
Age*white $^{-02}$	0.08	(1.75)	-0.18	(5.63)
Age*yellow-^{01}	0.03	(0.18)	1.54**	(0.60)
Age*orange-^{01}	0.10	(0.18)	1.500**	(0.61)
Asset*white	-0.31	(0.22)	0.48	(0.43)
Asset*yellow	-0.31	(0.22)	-1.07***	(0.41)
Asset*orange	-0.27	(0.22)	1.88***	(0.51)
ASC				
White	4.61***	(0.85)	5.17***	(1.69)
Yellow	3.84***	(0.85)	-0.71	(1.93)
Orange	5.30***	(0.85)	1.28	(2.13)
SD				
Education*white			0.38***	(0.04)
Education*yellow			0.60***	(0.05)
Education*orange			0.77***	(0.07)
N	6848		6848	
Log likelihood	-1630.61		-981.51	
AIC	3297.20		2005.00	
BIC	3420.20		2148.50	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Price (price of maize in Zambian Kwacha, 800, 1200, 1500, 2000); White (1 if white maize variety is chosen, 0 otherwise); Yellow (1 if yellow maize variety is chosen, 0 otherwise); Orange (1 if orange maize variety is chosen, 0 otherwise); Gender (1 if male, 0 otherwise); Age (years); Education (years);

Comparison of WTP estimates from the BDM and the non-hypothetical choice experiments

To make WTP estimates from the nHCE comparable to that of the BDM, the mean WTP from the choice experiment was computed using equation (14). Predicted values from the SCLS model are used from the BDM since bids are direct WTP estimates as earlier stated in equation (13). The equality of predicted WTP was tested using a two-sample t-test. Table 5 shows that WTP estimates from the two valuation techniques are significantly different (p<0.01), with WTP average from the BDM experiment lower than that of the nHCE.

Table 5 Comparison of WTP estimates for the orange maize variety from the BDM & non-hypothetical Choice Experiment

	Choice (mean)	BDM (mean)	Difference	(p-value)
WTP	2970	1691	1279	0.00

Design Factors Potentially Affecting Bidding and Choice Behaviour for the BDM experiment and the Choice Experiment.

Since the two experimental approaches generate significantly different WTP results, we investigate the experimental design factors (repeated-auctions, ordering-effects and lexicographic behaviour) and their effect on the observed WTP differences in the two valuation techniques. Under classical economic assumptions, WTP is a function of product attributes and demographics. To test the hypothesis of no design effects, WTP was allowed to also depend on experimental design features.

Repeated-Auctions

To account for possible effects of repeated-bidding on WTP in the BDM experiment, results from the two treatments prior the BDM bid are used. Table 6 reports the characteristics of the participants in the BDM under the two treatments. A total of 60 individuals participated in a single bid while 85 participated in repeated bidding prior the main BDM bidding. Results indicate that the mean characteristics of participants were similar.

Table 6 Characteristics of the BDM participants (means)

	Single bidding round	**Repeated bidding**	**Difference**	**(p-value)**
Age	39.47	37.94	1.53	0.48
Gender	0.53	0.60	-0.07	0.43
Marital status	0.92	0.84	0.08	0.15
Education	7.95	8.47	-0.52	0.28
Household size	8.95	8.51	0.44	0.59
No. of people sharing meals	8.80	7.76	1.04	0.13
Asset index	-0.13	-0.00	-0.13	0.44
N	60	85		

In the first treatment, respondents bid for a 2.5kg of orange maize grain in a single auction before their BDM bid. In the second treatment the auctions-rounds were increased to 10. Following Plott's (1996) preference hypothesis of learning through experience and market exposure, the study expects that subjects are more likely to reveal their true preferences in the BDM with repeated-auctions than in a single-auction. Further, it is expected that heuristics of the "buy low type" which subjects normally exhibit in a real market situation even when told to bid optimally (Drichoutis, et al., 2010) will be eliminated with repeated-auctions due to learning.

