The Formation of Reference Points in Consumer Choice Behavior: Experimental Evidence From A Fish Market in Nigeria

by Kelvin Mashisia Shikuku, Erwin Bulte, Carl Johan Lagerkvist, and Nhuong Tran

The Formation of Reference Points in Consumer Choice Behavior: Experimental Evidence from a Fish Market in Nigeria

Kelvin Mashisia Shikuku[a,*], Erwin Bulte[b], Carl Johan Lagerkvist[c], Nhuong Tran[d]

[a] WorldFish, Jalan Batu Maung, Batu Maung, 11960 Penang, Malaysia.

[b] Department of Social Sciences, Development Economics Group, Wageningen University, P.O. Box 8130, 6700EW Wageningen, The Netherlands.

[c] Department of Economics, Swedish University of Agricultural Sciences, P.O. Box 7013, 750 07 Uppsala, Sweden.

[d] WorldFish, c/o:Department of Fisheries, Bayint Naung Rd, West Gyogone, Insein Township,Yangon, Myanmar.

* Corresponding author; Email: k.shikuku@cgiar.org

A contributed paper submitted for presentation at the 31st Triennial International Conference of Agricultural Economists held virtually on August 17-31, 2021

Abstract

We study the role of endowments and expectations in the construction of reference points in the context of fish trading on fish markets in Nigeria. In our field experiment, consumers can trade a known food item for a novel, superior food item (and vice versa). Endowments matter for reference point formation, but the effect of expectations about future ownership is mixed. While expectations affect bidding behavior for subjects "trading up", it does not affect bids for subjects "trading down". We probe the role of aspirations as a mediating factor but document no evidence in support of it.

Key words: Expectations, reference-dependent utility, economic valuation, food safety, loss aversion

JEL Codes: C57, C93, D12, D44, D84

1. Introduction

In many circumstances, individuals make choices by evaluating possible outcomes against a reference outcome. This behavioral tendency is known as reference-dependence—a concept introduced by Kahneman and Tversky (1979). Reference-dependent utility and its implications have been studied in various settings, including technology adoption (Dupas, 2014), demand for housing (Simonsohn and Loewenstein, 2006), labor supply (Crawford and Meng, 2011; Bulte et al., 2020), and food choice (Caputo et al., 2019). The workhorse model of reference-dependent utility, proposed by Kőszegi and Rabin (2006), assumes that an individual's utility consists of two components: conventional consumption utility and so-called gain-loss utility. Gain-loss utility emerges as outcomes deviate from a reference point. Loss aversion implies that downward deviations from reference points cause greater utility losses than the gains generated by equal-sized positive deviations.

Reference points are "constructed", and the determinants of reference points have emerged as an important research topic. People may base reference points on current endowments (the status quo), but Kahneman and Tversky (1979) argue that "*there are situations in which gains and losses are coded relative to an expectation or aspiration level that differs from the status quo*." The suggestion that subjects may form reference points based on expectations was elaborated by Kőszegi and Rabin (2006). Early work, based on observational data, struggled to empirically distinguish between the roles of endowments, aspirations and expectations about ownership, as these concepts tend to correlate.[1] More recent experimental work on the origins of reference points produces ambiguous results (see below).

[1] Another challenge is distinguishing between how expectations affect reference points and other motivations affecting choice behavior. For example, Wenner (2015) uses expected prices as reference points, and must address the challenge that higher prices may also be correlated with expectations regarding product quality. Ericson and

It is difficult to overstate the importance of reference point formation for trading and purchasing behavior of consumers. For example, outcomes will depend on whether consumers construct reference points based on expected prices of goods, or on expectations with respect to owning the goods (Ericson and Fuster, 2011). Anticipating that gain-loss utility considerations enter consumer decision-making, firms should adjust price setting and marketing policies (managing expectations in a desirable direction). Gain-loss utility will also affect appreciation of public good supply, and should therefore enter in cost-benefit analyses of governments.

In this paper, we use a field experiment in a low-income country to study food trading behavior in a real market setting—a Nigerian fish market. We endow subjects with either a conventional fish commodity or a superior alternative item. In addition, we vary expectations about future ownership by assigning participants to treatment arms with high or low probabilities of being allowed to trade. Our experiment is based on a factorial design. Economic values are measured by eliciting willingness to pay (or accept) for trading the fish endowments in a Becker-DeGroote-Marschak (BDM) auction. Extending the literature, we ask participants to trade-off asymmetric commodities that can be ranked in terms of utility. The new commodity that we introduce on the local market is food-safety certified fish, which should be (weakly) preferable.

Our main objective is to examine whether endowments and expectations affect reference point formation for this novel sample population in this novel context. We find strong support for reference-dependence in our sample. The minimum willingness to accept (WTA) for exchanging the superior item for the conventional one is 40% greater than the maximum

Fuster (2011) emphasize the importance of "transparent randomization" into treatment as subjects may otherwise try to make sense out of the task at hand, which may affect choice behavior.

willingness to pay (WTP) for the reverse trade. Our second result is that expectations about future ownership appear to matter for the construction of reference points in some trades, but not others. Expectations affect bidding behavior for subjects endowed with the conventional fish commodity, but not for subjects endowed with the superior alternative. We probe whether this asymmetric response is caused by the mediating effect of aspirations (about "trying out" the new product), but our data do not support such an explanation.

Our experiment extends the literature along several dimensions. First, we identify expectation-based reference points in the field, rather than a lab setting. We implement an auction experiment in a real fish market with regular customers, and use two types of fish as experimental items. Second, the commodities that are traded are "different" in a meaningful way. While subjects in lab experiments are typically asked to choose between items of the same (token) value, such as university mugs versus pens, we use the same item (500 grams of live catfish) and introduce a distinction based on food safety certification. Food safety is a key concern for our sample population. Certified fish should be (weakly) preferred over uncertified fish, and is new to the consumers in our market context; it is a desirable item that consumers aspire to try out. Third, while we do not have exogenous variation in aspirations, or clean measurement of aspirations, our survey-based data allow us to probe the mediating role of aspirational factors in bidding behavior.

