

Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior*

Keith M. Marzilli Ericson[†] and Andreas Fuster[‡]

November 18, 2009

Abstract

Evidence on loss aversion and the endowment effect suggests that people evaluate outcomes with respect to a reference point. Yet little is known about what determines reference points. This experiment shows that expectations determine reference points. We endow subjects with an item and randomize the probability they will be allowed to trade it for an alternative. Subjects that are less likely to be able to trade are more likely to choose to keep their item, as predicted when reference points are determined by expectations, but not when reference points are determined by the status quo or when preferences are reference-independent.

JEL classification: C91, D11, D81

Keywords: reference points, loss aversion, endowment effect, experimental economics

*First draft: November 12, 2009. We thank Keith Chen, Lorenz Goette, Judd Kessler, David Laibson, George Loewenstein and Matthew Rabin for helpful comments and discussions. We gratefully acknowledge support from the Russell Sage Foundation through the Small Grants Program in Behavioral Economics.

[†]Address: Department of Economics, Harvard University, Littauer Center, 1805 Cambridge St., Cambridge, MA 02138. Phone: (617) 861-7673. E-mail: kericson@fas.harvard.edu

[‡]Address: Department of Economics, Harvard University, Littauer Center, 1805 Cambridge St., Cambridge, MA 02138. Phone: (857) 928-1122. E-mail: afuster@fas.harvard.edu

1 Introduction

A wide literature suggests that people evaluate outcomes with respect to a reference point. In Kahneman and Tversky's (1979) highly influential prospect theory, individuals value gains and losses and are loss averse around a reference point. Yet the determination of the reference point was left imprecise. It has often been taken to be the status quo (Samuelson and Zeckhauser, 1988), meaning that individuals are reluctant to give up things they currently possess. However, Kőszegi and Rabin (2006, 2007, 2009) (henceforth KR) argue that reference points are determined by recent expectations about outcomes, which need not correspond to the status quo. Since individuals often expect to keep the status quo, much evidence is consistent with either method of determining reference points. This study provides new experimental evidence that reference points are shaped by expectations.

In our experiment, we manipulate expectations separately from endowments and the status quo. We endow subjects with an item (a mug), and randomize the probability that they will be allowed to exchange that item for an alternative (a pen). We show that when subjects are more likely to *expect* to keep their endowed item (because they have a low probability of being allowed to exchange), they are more likely to *choose* to keep their item if given the opportunity to exchange. This result is predicted when expectations determine reference points, but not by theories that identify the reference point with the status quo or by theories in which preferences do not depend on reference points.

Evidence on whether reference points matter and on how they are determined is necessary to choose between competing models of reference-dependent preferences and to predict when such preferences are likely to manifest themselves (e.g., when we should expect to see loss aversion).¹ Moreover, our experiment and the theory of expectation-based reference points can reconcile different findings regarding the presence or absence of "endowment effects" in exchange experiments (Knetsch, 1989; Plott and Zeiler, 2007; Knetsch and Wong, 2009). Experimental procedures can influence subjects' perceptions of the likelihood that they will be able to exchange their endowed item for an alternative or that they might also be given the alternative. As our results show, such expectations can dramatically impact subjects' propensity to exchange. Thus, experimenters testing reference-dependent preference theories should either directly induce these expectations or at least try to elicit them from the subjects, for instance through debriefing questions.²

There is little other literature directly examining the effect of expectations on reference points. Smith (2008) experimentally tests a prediction of the KR (2006) theory, namely that a higher (lagged) probability of receiving an item increases one's valuation for that item. He cannot reject

¹Recent theoretical and empirical applications of reference-dependent preferences with expectation-based reference points include Mas (2006), Heidhues and Kőszegi (2008), Crawford and Meng (2009), and Herweg, Müller, and Weinschenk (2009).