A mean comparison of round-one bids revealed that there was no significant difference ($p=0.52$) in the mean bids for the single-auction treatment group (1469 ZMK) and the repeated-auction group (1427 ZMK). This suggests that learning levels for both groups at the beginning were similar and potential differences in their final BDM bid can only be attributed to repeated bidding.

To determine whether repeated auctions had an effect on the subject's bidding behaviour, the pooled BDM bids are used as a dependent variable while controlling for repeated auctions (Table 7). An indicator variable "repbid" is used which takes the value of 1 if one's BDM bid came after 10-auction rounds, zero otherwise. It is hypothesized that the 'repbid' coefficient should be different from zero if learning occurred with repeated auctions. Results indicate that this coefficient is positive and significant ($p<0.05$). Subjects that participated in the BDM after 10 auction-rounds on average bid 119.00 ZMK more than those who did not (column1). This finding suggest that repeated-auctions eliminated the heuristics of wanting to "bid low", and provided a better understanding of the experiment, thus yielded more realistic results than the single-auction treatment.

Order-effects

The order of sample tasting is known to affect the sensory ranking scores (O'Mahony, 1986) and since sensory attributes scores are expected to have an endogenous relationship with WTP (as they are both determined by demographics and attitudes), we tested whether this also impacted WTP for orange maize. The results are summarised in Table 7 (2). As stated earlier, all participants evaluated sensory attributes of the three maize varieties prior the experiments, in which 6 different orders of sample-tasting emerged from the data. Dummy variables for 6 different orders are included in modelling bidding behaviour while controlling for repeated-auctions.

Interactions terms of order and repeated-auctions are also included to determine whether sample tasting-order had different effects on the two treatment groups (single vs. repeated-rounds). Results shows that WTP estimates for respondents who tasted maize samples following orders 3 and 4 were significantly higher ($p<0.05$) than the reference category. Both of these orders began with the familiar white-maize. This indicates that if subjects begin the sensory tasting exercise with the conventional white maize, they are more likely to bid more for orange maize. Consistent with O'Mahony's, (1986) assertion, the familiar white maize could have masked the subject's sensory acuity for the new product. Results further revealed that participants in the repeated-round treatment were less likely to have order-effects than those in a single treatment. This was true for order 3 ($p< 0.05$), orders 4 and 6 ($p< 0.1$), suggesting that some of the order-effects observed could be due to inadequate familiarization of the experiment and were reduced through learning from the repetitive treatment.

Experimenter's-effect

The BDM is considered complex relative to the nHCE. Experimenters had to translate instructions from English to the respondent's native language, administer questionnaires and guide the entire bidding process. In light of this, the study determines whether in doing so, experimenters influenced respondents' bidding behaviour. Dummy variables for enumerators are included in the model. Results suggest that the effect of experimenters on WTP outcomes was limited, with only experimenters number 7 and 12 whose subjects bid higher than those for experimenter one at $p<0.1$(Table 7 (3)).

Table 7 WTP estimates from the BDM auction for different design factors

	(1) Repeated auctions		**(2) Order-effects**		**(3) Experimenter's-effect**	
Bid	**b**	**se**	**b**	**se**	b	se
Household size	4.36	(4.12)	2.52	(4.43)	1.34	(4.32)
Respondents' age	3.91	(11.69)	4.10	(11.58)	16.82	(10.84)
Age square	-0.07	(0.12)	-0.08	(0.13)	-0.22*	(0.12)
Gender	113.9*	(61.69)	131.8**	(58.91)	91.14	(55.50)
Education	-10.40	(9.84)	-16.27	(10.30)	-11.59	(8.97)
Asset index	2.23	(13.24)	2.65	(13.33)	0.19	(12.55)
Repbids	118.9**	(55.33)	282.1**	(115.2)	151.9***	(55.04)
Order2 (yellow-white-orange)			-52.45	(156.0)		
Order3 (white-orange-yellow)			281.7**	(128.3)		
Order4 (white-yellow-orange)			271.3**	(135.5)		
Order5 (orange-white-yellow)			14.04	(146.8)		
Order6 (yellow-orange-white)			192.7	(135.5)		
Order2 x Repbid			-30.99	(198.3)		
Order3 x Repbid			-373.7**	(157.8)		
Order4 x Repbid			-303.7*	(158.6)		
Order5 x Repbid			-5.34	(170.2)		
Order6 x Repbid			-348.9*	(182.1)		
Enum2					-230.4	(171.1)
Enum3					162.5	(139.9)
Enum4					-131.0	(180.7)
Enum5					130.8	(152.8)
Enum6					125.2	(155.8)
Enum7					264.6*	(146.7)
Enum8					60.51	(180.4)
Enum9					59.18	(173.1)
Enum10					76.89	(164.6)
Enum11					69.56	(158.3)
Enum12					263.7*	(150.7)
Enum13					152.4	(176.2)
Enum14					237.3	(151.1)
Enum15					165.1	(136.8)
Enum16					79.88	(164.0)
Constant	1561.3***	(254.1)	1514.6***	(265.6)	1235.5***	(286.4)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ Reference order category is 1 (orange-yellow-white)