This paper is organized as follows. In section 2 we briefly summarize the empirical literature on reference point formation. In section 3 we sketch the background and context of our own experiment, and introduce the experimental design and data. Section 4 presents the regression results about reference dependent utility and the formation of reference points, based on experimental and survey data. The final section concludes.

2. Reference point formation

The Kőszegi and Rabin (2006) model assumes that utility depends on a consumer's *k*-dimensional consumption vector $\boldsymbol{c}$ and on a reference vector $\boldsymbol{r}$, as follows:

$$u(\boldsymbol{c}|\boldsymbol{r}) = \sum_k m_k(c_k) + \sum_k \mu(m_k(c_k) - m_k(r_k)). \quad (1)$$

Utility depends on two components, separable across dimensions. The first term on the right hand side captures classical utility, or utility derived from consuming good *k*. The second term captures gain-loss utility, which is where reference-dependence enters. Value function μ is defined as: $\mu(x) = \eta x$ for $x > 0$ and $\mu(x) = \eta\lambda x$ for $x < 0$. Parameter η is the weight attached by the consumer to gain-loss utility and $\lambda > 1$ is the coefficient of loss aversion. The latter coefficient captures that utility losses associated with outcomes c_k below reference value r_k are greater than utility gains from equal-sized realizations in excess of that reference point.

Where does reference vector $\boldsymbol{r}$ in (1) come from? Early papers of reference-dependent utility demonstrated that reference points can be manipulated through random assignment of items, suggesting that status quo endowment levels determine reference points. For example, many subjects *ex-post* prefer an item randomly distributed to them over another item of the same monetary value (e.g., Kahneman and Tversky, 1979; Thaler, 1980; Knetsch, 1989). However, reference point formation may be a complex process, in which additional considerations could enter as well. One prominent alternative (or additional) candidate determinant of reference points is expectations about future ownership. According to the Kőszegi and Rabin (2006) theory, people base reference points on expectations about outcomes which, in turn, are the result of people's own anticipated behavior (which should be consistent with actual behavior). This is called a *personal equilibrium*.

As noted, it is hard to distinguish expectations from endowments (in observational data) if subjects expect to keep their endowment. However, expectations can be exogenously varied

in an experimental setting. Several lab experiments seek to manipulate expectations to probe their causal effects on valuation. Abeler et al. (2011) organized a real-effort experiment where subjects either received a fixed wage (50% probability) or a payment based on performance (50% probability). Raising the fixed wage caused subjects to work harder. This is consistent with the hypothesis that an increase in expected experimental income pushed up the reference point, crowding in additional effort because subjects want to avoid the disappointment of receiving less than their fixed wage if performance-based payment was selected for pay-outs.[2]

Another influential example is Ericson and Fuster (2011). In their first experiment, participants received a mug and were randomly assigned to one of two treatment arms. In one arm subjects were told that they could exchange their item for another one (a pen) with 10% probability. In the other arm, subjects could exchange with 90% probability. In both arms, participants had to decide whether they would trade their item, conditional on receiving permission to do so. Individuals in the low probability arm were more likely to keep their endowment. This is consistent with the hypothesis that expectations about future ownership shape reference points and economic values. In a follow-up experiment, Ericson and Fuster (2011) randomized the probability that participants would receive a mug, and subsequently elicited their WTA for it. When an individual expected to receive an item, their WTA for that item increased.[3]

The third (lab) experiment supporting the notion of expectation-based reference points is provided by Banerji and Gupta (2014). In a BDM auction setting they experimentally varied the probability of winning the auction (conditional on bidding), by varying the support from

[2] Observe that this experiment, with exogenous variation in expected income, may capture both an effect of expectations on reference points as well as an effect of aspirations.

[3] Findings by Crawford and Meng (2011) and Gill and Prowse (2012) are also consistent with the hypothesis that expectations influence behavior, in a labor supply and real-effort task context, respectively.

which the strike price is randomly drawn. An important finding is that loss averse subjects bid lower when their probability of winning goes down, or when the expected value of the strike price goes up.

However, not all experimental evidence supports the idea that expectations shape reference points. Heffetz and List (2014) presented participants with a mug and a pen, and randomly assigned one of the items to them. They manipulated expectations by assigning subjects to either a "weak expectations treatment" (where they received the assigned item with 1% probability, and with 99% probability they could choose one of the items) or a "strong expectations treatment" (where these probabilities were reversed). While participants were more likely to choose the item assigned to them by chance, choice behavior was statistically identical across the expectation treatments. Heffetz and List (2014) therefore conclude that the endowment effect exists, but is not driven by expectations. This conclusion is supported by Wenner (2015), who manipulated price expectations in a lab setting. Wenner (2015) found that buying behavior was not affected by the *ex ante* distribution of prices.[4]

[4] Finally, Smith (2019) explores whether lagged beliefs affect reference point formation. He sought to induce reference points by making subjects participate in a lottery with different probabilities of winning a prize, and then tested whether this probabilistic reference point affects valuation *after* the outcome of the lottery was realized (using a BDM design). While current endowments affected valuation, lagged beliefs did not affect bidding behavior. This finding may reflect that the timing structure of the experiment does not match actual reference point formation. The effect of lagged expectations about lottery outcomes may be swamped by the effect of the actual outcome of the lottery—the literature provides little guidance regarding the speed of reference point adjustment in response to new information (Smith 2019). Also see Karle et al. (2015) for theory and experimental work on price expectations, loss aversion and choice behavior.

3. Context and experiment

3.1 Context

We implemented our experiment in Lagos State, Nigeria. Fish accounts for over 40% of total protein intake in Nigeria, the largest aquaculture producer in sub-Saharan Africa (WorldFish, 2018). However, food safety is an issue, with potential health hazards arising from contaminants during farming, or quality loss due to poor handling and processing. Misuse of antimicrobials is recognized as a key driver of the emergence and spread of antimicrobial resistance. Foodborne illnesses caused 420,000 deaths and 33 million disability-adjusted life years (DALYs) worldwide in 2010 (World Health Organization, 2015). One approach to improve production standards is certification for food safety (e.g., Bush et al., 2013; Birol et al., 2015).