²When inducing expectations directly, experimenters have to be very careful to avoid potentially mistaken inferences by subjects. In our experiment, we transparently randomize subjects into low or high probability of getting the option to exchange, so it is clear that there is no information content in this probability. We also report the results of a pilot in which we did not transparently randomize, and subjects appeared to make value inferences from the probability with which they were allowed to exchange.

the null hypothesis that lagged expectations do not matter for valuation.³ Knetsch and Wong (2009) conduct exchange experiments using a variety of procedures, and discuss their results in terms of the KR theory. However, they do not manipulate expectations explicitly as we do, and do not directly test the theory of expectation-based reference points.⁴ Abeler, Falk, Götte, and Huffman (2009) conduct a real-effort experiment in which subjects are equally likely to be paid a piece rate or to receive a fixed payment, and only learn which case applies after they stop working. When the amount of the fixed payment is raised, subjects work more, consistent with KR and related theories. Their study therefore provides evidence for an effect of expectation-based reference points on labor supply, and is thus complementary to our experiment which instead focuses on a simple exchange situation.

2 Theoretical Framework

In this section, we compare how expectations affect the choices of three types of individuals: an individual with classical preferences (no loss aversion), one with loss aversion around a reference point given by the status quo, and one with loss aversion around a reference point determined by expectations (as in KR 2006).⁵ Consider an individual who is endowed with a mug and expects to have the option to exchange this mug for a pen with probability p . After being given some time to think about the situation, she is then asked to register a decision: if the option to exchange materializes, would she like to exchange?

Let individuals have utility functions $u(c|r)$ that can depend both on consumption c and reference r . Consumption and reference levels have multiple dimensions (a “mug” dimension and a “pen” dimension in this application), as in the KR extension of prospect theory. Utility on the k^{th} dimension is composed of direct consumption utility u_k and gain-loss utility with respect to that dimension’s reference level of utility u_k^r . Total utility is as follows:

$$u(c|r) = \sum_{k \in \{\text{mug}, \text{pen}\}} \left\{ \underbrace{u_k}_{\text{Consumption utility}} + \underbrace{\mu(u_k - u_k^r)}_{\text{Gain-loss utility}} \right\},$$

where μ has the properties of the Kahneman-Tversky value function. For simplicity, we follow Section IV of KR (2006) and assume μ to be a piecewise linear function with a kink at zero that captures loss aversion: $\mu(x) = \eta x$ for $x > 0$ and $\mu(x) = \eta \lambda x$ for $x \leq 0$, where $\eta \geq 0$ is the weight on gain-loss utility and $\lambda > 1$ is the individual’s loss-aversion coefficient. This specification nests classical preferences that do not feature gain-loss utility ($\eta = 0$). Throughout, when referring to individuals with reference-dependent preferences, we assume $\eta > 0$.

³However, his standard errors are large enough that he would not have had the power to detect reasonably sized effects (e.g. a \$1, or 15-25%, change in value). Another potential reason why he finds no effect may be that valuations are elicited after subjects learn whether or not they receive the item, so that the reference point may already have adjusted to having the item (or not having it) for sure.

⁴Section 5 will discuss their experiments in more detail.

⁵A more detailed exposition is provided in the Theory Appendix.

Now we consider two different specifications of the reference point. If the reference point is given by the status quo or by endowments, then we simply have $u_{pen}^r = 0$ and $u_{mug}^r = u_{mug}$, since the individual owns the mug.

In contrast, in KR’s specification, the reference point is determined by the individual’s probabilistic beliefs about outcomes, or in other words, her expectations. KR assume that individuals calculate an outcome’s gain-loss utility as follows: they compare each dimension’s consumption utility from the actual outcome to that of each possible outcome, producing a series of gain-loss $\mu(\cdot)$ terms. Each possible outcome’s $\mu(\cdot)$ is then weighted by the probability with which the individual expected that outcome to occur.⁶ Furthermore, an individual’s reference point is *endogenous* to her plans, as the probabilities of different states of the world occurring may depend at least in part on her decisions.