Price Insensitivity (Lexicographic Preferences)-Choice Experiment

In the nHCE, it was found that 31% of the respondents consistently chose one variety in all 16-choices regardless of the price offered. This was mostly observed with orange maize (21%), followed by white maize (6.54 %) and yellow maize (3.74%). Reasons for such behaviour include: complex or poorly explained experiment; boredom or fatigue from repeated choice-tasks; omission of the relevant attribute. To account for lexicographic behaviour, we allowed the price coefficient associated with lexicographic responses to be zero such that the systematic component in equation (5) is equal to the alternative specific constant i.e., $V_{ij} = \beta_{1j}$ (Campbell , et al., 2006).

Results (Table 8) reveal that controlling for lexicographic behaviour increased model fitness based on the AIC, BIC and log likelihood values. The coefficient signs, significance levels and magnitude changed on some explanatory variables. Notably, the asset-index variable no longer has an effect on orange and yellow maize choices, while education no longer has an effect on orange maize choices. Further, WTP estimates for all varieties are significantly reduced, suggesting overestimation of these estimates without accounting for lexicographic responses.

Table 8 Parameter Estimates from the BDM and non-hypothetical choice experiment for orange maize variety after controlling for price insensitivity (Lexicographic Preferences)

Choice	**No lexicographic preference control**		**Control-lexicographic preferences**	
	β	***Se***	**β**	***Se***
Price.x10^{-02}	-0.28***	(0.06)	-0.39***	(0.04)
Gender*white	0.86	(0.89)	0.90	(0.70)
Gender*yellow	0.99	(0.94)	1.21	(1.10)
Gender*orange	1.14	(0.98)	-0.12	(0.73)
Age*price.x10^{-03}	-0.025*	(0.01)	-0.01	(0.01)
Asset*price x10^{-02}	-0.03	(0.02)	-0.042**	(0.02)
Age*white	0.00	(0.06)	0.00	(0.04)
Age*yellow	0.15**	(0.06)	0.08**	(0.04)
Age*orange	0.15**	(0.06)	0.10**	(0.04)
Asset*white	0.48	(0.43)	0.19	(0.39)
Asset*yellow	-1.07***	(0.41)	-0.53	(0.49)
Asset*orange	1.88***	(0.51)	0.75	(0.46)
Education*white	0.44***	(0.16)	0.30***	(0.11)
Education*yellow	0.16	(0.17)	0.08	(0.11)
Education*orange	0.39**	(0.18)	0.07	(0.10)
ASC				
White	5.17***	(1.69)	5.89***	(1.40)
Yellow	-0.71	(1.93)	2.02	(1.93)
Orange	1.28	(2.13)	5.39***	(1.56)
SD				
Education*white	0.38***	(0.04)	0.24***	(0.04)
Education*yellow	0.60***	(0.05)	0.50***	(0.05)
Education*orange	0.77***	(0.07)	0.42***	(0.05)
N	6848		6848	
Log likelihood	-981.51		-941.74	
AIC	2005.00		1925.50	
BIC	2148.50		2068.90	
WTP(ZMK)				
White	2430.98		2163.40	
Yellow	2003.48		1725.26	
Orange	2970.24		2378.19	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Comparing WTP after controlling for design effects

WTP was computed in each experiment after accounting for repeated bids and lexicographic responses in the BDM and nHCE respectively. Results (Table 9) revealed a WTP gap reduction by half, although the difference is still significant.