Nigeria's Federal Department of Fisheries and Aquaculture initiated the process of voluntary certification and standardization of fisheries and aquaculture products in 2009. As part of this process, operational guidelines and criteria for certification of aquaculture products in the country were developed to standardize operation of fish farms and to minimize hazards to human health. The aim is to increase consumer benefits, confidence, and traceability in aquaculture production, processing, and marketing. However, while consumers are well-aware of the various food safety concerns associated with the consumption of uncertified fish, certified fish products are currently unavailable in the fish markets where we conducted our study. Certified fish is a "novel" and salient product for our subjects.

3.2 Experimental design

We partnered with Lagos State University to conduct a framed field experiment in October–November 2019 in a real fish market setting, attended frequently (often on a daily basis) by our participants. A pre-analysis plan (PAP) was registered at

https://aspredicted.org/rj2mu.pdf on September 27th, 2019 prior to implementation of the experiment.[5] We recruited participants on the market, seeking to purchase fish. As part of the experiment, we used two distinct catfish commodities for trading, similar in all but one important aspect—whether or not it was certified for food safety by the government (introduced by us on the local market).[6] We collected experimental data on two fish markets.

The experiment was pre-tested. Prior to the experiment, market leaders were visited by the research team to explain the purpose of the study, introduce the field team and identify suitable places to set up the experiment. Each second consumer entering the market was approached and asked to participate in the experiment, until we reached a sample size of 400 participating market visitors. 19% of those invited consumers agreed to participate, with the remaining 10% declining because they could not spare 30 minutes—the duration of the experiment. Participants were informed that they participated in a field experiment run by a research team (not by salespersons for any particular type of fish). Participants received a show-up fee of 1,000 Naira (≈ USD 2.8) in an envelope and their fish endowment.

We implemented a factorial design. In Arm 1, 200 consumers received 500 grams of uncertified live catfish. It was explained that the fish was obtained from the local market, so that conventional production and handling standards applied.[7] In Arm 2, 200 other consumers received 500 grams of safety-certified live catfish. We clearly explained that the fish was

[5] The research protocol was reviewed and approved by the Institutional Review Board (IRB) of the International Food Policy Research Institute (IFPRI)—IRB application approval number: MTID-19-1159.

[6] Prior to the experiment, a list of certified producers was obtained from the government. Certified producers were contacted to collect information about types of fish produced, the different forms in which products are sold, main buyers, typical sizes sold, prices charged, and whether fish would be available for selling during the study.

[7] Obtaining uncertified fish from the local fish traders also helped to increase cooperation and to avoid creating the perception that our enumerators were competing with the traders.

produced by a fish farmer approved by the government for following food safety guidelines, and what this meant. In Arm 1, we measured WTP to "trade up" and exchange the standard fish for certified fish. We measured WTA for subjects in Arm 2 to "trade down" and exchange their certified endowment for the standard quality. Orthogonal to the endowment arms, we randomly varied the probability that subjects were obliged to keep their endowment, or were allowed to trade it for the other item (conditional on their bid in the auction). In other words, we have the following design.

<< Insert Figure 1 about here >>

To probe the role of reference-dependent utility in fish purchasing, and the potential role of expectations in the formation of reference points, we implemented a 6-step protocol:

Step 1 (Introduction). Subjects receive their show up fee and endowment, and information about the fish item. Next, they are informed that they may be able to trade their fish endowment for another one. Specifically, in Arm 1, subjects learn that they may be able to exchange their uncertified fish for certified fish (and we explained what this meant). Similarly, in Arm 2, subjects learn that they may be able to trade their certified fish for uncertified fish.

Step 2 (Expectation manipulation). Subjects are randomly (and transparently) assigned to one of two treatments, with either "low" or "high probability" of being allowed to exchange. A coin is flipped in front of the participant, with sides labelled "1" and "9". The participant then receives an index card with the resulting number on it, and is told that they can exchange their endowment for the other commodity if and only if a 10-sided die to be rolled at the end of the session comes up *lower* than the number on the index card. Hence, a participant whose coin comes up "1" has a 10% probability of being allowed to exchange. In contrast, a

participant whose coin comes up "9" has a 90% chance of being able to exchange.[8] This was clearly explained (and practiced). This approach introduces exogenous variation in expectations as subjects throwing a "1" expect to go home with their endowment.

Step 3 (Survey 1). Participants answer a few demographic questions.[9] The purpose is to allow time for reference points to form in response to an individual's plans regarding whether to exchange fish if given the opportunity to do so.

Step 4 (Valuation). All subjects participate in a BDM auction, which was clearly explained with trial runs. Using a multiple price list format, participants in Arm 1 are asked *ex ante* to state their WTP to exchange their uncertified fish for certified fish. The highest that respondents were willing to make was compared with a randomly drawn strike price (between 0–400 Naira) in a sealed envelope. Similarly, participants in Arm 2 are asked *ex ante* to state their WTA to exchange their certified fish for uncertified fish (and their minimum bid was compared to a random strike price between 0–400 Naira). Before bidding, all participants are reminded that effectuating the exchange depend on both the roll of the ten-sided die, and the level of their bid relative to the strike price.

Step 5 (Survey 2). Participants fill out a second short survey including questions about consumption and expenditures, and food safety knowledge. We also used a 5-point Likert scale, from "disagree strongly" to "agree strongly", to measure the extent to which subjects agreed

[8] In order to ensure full understanding by participants about the idea of chance, we used blue and orange balls to explain the chance of success and failure, respectively. Specifically, the 10% chance was explained by showing a combination of one blue ball and nine orange balls. Similarly, the 90% chance was explained by showing a combination of nine blue balls and one orange ball.