In our example, if the KR individual plans not to exchange, then her reference point is the same as in the status quo case — she keeps the mug for sure. If instead she decides to exchange conditional on having the possibility to do so, the two states of the world she considers for her gain-loss utility are $r_{end\ w/pen} = \{1\ \text{pen}, 0\ \text{mugs}\}$ and $r_{end\ w/mug} = \{0\ \text{pens}, 1\ \text{mug}\}$, which occur with probability p and $1 - p$, respectively. Her utility of consumption bundle c given her expectations is then $pu(c|r_{end\ w/pen}) + (1 - p)u(c|r_{end\ w/mug})$. If the exchange occurs, then she gains or loses nothing with respect to $r_{end\ w/pen}$, but with respect to $r_{end\ w/mug}$, she gains u_{pen} and loses u_{mug} . Thus, her total utility of this outcome would be her consumption utility u_{pen} plus her gain-loss utility $(1 - p)\eta(u_{pen} - \lambda u_{mug})$. Similarly, if the exchange does not occur, her total utility would be $u_{mug} + p\eta(u_{mug} - \lambda u_{pen})$. Following KR, we assume that the individual acts so as to maximize her expected utility given her reference point (which itself is determined by her plan), and chooses the plan that leads to the highest expected utility, given her rationally-forecasted future actions.⁷

Now, let individuals vary in their value for the mug and for the pen; assume that (u_{mug}, u_{pen}) are distributed according to some population distribution function. The following proposition shows that increasing the probability p with which an individual is allowed to exchange increases the expected proportion of individuals who choose to exchange when reference points are determined by expectations, but not when reference points are given by the status quo or when individuals do not have gain-loss utility.

Proposition 1 *If reference points are given by the status quo, or if there is no gain-loss utility ($\eta = 0$), then the probability p of having the option to exchange does not affect the decision to exchange. If reference points are determined by expectations, then increasing p increases the expected proportion of individuals who choose to exchange.*

Proof. When reference points are given by the status quo, the utility of exchanging is greater

⁶In the application considered in this paper, no predictions would change if reference points were instead simply given by the expected consumption utility on each dimension.

⁷In the language of KR, we assume throughout the main part of the paper that the individual plays the “preferred personal equilibrium.” However, the theoretical prediction on the effect of p when using the less restrictive “personal equilibrium” concept is the same. See the Theory Appendix for more details.

than the utility of keeping the mug if and only if $[u_{pen} + \eta(u_{pen} - \lambda u_{mug})] > u_{mug}$, which does not depend on p . When $\eta = 0$, the individual desires to exchange if and only if $u_{pen} > u_{mug}$.

However, when reference points are determined by expectations, an individual desires to exchange if and only if the expected utility of planning to exchange (and following through on the plan) is greater than the expected utility of planning to keep the mug (and keeping it). This is true when $(u_{pen} - u_{mug}) + \eta(1 - p)(1 - \lambda)(u_{pen} + u_{mug}) > 0$. Note that $u_{pen} > u_{mug}$ is a necessary condition for the individual to desire to exchange, as for $p < 1$ the second term is negative due to loss aversion. The second term gets less negative as p increases. Therefore, an individual who chooses to keep the mug under one value of p will always do so for lower values of p , but may instead choose to exchange for higher values of p . ■

Intuitively, for an individual who has classical preferences, the decision of whether to exchange is influenced only by the consumption utilities she derives from the two items. On the other hand, an individual who is loss-averse around the status quo may not be willing to give up the mug he owns for the pen, even if the pen has higher consumption utility, due to the loss he feels from moving away from the status quo. However, this loss, and therefore the decision, is independent of whether the chance of the exchange occurring is high or low.