Table 9 Comparison of WTP estimates from the BDM & non-hypothetical Choice Experiment after controlling for design effects

	RCE (mean in ZMK)	BDM (mean in ZMK)	Difference	(p-value)
WTP before	2970	1691	1279	0.00
WTP after	2378	1,747	*632*	0.00

Conclusion

Are WTP estimates from the Becker-DeGroot-Marschak (BDM) and the non-hypothetical choice experiment (nHCE) comparable under the same research context as economic theory would predict? If not, can design factors account for some of the differences observed under the two mechanisms? To answer these questions, we use data from a consumer acceptance study for a new improved maize variety enriched with Vitamin A in rural Zambia. White maize varieties are preferred which raises concerns since the new variety is orange in colour.

Average WTP estimates were derived from both the BDM and the nHCE in an actual payment setting (i.e., using real money and real maize variety products). This provided a platform for consumers to reveal their true value for the new orange maize variety under both methods. Participants tasted and evaluated the sensory attributes of the maize variety samples to familiarize themselves with all varieties before participating in the experiments. There was no information provided about the nutrition value of the maize varieties, hence participants bid and chose based on their own sensory evaluations of maize samples.

A comparison of mean WTP estimates elicited from the BDM and the nHCE for orange maize revealed that estimates from the nHCE were significantly higher than those from the BDM experiment, confirming results from previous findings. We further show that this disparity can be reduced by half by employing extra bidding rounds in the BDM experiment and controlling for lexicographic behavior (choosing one maize variety regardless of price) in the choice experiment. However, controlling for these designs (lexicographic behaviour and extra bidding rounds), is not statistically significant in reducing the WTP gap between the two experimental designs, but could be meaningful from the economic point of view. These results are consistent with Banerji, et.al., (2013) who also find that controlling for censored bids and lexicographic answering make the estimates under the two valuation techniques more comparable.

Participants in the BDM were also found to be susceptible to order-effects or the order in which participants tasted maize varieties' samples. Although the order effects of food sample tasting have been associated with affecting sensory scores (O'Mahony, 1986), this study indicated that this effect could persist in the BDM experiment by affecting WTP scores. This persistence, however, was only observed in participants who completed fewer auction rounds in the experiment, suggesting lack of understanding of the experimental procedure. Given that the BDM has been considered complex (Cason & Plott, 2014), it is likely that repeated bidding could have made the bidding decisions more clear to the respondents. This suggests that future research should take steps to either simplify the procedures, or provide more practice with them.

The reasons for lexicographic behaviour are beyond the scope of this paper, but need further research. Various reasons have been proposed, including: complicated or poorly explained choice tasks; poorly designed experiments (missing relevant attribute); being truly lexicographic (one attribute is truly desired), as outlined by Killi, et al., (2007), who suggest adding qualitative questions covering reasons for such a behaviour.

Our results also indicate that individuals exhibit different behaviors under the two valuation techniques. While gender explained the bidding behaviour in the BDM experiment, age influenced the choice behaviour in the nHCE for orange maize. In interpreting these results, we acknowledge one limitation in our study. Although participants under both experiments were similar in most characteristics, the participants under the nHCE were 3 years older. This could be one of the reasons why mean WTP for orange maize was higher for the nHCE given the positive effect of age. However, we did not expect any difference in the marginal WTP or any systematic error since there was no influence of age on other varieties in the nHCE.

Although it is not yet clear how the BDM and the nHCE can result in significantly different empirical estimates when addressing the same research question under similar conditions, these results suggest that part of the differences observed could be as a result of design effects from the two techniques. Adequate training in the BDM seems to be necessary to fully reveal its incentive structure. Similarly, controlling for lexicographic behaviour or insensitivity to other attributes is necessary in the choice experiment to attain valid estimates.