[9] The questions focused only on demographic characteristics to avoid priming the respondent. Enumerators were strictly advised not to ask any additional questions at this stage. The questions included age and main source of livelihood of the respondent; household size and number of resident household members under five years of age; and highest level of education in the household.

with the following statements: (*a*) "Since the beginning of the session, I have spent some time thinking about how I would use the certified fish"; (*b*) "Since the beginning of the session, I have spent some time thinking about how I would use the non-certified fish"; and (*c*) "Since the beginning of the session, I have spent more time thinking about the non-certified fish than about the certified fish". We use these data to construct a rough proxy for aspirations, below.

Step 6 (Implementation). The die is rolled and, depending on the outcome (and bidding behavior), the trade is effectuated—fish is exchanged and corresponding payments are made. None of the subjects reneged and changed their minds.

To mitigate potential "house money" effects arising from the unexpected receipt of money (List and Price, 2016), we used a cheap talk script.[10] Some 95% of the participants did not open the envelope containing the participation fee; instead, they used their own money in the experiment. To mitigate concerns about contamination due to social learning (e.g., Magnan et al. 2016) and participants "gaming" the experiment, we worked with a team of 10 enumerators to keep the duration of the study as short as possible, and changed locations randomly within the same market.

4. Empirical strategy

For our main empirical strategy, we exploit both experimental margins: variation in endowments and variation in expectations. To test whether endowments matter for the formation of reference points we follow the literature and analyze whether endowments affect bidding behavior. We compare average WTP for trading up in Arm 1 (groups 1 and 2) to average WTA for trading down in Arm 2 (groups 3 and 4). To test whether respondents use their endowment as the reference point, we therefore estimate the following model:

[10] House money effects may emerge if participants spend part of their endowment in the experiment since they do not consider this their own money (windfall), or because they feel obliged to reciprocate to the experimenter.

$$y_i = \alpha + \beta Certify_i + \boldsymbol{\delta_i x_i} + \boldsymbol{C_c} + \epsilon_i. \quad (2)$$

In (2), y_i measures the natural log of consumer i's bid. The variable $Certify_i$ denotes the dummy associated with Arm 2, equal to one if consumer i is assigned to receive certified fish, and zero otherwise. The vector $\boldsymbol{x_i}$ captures demographic characteristics: *age*, *sex*, *education*, and *knowledge* (included in some models). We further control for enumerator effects, vector $\boldsymbol{C_c}$ captures market fixed effects, and ε_i is a random error term. Anchoring on your own endowment drives up WTA for participants in Arm 2, and drives down WTP for participants in Arm 1 (compared to consumers with a conventional utility function, without a gains-loss utility term). We therefore expect $\beta > 0$. We also compare cumulative bid distributions under arms 1 and 2 to check whether one bid distribution stochastically dominates the other one.

To test whether expectations matter for the creation of reference points, we exploit exogenous variation in the probability of retaining the endowment (due to random differences in the probability of being allowed to trade). That is, we compare bids of respondents from groups 1 and 3 to bids from groups 2 and 4. We identify the effect of expectations by regressing bidding behavior of participant i on a treatment dummy capturing whether this participant has a high probability of being able to trade ($Trade_i$). However, observe the following. If gain-loss utility matters, respondents expecting to keep their endowment should *bid less in Arm 1 (as they anchor on uncertified fish) and bid more in Arm 2 (as they anchor on certified fish)*. Lumping the two effects together obfuscates results. Therefore, we must estimate a model that allows identification of the effects of expectations for these groups of subjects separately:

$$y_i = \alpha + \beta Certify_i + \gamma Trade_i + \theta Trade_i Certify_i + \boldsymbol{\delta_i x_i} + \boldsymbol{C_c} + \epsilon_i. \quad (3)$$

In (3), parameter β (again) captures the effect of endowments, γ captures the effect of expectations for subjects from Arm 1, and $\gamma + \theta$ captures the effect of expectations for subjects from Arm 2. As before, we expect $\beta > 0$. We also expect $\gamma > 0$ as subjects expecting to trade

are less likely to anchor on their endowment of uncertified fish, and therefore WTP more for an upgrade. Finally, we expect $\theta < 0$, reflecting that subjects with certified fish who expect to trade are less likely to anchor on their endowment of certified fish, and hence WTA less for a downgrade.

4.1 Data

Comparing respondents across arms reveals that they are similar in terms of observables and stated preferences. This is shown in Appendix Table A.1. Table 1 demonstrates that random assignment to the trade treatment within the arms also succeeded in generating comparable groups. This is tested formally through an *F*-test of joint orthogonality using a logit regression, which tests whether the observable characteristics in Table 2 are jointly unrelated to treatment status. We cannot reject this null hypothesis (*p*-value = 0.685 for Arm 1 and *p*-value = 0.381 for Arm 2), suggesting that randomization succeeded in achieving balance.

<< Insert Table 1 about here >>

Some 80% of our respondents are female, and the average age is 38 years. Nearly 90% of the respondents had completed at least primary level education. On average, 14% of the household members living with our participants are infants—a demographic group particularly vulnerable to foodborne illnesses (Jaffee et al., 2019). Very few study participants (3%) engaged in salaried employment. Participation in salaried employment was significantly higher in the high probability of exchange treatment group than in the group with low probability to exchange, in Arm 1. Households spend, on average, 6,980 Naira per person every month for food purchases. Per capita fish consumption equals 0.5 kg per week—illustrating the importance of aquaculture for the livelihoods of our respondents. On average, subjects are aware of food safety issues. 85% of study participants in Arm 1 and 96% in Arm 2 correctly identified at least one reason why consumption of fish can be unsafe. In addition, 82% of

participants in Arm 1 and 84% in Arm 2 were aware of the risks associated with eating unsafe fish. There was no statistically significant difference in awareness between groups 1 and 2 in Arm 1 and groups 3 and 4 in Arm 2. Furthermore, a *t*-test showed that consumers who completed primary education had greater knowledge about food safety than those who did not (*p*-value=0.019).[11]

5. Results

5.1 *Endowments and reference-dependent utility*

We first examine whether Nigerian fish consumers anchor on their (random) endowment, and compare WTP for trading up in Arm 1 with WTA for trading down in Arm 2. If participants use their endowment as a reference point for a possible gain-loss utility term, we expect reference-dependence to push up WTA and push down WTP. The results are consistent with this prediction. We report results with and without controls, and always include market fixed effects.