Meanwhile, for an individual with KR-style preferences, the cost imposed by the gain-loss utility in case she decides to exchange must be outweighed by the gain in consumption utility from getting the pen instead of the mug, and this is less likely to happen the lower is p . On the other hand, as $p \rightarrow 1$, gain-loss utility becomes relatively less important. If such an individual were certain to be able to exchange, she would prefer to do so if and only if the consumption utility from the pen exceeds the one from the mug, exactly like a classical individual.

One underlying psychological mechanism that may be driving the effect of p is that a higher probability of ending up with an item increases the time or intensity with which an individual thinks about the item, and leads her to become attached to the item.⁸ Though the probability of ending up with an item is to some extent endogenous to her plans, by lowering the probability of being able to exchange, we “force” the individual to think about the item she is endowed with. This in turn increases the chance that the reluctance to lose this “endowment” created by expectations outweighs the potential consumption utility surplus the alternative item may provide.

3 Experimental Design

In this section, we describe the design of our experiment and the treatment conditions we use to test our hypothesis, based on Proposition 1, that increasing the probability that subjects have the option to exchange increases the proportion that choose to exchange, as predicted when reference points are determined by expectations. Our experiment is a variation of the classic Knetsch (1989) experiment,

⁸This is related, for instance, to the “emotional attachment mechanism” discussed by Ariely, Huber, and Wertenbroch (2005).

with many procedural details following Plott and Zeiler’s (2007) “loss emphasis treatment.”⁹ We endow subjects with an item (a Harvard university travel mug) and randomly and transparently manipulate the individual probability with which a subject will have the opportunity to exchange her item for an alternative (a silver metal Harvard university pen).¹⁰

Upon arrival at the lab, each subject is seated at a carrel with a mug and a pen on it. We then flip a coin in front of each subject individually. The coin’s sides are labeled “1” and “9”, and we give each subject an index card with their resulting number on it. Subjects then start reading instructions on the computer screen in front of them. Subjects are told that they own the mug in front of them, and that each participant will leave the experiment with either a mug or a pen. They are then informed that at the end of the session, they may have the option to exchange their mug for the pen, if they so desire, and that this option will occur if and only if a ten-sided die that we roll individually for each subject at the end of the study comes up lower than or equal to the number on their index card. Thus, subjects whose coin came up “1” have a 10% chance of having the option to exchange (we will refer to them as being in *Treatment L*, for low probability) while subjects who got a “9” have a 90% chance of being able to exchange their mug for the pen if they would like (*Treatment H*).

After this first round of instructions, and after an experimenter reads the most important parts out loud, subjects answer some demographic questions and then fill out the first part of a 44-item “Big Five” personality questionnaire (John, Donahue, and Kentle, 1991). As usual, the goal of such a “filler” questionnaire is to distract subjects from the main decision we are interested in, and also to provide them with some time to plan their decision as to whether or not to exchange the mug. After they finish answering the first 22 questions, we remind them of the instructions and procedures for the (possible) mug-pen exchange, in order to make sure they understand, and also to make them think about their choice. Then, after they answer the second 22 questions of the Big Five questionnaire, subjects are asked to make a choice *conditional* on the die coming up lower than or equal to the number on their index card.¹¹ This allows us to observe a decision for each subject, not just the ones for which the decision actually turns out to apply, which is similar to the strategy method often used in experimental games.¹² Before rolling the die, we ask some additional questions, as described in the next section.

The experiments were run at a lab at the Harvard Decision Sciences Lab. A total of 45 subjects (23 females; mean age 21), all of them undergraduate students at Harvard, participated in our main sessions. We conducted 10 sessions with between 3 and 7 subjects per session. Half the sessions

⁹More details are provided in a Methods Appendix available from the authors.

¹⁰Both items have a retail value of around \$8. In a usual “endowment effect” experiment, half the subjects would be endowed with a mug and the other half with a pen. We chose to endow all our subjects with the same item (which we randomly picked by flipping a coin before our very first session) in order to get sufficient power for testing our hypothesis of interest.