References

Banerji, A., Chowdhury, S., de Groote, H., Meenakshi, J. V., Haleegoah, J., and Ewool, M. *Using elicitation mechanisms to estimate the demand for nutritious maize. Evidence from experiments in rural Ghana.,* Working Paper No. 10 (HarvestPlus: 2015).

Becker , G. M., DeGroot, M. H. and Marschak, J. 'Measuring utility by a single-response sequential method', *Behavioral Science,* Vol. 9, (1964) pp. 226–232.

Berry, J., Fischer, G. and Guiteras, R. *Incentive compatibility in the field: A test of the Becker-Degroot-Marschak mechanism.* 2007. Available at: http://www.ncsu.edu/cenrep/workshops/TREE/documents/Guiteras_April2011.pdf (last accessed 12 August, 2017)

Cameroon, C. A. and Trivedi, P. K. *Microeconometrics Using Stata.* Revised edition (Stata Press, 2010).

Campbell , D., Hutchinson, G. W. and Scarpa, R. *Lexicographic preferences in Discrete Choice Experiments: Consequences on Individual-Specific Willingness to Pay Estimates.* Working Paper No.128 (Fondazione Eni Enrico Mattei: Nota Di Lavoro Italy, 2006).

Cason, T. and Plott, C. 'Misconceptions and game form recognition: Challenges to theories of revealed preference and framing', *Journal of Political Economy,* Vol. 122(6), (2014) pp. 1235-1270.

Chowdhury, S., Meenakshi, J. V., Tomlins, K. and Owori, C. 'Are consumers in developing countries willing to pay more for micronutrient dense biofortified foods? Evidence from a field experiment in Uganda', *American Journal of Agricultural Economics,* Vol. 93(1), (2011) pp. 83-97.

Cummings, R. G. and Taylor, L. O. 'Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method', *American Economic Review.* Vol. 89(3), (1999) pp. 649-665.

De Groote, H., Chege, S. K., Tomlins, K. and Gunaratna, N. S. 'Combining experimental auctions with a modified home-use test to assess rural consumer's acceptance of quality protein maize, a biofortified crop', *Food Quality and Preference,* Vol. 38, (2014) pp. 1-13

De Groote, H. & Kimenju, S. C. 'Comparing consumer preferences for color and nutritional quality in maize: Application of a semi-double-bound logistics model on urban consumers in Kenya', *Food Policy,* Vol .33(4), (2008), pp. 362-370.

Drichoutis, C. A., Nayga Jr., R. M. and Lizaridis, P., 2010. 'The Role of Training in Experimental Auctions', *American Journal of Agricultural Economics,* Vol. 93(2), (2010), pp. 521-527.

Filmer, D. & Pritchett, L. H. 'Estimating wealth effects without expenditure data – or tears: An application to educational enrollments of India'. *Demography,*Vol. 38(1), (2001), pp. 115-132.

Gracia, A. L., Loureiro, M. L. and Nayga, R. M. J. 'Are valuations from non-hypothetical choice experiments different from those of experimental auctions?' *American Journal of Agricultural Economics,* Vol. 93(5), (2011), pp. 1358-1373.

Hole, A. R. and Kolstad, J. R. ‘Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and WTP space using data from a health-related choice experiment’, *Empirical Economics*, Vol.42(2), (2012), pp. 445-469.

Holt, C. ‘Preference reversals and the independence axiom’, *American Economic Review,* Vol. 76, (1986) pp. 508-515.

Horowitz, J. K. The Becker-DeGroot-‘Marschak mechanism is not necessarily incentive compatible, even for non-random goods’, *Economic Letters,* Vol. 93, (2006) pp. 6-11.

Karni, E. and Safra, Z. ‘Preference reversal' and the observability of preferences by experimental methods’, *Econometrica,* Vol. 55(3), (1987) pp. 675-685.