<< Insert Table 2 about here >>

As is evident from the top row of columns (1-4) in Table 2, subjects endowed with certified fish on average require greater compensation to part with their endowment than subjects endowed with uncertified fish are willing to pay. Endowments matter for the formation of reference points. This finding is robust across specifications. Turning to an unconditional comparison of sample means, we find that average WTA equals 267 Naira and average WTP equals 187. In words, WTA is 43% higher than WTP, which is a gap that is not only statistically

[11] We asked about reasons why fish can be unsafe to eat; the risks associated with eating unsafe fish; ways to protect themselves from eating contaminated fish; and whether they knew approaches used by farmers, traders, and the government to ensure production and consumption of safe fish products. Knowledge levels were computed as the sum of all correct answers.

but also economically significant. The standardized mean difference equals 0.58, which exceeds the common threshold of 0.25 for "substantial effects" (Cochran and Rubin 1973).[12]

<< Insert Figure 1 about here >>

Figure 1 plots the cumulative density functions of WTA and WTP. Arm 1's bid distribution stochastically dominates the one from Arm 2. The *p*-value for the Wilcoxon rank-sum test, at <0.001, provides strong confirmation of stochastic dominance.

5.2 Expectations and reference point formation

We now turn to the role of expectations and reference point formation. The empirical support for these predictions is mixed, and suggest that expectations matter for reference point formation and economic valuation in some contexts, but not others. This follows from our (OLS) regression analysis, of which results are summarized in Columns (3-4) of Table 2.

First consider the effect of expectations for subjects in Arm 1. The *Trade* variable enters significantly for this subsample, consistent with the idea that expectations about future ownership are weaker which increases WTP. Hence, $\gamma > 0$, as predicted. A *t*-test indicates a statistically significant difference in stated bid to "trade up" equal to 76 Naira ($p < 0.10$).

However, there is no evidence that expectations affect WTA to trade down. It is clear from columns (3-4) that $\gamma + \theta \approx 0$. This is confirmed by a robustness analysis reported in Appendix Table A.2, where we split the sample and analyze the bidding behavior of subjects from Arm 1 and Arm 2 separately. While the *Trade* variable is significant for subjects from

[12] Similar outcomes are found when we compare WTP and WTA distinguishing between the two treatment arms. For example, for the high probability of exchange treatment, Mann-Whitney/Wilcoxon rank sum test for equality of distributions rejected the hypothesis that the two distributions are equal ($p = 0.056$).

Arm 1, we obtain a precisely estimated null result for subjects from Arm 2.[13] The former result is consistent with Ericson and Fuster (2011), and the latter result is consistent with findings by Heffetz and List (2014). This raises an important question; what explains the asymmetric response to expectations in bidding behavior across experimental arms?

5.3 Aspirations and reference dependent utility: thinking about fish

Kahneman and Tversky (1979) speculate that, in addition to endowments and expectations, aspirations may matter for reference point formation. While our experimental design does not allow identification of the effect of aspirations, we tentatively probe the aspirations thesis as an explanation for the above-documented asymmetry. We believe that the certified commodity is a salient new product, which many of our subjects aspire to try out. All subjects in Arm 2 wishing to try the new certified product can increase the probability of consuming it by increasing their minimum WTA. However, there is little that subjects in group 1 can do to obtain the good—they are not endowed with it, and are extremely unlikely to obtain it via trade. The literature suggests that unrealistic aspirations can be a source of frustration and anxiety (Genicot and Ray 2020). While aspirations can be inspiring and motivational, to aspire for things that are unlikely to happen is likely to be welfare-reducing.

This implies that the asymmetric response to variation in expectations across experimental arms may be due to expectations-mediated aspirations. Specifically, while all subjects in Arm 2 (groups 3 and 4) may "safely aspire" to consume the new commodity, expectations-mediated aspirations imply that subjects from group 2 may more safely aspire to consume certified fish than subjects from group 1. This could explain why, on average, subjects

[13] A Mann-Whitney/Wilcoxon test for equality of distributions of WTA between the two treatment arms fails to reject the null hypothesis of equal distributions ($p = 0.742$). But a Mann-Whitney/Wilcoxon test for equality of WTP distributions rejects the null hypothesis of equal distributions ($p = 0.032$).

from group 2 bid more for trading up than subjects from group 1, but groups 3 and 4 demand similar levels of compensation for trading down.

Of course it is problematic, from an empirical perspective, that we do not observe aspirations directly. As a proxy, we use a survey-based measure capturing whether subject i actively thought about the certified commodity and imagined using it ($Think_i$). Consider the survey responses summarized in Table 3.