¹¹They are asked to check one of the following: If the die comes up [1 in treatment L; 1 to 9 in treatment H]:
 I want to keep the mug; or I want to trade the mug for the pen.

¹²As is further discussed in the Theory Appendix, eliciting a subject’s decision conditional on a choice set being reached is theoretically equivalent to asking the subject to choose once the choice set is reached, if by then the subject has made her plans (which determine her reference point).

were run on one day in late October 2009, the other half over two days in early November. Subjects received a show-up fee of \$10, and the experiment took about 20 minutes.¹³

4 Results

4.1 Expectations Affect Exchange Behavior

Our main result shows that the proportion of subjects who indicate that they want to exchange their mug for the pen is significantly higher in the treatment where an exchange, if desired, is more likely to occur:

RESULT 1: Subjects that have a 10% chance of being able to exchange are significantly less likely to be willing to exchange than subjects that have a 90% chance. The proportion of subjects who choose to exchange in Treatments L and H are 22.7% and 56.5%, respectively ($p=0.021$, two-sided two-sample test of proportions).

Thus, we confirm the prediction of KR (2006) that reference points matter for choice, and that they are determined by expectations. The effect is large: subjects in Treatment H are 34 percentage points, or one and a half times, more likely to choose to exchange than subjects in Treatment L. As a test of the robustness of the result, Table 1 reports the estimated marginal effects from probit regressions that predict the probability a subject chooses to exchange from a treatment indicator and other covariates that may be related to their choice. Among the covariates we consider, gender significantly predicts the desire to exchange: column (2) shows that females seem to like the pen relatively more than males. However, the gender effect does not drive our result, as our treatments were perfectly gender-balanced.¹⁴ Column (3) shows that when both treatment and gender indicators are included as regressors, the indicator for Treatment H is still significant at $p < 2\%$. Furthermore, column (4) shows that adding subject age and an indicator for the month in which the session took place does not change the result either. If anything, the estimated treatment effect becomes even larger.¹⁵

4.2 Expectations and Subject Evaluations

RESULT 2: According to questionnaire responses, subjects in treatment L like the mug more than subjects in Treatment H, and more strongly agree with the statement that they spent more time thinking about the mug than about the pen.

¹³In the November sessions, the subjects afterwards participated in a second, completely unrelated experiment. The fact that these experiments were unrelated was made very clear to participants, and they were told that the first experiment involved the mugs and pens while the second involved choosing between various payment amounts. The results below demonstrate that behavior was similar in the October and November sessions.

¹⁴We had 11 males and females each in Treatment L, and 11 males and 12 females in Treatment H.

¹⁵The robustness of the effect across subsets of sessions is surprisingly strong: the proportions of subjects in treatments L and H who wanted to exchange were 23.1% and 53.9% in our October sessions (13 subjects per treatment) and 22.2% and 60.0% in our November sessions (9 subjects in treatment L, 10 subjects in treatment H).

Before rolling the die to determine whether a subject’s decision applies, we asked the subjects to indicate agreement or disagreement with the following statements (on a scale from 1 (disagree strongly) to 5 (agree strongly)):

1. I like the mug better than the pen.
2. Since the beginning of the session, I have spent some time thinking about how I would use the pen.
3. Since the beginning of the session, I have spent some time thinking about how I would use the mug.
4. Since the beginning of the session, I have spent more time thinking about the mug than about the pen.

Summary statistics on the responses by treatment are provided in Table 2. The table also provides p-values from two-sided Mann-Whitney tests of the null hypothesis of equal distributions of responses across treatments. The table shows that subjects in Treatment L (mildly) significantly more strongly agree with the statement that they like the mug better than the pen, which is of course consistent with the behavioral evidence. Subjects in both conditions are about equally likely to agree to have spent *some* time thinking about using each item. However, subjects in Treatment L more strongly agree to having spent *more* time (between the moment we explained the decision situation and the moment they made their decision) thinking about the mug than about the pen. This latter fact is consistent with a potential psychological mechanism through which the reference point may be formed: the more likely you are to receive something, the more time you spend thinking about it, which increases your “attachment” and would generate the feeling of a loss in case you do not receive it.