Killi, M., Nossum, A. and Veisten, K. ‘Lexocographic answering in travel choices: Insufficient scale extensions & steep indifference curves?’ *EJTIR,* Vol 7(1), (2007)pp. 39-62.

Lancaster, K. J. ‘A new approach to theory’, *Journal of Political Economy,* Vol.74(2), (1966) pp. 132-157.

Louviere, J. J. and Woodworth, G. ‘Design and analysis of simulated consumer choice or allocation experiments: An approach based on aggregate data’, *Journal of Marketing Research,* Vol. 20, (1983) pp. 350-367.

Lusk, J., Alexander, C. and Rousu, M. ‘Designing experimental auctions for marketing research: The effect of values, distributions, and mechanisms on incentives for truthful bidding’, *Review of Marketing Science,* Vol.5 (1), (2007) pp. 1-32.

Lusk, J. L. and Schroeder, T. C. ‘Auction bids and shopping choices’, *Advances in Economic Analysis & Policy,*Vol. 6(1), (2006) pp. 1-37.

Lusk, J. L. and Shogren, J. F. *Experimental Auctions: Methods and applications in economic and marketing research (*Cambridge, UK: Cambridge University Press, 2007).

McFadden, D. *Conditional logit analysis of qualitative choice behavior. In Frontiers in Econometrics* (New York: Academic Press, 1974).

Meenakshi, J. V., Banerji, A., Manyong, V., Tomlins, K., Mittal, N., and Hamukwala, P. ‘Using discrete choice experiment to elicit the demand for a nutritious food: Willingness to pay for orange maize in rural Zambia’, *Journal of Health Economics,* Vol. 31, pp (2012). 62-71.

Morawetz, U. B., De Groote, H. and Kimenju, S. C. ‘Improving the use of experimental auctions in Africa: Theory and evidence’, *Journal of Agricultural and Resource Economics,* Vol. 36(2), (2011) pp. 263-279.

O’Mahony, M., 1986. ‘Sensory adaptation’, *Journal of Sensory Studies, Vol.*1(3), (1986) pp. 237-258.

Oparinde, A., Birol, E., Murekezi, A., Katsvairo, L., Diressie, M. T., Nkundimana, J. D., and Butare, L.*Consumer acceptance of biofortified iron beans in Rwanda: Experimental evidence,*Working paper No. 18 (HarvestPlus: 2015).

Oparinde, O., Banerji, A., Birol, E. and Ilona, P..*Information and consumer willingness to pay for biofortified yellow cassava: Evidence from experimental auctions in Nigeria,* Working paper No. 13 (HarvestPlus:2014).

Plott, C. R. 'Rational Individual Behaviour in Markets and Social Choice Processes: The Discovered Preference Hypothesis', in Arrow K. J. et al, (eds.), The Rational Foundations of Economic Behaviour (Basingstoke: International Economic Association and Macmillan, 1996, pp. 225–250).

Powell, J. L. 'Least Absolute Deviations Estimation for Censored regression model', *Journal of Econometrics,* Vol. 25(3), (1984) pp. 303-325.

Powell, J. L. 'Symmetrically trimmed least squares estimation for Tobit models', *Econometrica,* Vol. 54(6), (1986).pp. 1435-1460.

Stevens, R. & Winter-Nelson, A. 'Consumer acceptance of pro-vitamin A biofortified maize in Maputo', *Food Policy,* Vol. 33(4), (2008) pp. 341-351.

Tobin, J. 'Estimation of relationships for limited dependent variables'. *Econometrica,* Issue January, 26, (1958) pp. 24-36.

Voelckner, F. 'An empirical comparison of methods for measuring consumer's willingness to pay', *Marketing Letters,* Vol. 2, (2006) pp. 137-149.

Whitting, D. Improving the Performance of Contigent Valuation Studies in Developing Countries, *Environmental & Resource Economics,*Vol. 22(1), (2002) pp. 323-367.

Zawojska, E. and Czajkowski, M. *Re-examining empirical evidence on contingent valuation-Importance of incentive compatibility,* Working Paper No. 8/2015 (156) (Warsaw: 2015).