<< Insert Table 3 about here >>

From the bottom panel, not surprisingly, we learn that consumers endowed with certified fish think about how to prepare it, and are much more likely to think about cooking certified fish than uncertified fish (83% versus 20%). The responses by participants from Arm 1 are not the "mirror" outcomes of responses from subjects in Arm 2—they are qualitatively different. In fact, while participants endowed with the new product think about that product, participants endowed with the conventional product *also think about the new product* instead of their own endowment.[14]

Our new variable, $Think_i$, captures whether subjects strongly agree with the statement that they thought about how to use certified fish (similar results eventuate when we collapse

[14] While endowed with uncertified fish, these participants are as likely as participants from the other experiment to think about preparing certified fish (83% versus 82%). Surprisingly, a greater share of these consumers thought about cooking certified fish than uncertified fish (82% versus 67%). Moreover, while the great majority of consumers endowed with certified fish (strongly) disagrees with the statement that they "spend more time thinking about uncertified fish than about certified fish" (78%), we find that consumers endowed with uncertified fish also tend to disagree with this statement. The share of participants that (strongly) disagrees is even greater than the share of consumers that (strongly) agrees (45% versus 35%).

the answers to the three questions into one index variable). We regress this variable on our experimental variables and controls:

$$Think_i = \alpha + \beta Certify_i + \gamma Trade_i + \theta Trade_i Certify_i + \boldsymbol{\delta}_i \boldsymbol{x_i} + \boldsymbol{C}_c + \epsilon_i. \quad (4)$$

Based on the reasoning above, we hypothesize that being endowed with certified fish promotes thinking about it, $\beta > 0$. Subjects from group 2 expect to be able to obtain it through trade, and therefore can also aspire to consume it, hence $\gamma > 0$. Similarly, if subjects endowed with certified fish believe they may lose their endowment, then they are expected to spend less time thinking about it, hence $\theta < 0$.

Results are summarized in columns (5-6) of Table 2. As expected, and consistent with Table 3, subjects endowed with certified fish are more likely to think about it ($\beta > 0$). But we do not find that expectations matter for the direction of thinking. Subjects endowed with uncertified fish who can possibly trade up do not spend more time thinking about certified fish than their peers who are unlikely to trade ($\gamma = 0$). Expectations also do not matter for subjects endowed with certified fish ($\theta = 0$). Overall, this provides no support for expectations-mediated aspirations as an explanation for asymmetric bidding in the experiment.

6. Conclusion

We examine reference-dependent utility and the formation of reference points in the field, working with consumers trading real food items in a real market setting. An important addition to the literature is that we study trading behavior in *asymmetric commodities*—one is novel and (weakly) superior for consumers in the context we study, and therefore more salient than the other.

Consistent with existing (lab) evidence, we document that endowments matter for the construction of reference points. The literature is more divided on the role of expectations as a determinant of reference points. Our evidence is also mixed, and we document a clear

asymmetry in bidding behavior across the two commodities. While expectations with respect to future ownership matters for subjects "trading up", they are unimportant for subjects who may "trade down"—selling the superior commodity in order to obtain the conventional one. It is an open question to what extent this finding explains diverging patterns in the literature regarding the role of expectations for reference point formation—this depends on how subjects rank the items offered to them. Further research could usefully look into this.

We probe the role of aspirations to explain the asymmetric response to expectations, and conclude that there is no support for an expectations-mediated effect of aspirations. Reference point formation is a complex process affected by multiple factors. Moreover, depending on the context, the relative importance of endowments, expectations, and other factors likely varies.

Acknowledgement

This work received direct funding support from the CGIAR Research Program (CRP) on Policies, Institutions, and Markets (PIM) led by IFPRI. The authors also acknowledge support from the CGIAR Research Program on Fish Agri-food Systems (FISH). The opinions expressed are those of the authors and do not necessarily reflect those of PIM or FISH. The authors take full responsibility for errors and omissions. The authors thank Vivian Hoffmann at IFPRI for her comments on the research protocol; Shehu Latunji Akintola and Kafayat Adetoun Fakoya, both at Lagos State University, for their help during field implementation; and the team of 10 enumerators at Lagos State University for outstanding research assistance..

REFERENCES

Abeler, J., Falk, A., Götte, L., and Huffman, D., 2011. Reference points and effort provision. *American Economic Review*, 101, 470–492.

Banerji, A., and Gupta, N., 2014. Detection, identification, and estimation of loss aversion: evidence from an auction experiment. *American Economic Journal: Microeconomics*, 6(1): 91–133.

Becker, G.M., DeGroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. *Behavioral Sciences*, 9, 207–299.

Birol, E., Karandikar, B., Roy, D. and Torero, M., 2015. Information, certification and demand for food safety: Evidence from an in-store experiment in Mumbai. *Journal of Agricultural Economics*, 66(2), pp.470-491.

Bulte, E., List, J.A., and van Soest, D., 2020. Toward an understanding of the welfare effects of nudges: evidence from a field experiment in Uganda. *The Economic Journal*, 1–25 DOI: 10.1093/ej/ueaa054.

Bush, S.R., Belton, B., Hall, D., Vandergeest, P., Murray, F.J., Ponte, S., Oosterveer, P., Islam, M.S., Mol, A.P.J., Hatanaka, M., Kruijssen, T.T.T., Little, D.C., and Kusumawati, R., 2013. Certify sustainable aquaculture? *Science*, 341 (6150), 1067-1068.

Caputo, V., Lusk, J.L., and Nayga, R.M., 2019. Am I getting a good deal? Reference dependent decision making when the reference price is uncertain. *American Journal of Agricultural Economics*, 1–22, doi: 10.1093/ajae/aaz042.

Cochran, W.G. and Rubin, D. B., 1973. Controlling bias in observational studies: a review. *Sankhya: The Indian Journal of Statistics, Series A*, 35(4), 417–446.

Crawford, V., and Meng, J., 2011. 'New York city cabdrivers' labor supply revisited: reference-dependence preferences with rational expectations targets for hours and income', *American Economic Review*, 101(5), 1912–32.

Dupas, P., 2014. Short-run subsidies and long-run adoption of new health products: evidence from a field experiment. *Econometrica*, 82 (1), 197–228.

Ericson, K.M. and Fuster, A., 2011. Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *Quarterly Journal of Economics*, 126 (4), 1879–1907.

Genicot, G. and Ray, D., 2020. Aspirations and economic behavior. *Annual Review of Economics* 12: 715-746

Gill, D., and Prowse, V., 2012. A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, 102(1): 469–503.

Heffetz, O., and List, J.A., 2014. Is the endowment effect an expectations effect? *Journal of the European Economic Association*, 12(5), 1396–1422.