In the debriefing questionnaire, we ask subjects which item they think has the higher retail value, on a scale from “definitely pen” (1) to “definitely mug” (5). The final row of Table 2 shows that the answers do not differ significantly across treatments, suggesting that our explicitly random assignment to treatments prevented subjects from inferring anything about the relative values of the two items.¹⁶

4.3 A Cautionary Tale: Results from a Pilot

Before settling on the design reported in the previous sections, we ran sessions in which we did not make the random assignment to treatments obvious to subjects.

¹⁶Before rolling the die, we also asked subjects how much they would be willing to pay for the pen (hypothetically). For subjects that chose to keep the mug, the question read “Suppose you had the chance to buy the pen. What is the maximum you would be willing to pay for the pen?” For subjects that chose to exchange, the question was instead “Suppose the dice roll comes up so that you must keep your mug, but you also had the chance to buy the pen. What is the maximum you would be willing to pay for the pen?” Surprisingly, subjects who chose to keep the mug indicate a mildly significantly higher willingness-to-pay for the pen than subjects who would like to receive the pen instead of the mug (means: \$2.25 and \$0.99; two-sided t-test $p=0.086$). However, there are no significant differences across treatments.

A total of 63 subjects participated (36 undergraduates, 23 summer school students, four graduate students, all at Harvard; mean age 21.6) in 15 sessions conducted at the end of July 2009. Of the subjects, 32 were randomly assigned (without their knowledge) to Treatment L' and 31 to Treatment H'. Subjects in these treatments had the same probabilities of being able to exchange as in L and H, respectively, and received very similar instructions, except that they were not told the source of the probability they would be permitted to exchange (see Methods Appendix for more details).

The results were the following: in Treatment H', 29.0% of subjects chose to exchange, while in Treatment L', 62.5% chose to exchange, and the difference in proportions is statistically significant at $p < 0.01$ (two-sided two-sample test of proportions). Thus, the results went in the opposite direction of the ones in our main sessions reported above, and strongly contradicted our hypothesis that subjects with a higher probability of being able to exchange should be more willing to do so. After the first few sessions, we realized that this may have been due to a "value inference" effect: subjects who were given a low (10%) probability of being able to exchange their mug for the pen may have inferred that the pen must therefore be more valuable. We then added the question to the debriefing survey in which we ask subjects which item they believe has higher retail value. Consistent with the value inference hypothesis, among the 19 subjects in Treatment L' who answered the question, 13 (or 68.4%) indicated that they believed that the pen "definitely" or "probably" had higher retail value, while the same was true for only four out of 14 (28.6%) of subjects in Treatment H'. A two-sided Mann-Whitney test indicates that the distributions of responses are significantly different at $p < 0.03$.

Thus, we conclude that, in the pilot, our test of expectation-based reference points was confounded by an experimental design that created subject misconceptions about the relative values of the two items. However, the pilot may also speak to the relative "strength" (or lack thereof) of expectation-based reference points. While we find evidence in favor of the theory in our clean treatments, the effect of loss aversion can be more than outweighed by (perceived) value signals provided by probabilities. Our results therefore also provide evidence on how subtle cues in the experimental environment can have a strong influence on subjects' behavior.

5 Discussion of Related Experiments

In this section, we discuss how our findings relate to previous experimental evidence on the presence or absence of the endowment effect, and show that interpreting previous studies in terms of the KR theory can reconcile their varied findings.

The classic early experiment demonstrating the endowment effect was conducted by Knetsch (1989), who found that only about 10 percent of subjects that he endowed with either a mug or a chocolate bar were subsequently willing to give up their assigned item in exchange for the other item, while standard theory predicts that 50 percent should exchange. Such exchange asymmetries have usually been interpreted as resulting from loss aversion around a reference point given by

current endowments. However, a natural interpretation in terms of the KR theory is that, until the opportunity to exchange is offered to participants, they fully expect to leave the experiment with the item they were endowed with, so that expectations and endowments coincide.

Plott and Zeiler (2007) (henceforth PZ) argue that Knetsch's findings (and those of other researchers who replicated Knetsch's experiment) were largely driven by certain features of his experimental procedures. PZ alter Knetsch's procedures in various ways and demonstrate that such changes can have a large impact on the existence and magnitude of exchange asymmetries. They interpret this result as evidence that the endowment effect observed in earlier studies is not due to non-standard features of subjects' preferences.

However, when experimental procedures are altered, subjects' expectations of which item(s) they will leave the experiment with may also change. In particular, PZ's procedures may have made subjects believe that they would also be given a pen, or that there would be an opportunity to exchange the mug for the pen. Their two treatments in which no exchange asymmetries are observed, the "full set of controls" and the "loss emphasis treatment," begin as follows (p. 1459):

We began these sessions by informing the subjects that mugs and pens would be used during the experiment. Subjects were then told that a coin was flipped before the start of the experiment to determine which good, the mug or the pen, would be distributed first. We then distributed mugs to the subjects and announced, 'These mugs are yours.'¹⁷

In terms of the KR theory, the reference point of a subject who expects to receive both a mug and a pen is u_{mug} in the mug dimension and u_{pen} in the pen dimension. When she then gets surprised by the announcement that to obtain a pen, she has to give up her mug, she necessarily feels a loss in one dimension, and therefore chooses the item with the higher consumption utility. Similarly, a subject who expects to be able to exchange for sure will optimally plan to do so if the consumption utility of the pen is higher than the one provided by the mug, and therefore will not exhibit behavior predicted by theories that take the status quo as the reference point.

Thus, while PZ show that the endowment effect can disappear when certain experimental procedures are used, they do not directly manipulate or measure subjects' expectations that they will leave with a mug or pen. While they argue that previously observed exchange asymmetries are not due to reference-dependence but rather classical preferences interacting with experimental procedures, their results are also consistent with the endowment effect being driven by reference-dependent preferences, but with a reference point determined by expectations, not current endowments.

¹⁷In their "loss emphasis treatment," PZ use somewhat stronger language to convey subjects' entitlement to the endowed good; they say "The mug is yours. You own it." In footnote 13 (p. 1454), PZ give another account of the specific language they use at the beginning of the experiment: "Specifically, subjects were told, 'Before the start of the experiment, a coin was flipped to determine which good, the mug or the pen, to distribute. It came up heads, which means that we start with the mugs (or pens). The subjects in the other room will start with the pens (or mugs).'"

Other work has found an “endowment effect” even when subjects do not formally own an item. Knetsch and Wong (2009) find exchange asymmetries in an experimental treatment in which they give subjects an item, but tell them they do not yet own the item (but will own it at the end of the session). They interpret this finding as evidence in favor of KR, and interpret the PZ procedures as providing only a “weak reference state.”¹⁸ However, an alternative explanation for their finding is that it may be driven by subjects inferring something about the relative values of the items, as in that particular treatment it is not revealed to the subjects that the item they received first was chosen randomly. As emphasized by PZ (and also confirmed by the findings from our pilot experiment reported in section 4.3), such seemingly innocuous procedural details can have a major impact on exchange behavior. Nevertheless, the basic idea and the findings of their experiment are certainly consistent with KR.

6 Conclusion

In a simple decision context (keep or exchange an item), increasing subjects’ expectations of leaving with an item makes them less willing to exchange this item for an alternative. We have provided evidence that individual’s reference points are determined, at least in part, by expectations. Merely manipulating the probability that subjects will have the opportunity to exchange, and explicitly randomizing this probability so subjects know there is no information content in it, yields a strong effect on exchange behavior.