Ifft, J., Roland-Holst, D., and Zilberman, D., 2012. Consumer valuation of safety-labeled free-range chicken: results of a field experiment in Hanoi. *Agricultural Economics*, 43, 1–14.

Karle, H., Kirchsteiner, G., and Peitz, M., 2015. Loss aversion and consumption choice: Theory and experimental evidence. *American Economic Journal: Microeconomics* 7(2), 101-120

Knetsch, J.L., 1989. The endowment effect and evidence of nonreversible indifference curves. *American Economic Review*, 79(5), 1277–1284.

Knetsch, J.L., and Wong, W., 2009. The endowment effect and the reference state: evidence and manipulations. *Journal of Economic Behavior and Organization*, 71, 407–413.

Kopetz, C., Kruglanski, A., Arens, Z., Etkin, J., and Johnson, H., 2012. The dynamics of consumer behavior: A goal systemic perspective. *Journal of Consumer Psychology* 22: 208-223

Kőszegi, B. and Rabin, M., 2006. A Model of Reference-dependent Preferences. *Quarterly Journal of Economics*, 121 (4): 1133–65.

List, J.A., and Price, M.K., 2016. The use of field experiments in environmental and resource economics. *Review of Environmental Economics and Policy*, 10(2), 206–225.

Magnan, N., D.J. Spielman, T. Lybbert, and K. Gulati., 2015. Leveling with friends: Social networks and Indian farmers' demand for a technology with heterogeneous benefits. *Journal of Development Economics* 116: 223–251.

Simonsohn, U., and Loewenstein, G., 2006. "Mistake #37: The effect of previously encountered prices on current housing demand," *Economic Journal*, 116 (508), 175–199.

Smith, A., 2019. Lagged beliefs and reference-dependent utility. *Journal of Economic Behavior and Organization*, 167, 331-340.

Thaler, R., 1980. Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1(1), 39–60.

Tversky, A., and Kahneman, D., 1991. Loss aversion in riskless choice: a reference-dependent model. *Quarterly Journal of Economics*, 106, 1039–1061.

Wenner, L.W., 2015. Expected prices as reference points—theory and experiments. *European Economic Review*, 75, 60–79.

World Health Organization, 2015. WHO estimates of the global burden of foodborne diseases: foodborne disease burden epidemiology reference group 2007–2015.

WorldFish, 2018. WorldFish Nigeria Strategy: 2018-2022. Penang, Malaysia: WorldFish. Strategy: 2018-09.

TABLES

Table 1. Balance tests and summary statistics

Variable	Probability of being able to exchange					
	Arm 1 (WTP)			Arm 2 (WTA)		
	Low prob. (0.1)	High prob. (0.9)	Std diff	Low prob. (0.1)	High prob. (0.9)	Std diff
Age	38.82 (12.75)	38.82 (11.49)	0.02	36.80 (9.86)	37.44 (10.52)	0.06
Sex	0.89 (0.32)	0.87 (0.33)	0.06	0.74 (0.44)	0.65 (0.48)	0.20
Education	0.95 (0.21)	0.95 (0.21)	0.00	0.82 (0.39)	0.81 (0.40)	0.03
Percentage of infants in household	13.89 (16.29)	17.25 (17.14)	0.20	12.80 (15.29)	13.16 (15.17)	0.02
Salaried employment	0.03 (0.17)	0.01 (0.11)	0.14	0.01 (0.10)	0.09** (0.28)	**0.37**
Per capita monthly expenditure on food (Naira)	7,608 (5,266)	8,317 (4,934)	0.14	6,245 (4,466)	5,828 (3,388)	0.11
Household weekly per capita fish consumption	0.54 (0.59)	0.61 (0.66)	0.11	0.45 (0.38)	0.39 (0.37)	0.16
Preference for brand purchase	23.70 (7.91)	23.05 (7.80)	0.08	25.73 (7.53)	24.88 (7.02)	0.12
Trust in local market inspectors	0.65 (0.48)	0.74 (0.44)	0.20	0.75 (0.43)	0.83 (0.38)	0.20
Trust in government inspectors of food safety.	5.22 (3.34)	5.43 (3.34)	0.06	5.73 (2.34)	5.60 (2.35)	0.06
Knowledge about food safety	8.34 (3.40)	7.83 (3.23)	0.15	6.77 (2.17)	6.72 (1.62)	0.03
Number of observations	195			199		
Joint orthogonal test (*p*-value)	0.685			0.381		

Note: Standard deviations appear in parentheses. The *p*-value for the joint orthogonality test is obtained from a logit regression of the treatment variable on the covariates. Std diff means standardized difference in means. Bold figures show that the standardized difference in means is above the threshold of 0.25.

Table 2. Endowments, expectations and bidding behavior

	Dependent variable: natural log of consumer bids				Dependent variable: Think	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Certify	0.90***	0.91***	1.07***	1.10***	0.21***	0.25***
	(0.15)	(0.15)	(0.21)	(0.21)	(0.08)	(0.08)
Trade	-	-	0.42*	0.44*	0.01	0.02
			(0.24)	(0.24)	(0.06)	(0.06)
Certify x Trade	-	-	-0.39	-0.41	-0.08	-0.08
			(0.27)	(0.27)	(0.10)	(0.10)
Market fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls	No	Yes	No	Yes	No	Yes
Constant	4.37***	4.02***	4.18***	3.83	0.35***	0.06
	(0.17)	(0.41)	(0.22)	(0.43)	(0.09)	(0.14)
R-squared	0.17	0.19	0.18	0.20	0.08	0.10
Observations	394	394	394	394	394	394

Note: OLS regression estimates. Robust standard errors are reported in parentheses. Think is a dummy variable equal to one if the response to the statement, "Since the beginning of the session, I have spent some time thinking about how I would use the certified fish", is (strongly) agree. ***p<0.01, **p<0.05, *p<0.10.

Table 3. Survey responses

	(Strongly) disagree	Unsure	(Strongly) agree
Arm 1: WTP (N=195)			
S1. Since the beginning of the session, I have spent some time thinking about how I would use the certified fish;	8.2	10.3	81.6
S2. Since the beginning of the session, I have spent some time thinking about how I would use the uncertified fish;	20.5	12.8	66.6
S3. Since the beginning of the session, I have spent more time thinking about the uncertified fish than about the certified fish.	45.2	20.0	34.9
Arm 2: WTA (N=199)			
S1. Since the beginning of the session, I have spent some time thinking about how I would use the certified fish;	13.5	3.5	82.9
S2. Since the beginning of the session, I have spent some time thinking about how I would use the uncertified fish;	76.4	4.0	19.6
S3. Since the beginning of the session, I have spent more time thinking about the uncertified fish than about the certified fish.	77.9	7.0	15.1

Note: Figures reported are percentages of responses.

FIGURES

	Low probability of being allowed to trade	*High probability of being allowed to trade*
Subject received low value fish (*Arm 1*)	Group 1 (N_1=111)	Group 2 (N_2=89)
Subject received certified fish (*Arm 2*)	Group 3 (N_3=105)	Group 4 (N_4=95)

Figure 1. The factorial design of the framed field experiment

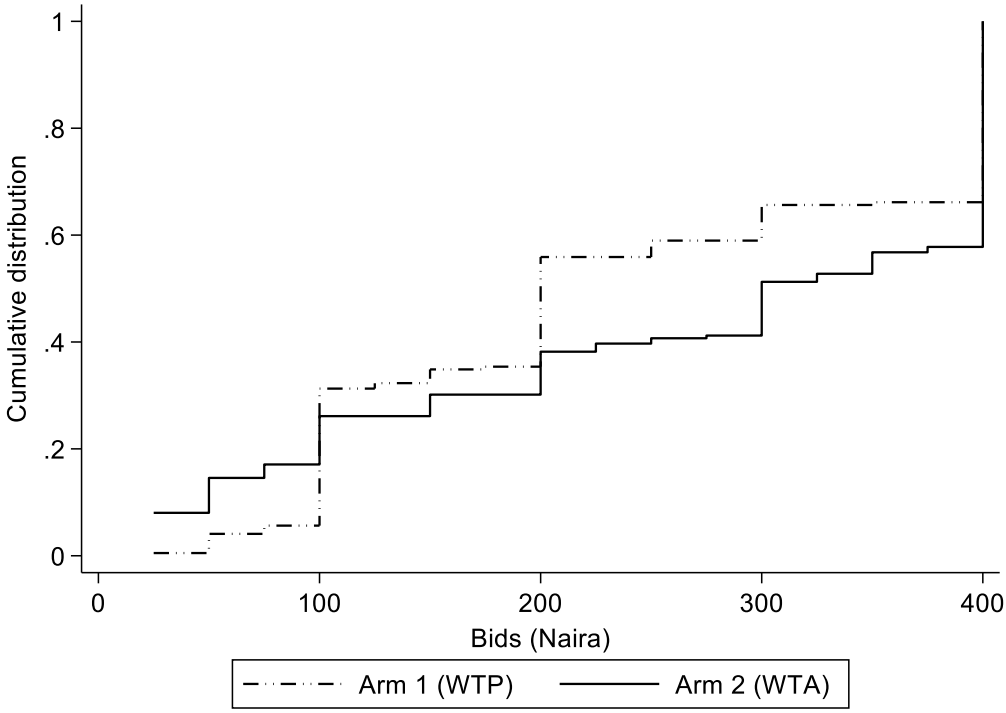

Figure 2. Empirical distribution functions of bids from Arm 1 (WTP) and Arm 2 (WTA)

Appendix

Table A.1. Summary statistics of participants in the experiment

Variable	Arm 1 (WTP)	Arm 2 (WTA)	*p*-value
Age (years)	37.10	38.69	0.466
	(10.15)	(12.17)	
Sex (1=female; 0=otherwise)	0.70	0.88	0.385
	(0.46)	(0.32)	
Education	0.81	0.95	0.235
	(0.39)	(0.21)	
Percentage of infants in household	12.97	15.39	0.342
	(15.20)	(16.72)	
Respondent participates in salaried employment	0.05	0.02	0.507
	(0.21)	(0.14)	
Per capita monthly expenditure on food (Naira)	6,000	7,900	0.152
	(3990)	(5,120)	
Household weekly per capita fish consumption	0.42	0.57	0.518
(kg)	(0.38)	(0.62)	
Preference for brand purchase	25.33	23.41	0.346
	(7.29)	(7.84)	
Trust in local market inspectors	0.79	0.69	0.417
	(0.41)	(0.46)	
Trust in government inspectors of food safety.	5.67	5.31	0.631
	(2.34)	(3.33)	
Knowledge about food safety	6.75	8.11	0.384
	(1.92)	(3.33)	
Number of observations	195	199	

Note: *p*-values are from *t*-test of difference in means, with clustering at local government area. WTA=willingness to accept; WTP=willingness to pay. In parentheses are standard deviations.

Table A.2. Splitting the sample expectations and valuation of certified fish

	Dependent variable: natural log of consumer WTA-bids		Dependent variable: natural log of consumer WTP-bids	
	(1)	(2)	(3)	(4)
Trade	0.01	0.02	0.42*	0.45*
	(0.12)	(0.12)	(0.24)	(0.24)
Market fixed effects	Yes	Yes	Yes	Yes
Additional controls	No	Yes	No	Yes
Constant	5.72***	6.13***	3.62***	2.78***
	(0.10)	(0.31)	(0.24)	(0.27)
R-squared	0.06	0.09	0.15	0.17
Observations	199	199	195	195

Note: OLS regression estimates. Robust standard errors are reported in parentheses. The dependent variables are: in columns (1 and 2), log of WTA to downgrade from certified to uncertified fish in experiment 1; in column (3 and 4), log WTP to upgrade from uncertified for certified fish in experiment 2. Asterisks indicate the following: ***p<0.01, *p<0.1