Experimental and empirical findings of individuals being reluctant to part with items they currently own has traditionally been interpreted as evidence of individuals being loss averse around the status quo. However, endowments often correspond to expectations, so that an interpretation of these exchange asymmetries in terms of expectation-based reference points is equally compelling. Furthermore, expectation-based reference points explain why some studies can endow subjects with an item, yet not observe an endowment effect, while other studies observe an “endowment effect” despite never having endowed subjects with an item. Future work should assess whether endowments also directly affect reference points (or some other feature of preferences) independently of expectations, though our interpretation of Plott and Zeiler (2007) suggests that they do not. Our experiment provides clean evidence that expectations shape the reference point, and indicates that the theory of Kőszegi and Rabin (2006) may be useful for economists who wish to model reference-dependent preferences in economic applications.

¹⁸In this treatment, there is no upfront announcement to subjects that they will be able to alternatively get the other item at the end. In another treatment, Knetsch and Wong closely follow PZ’s procedures, except that they explicitly tell subjects that at the end they will have the option to exchange, and confirm the absence of an endowment effect under these conditions. They do not discuss that the announcement of having the possibility to exchange for sure will lead KR to predict no endowment effect.

References

- ABELER, J., A. FALK, L. GÖTTE, AND D. HUFFMAN (2009): “Reference Points and Effort Provision,” *American Economic Review*, forthcoming.
- ARIELY, D., J. HUBER, AND K. WERTENBROCH (2005): “When do losses loom larger than gains?” *Journal of Marketing Research*, 42, 134–138.
- CRAWFORD, V. P. AND J. MENG (2009): “New York City Cabdrivers’ Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income,” <http://weber.ucsd.edu/~vcrawfor/cabdriver11.3.pdf>.
- HEIDHUES, P. AND B. KŐSZEGI (2008): “Competition and price variation when consumers are loss averse,” *American Economic Review*, 98, 1245–1268.
- HERWEG, F., D. MÜLLER, AND P. WEINSCHENK (2009): “Binary Payment Schemes: Moral Hazard and Loss Aversion,” *American Economic Review*, forthcoming.
- JOHN, O. P., E. M. DONAHUE, AND R. L. KENTLE (1991): “The Big Five Inventory—Versions 4a and 54,” *Berkeley: University of California, Berkeley, Institute of Personality and Social Research*.
- KAHNEMAN, D. AND A. TVERSKY (1979): “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 47, 263–292.
- KŐSZEGI, B. AND M. RABIN (2006): “A Model of Reference-Dependent Preferences,” *Quarterly Journal of Economics*, 121, 1133–1165.
- (2007): “Reference-Dependent Risk Attitudes,” *American Economic Review*, 97, 1047–1073.
- (2009): “Reference-Dependent Consumption Plans,” *American Economic Review*, 99, 909–936.
- KNETSCH, J. L. (1989): “The Endowment Effect and Evidence of Nonreversible Indifference Curves,” *American Economic Review*, 79, 1277–1284.
- KNETSCH, J. L. AND W.-K. WONG (2009): “The endowment effect and the reference state: Evidence and manipulations,” *Journal of Economic Behavior and Organization*, 71, 407–413.
- MAS, A. (2006): “Pay, Reference Points, and Police Performance,” *Quarterly Journal of Economics*, 121, 783–821.
- PLOTT, C. R. AND K. ZEILER (2007): “Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory?” *American Economic Review*, 97, 1449–1466.
- SAMUELSON, W. AND R. ZECKHAUSER (1988): “Status quo bias in decision making,” *Journal of Risk and Uncertainty*, 1, 7–59.

SMITH, A. (2008): “Lagged Beliefs and Reference-Dependent Utility,” *University of Arizona Working Paper 08-03*.

Table 1: Determinants of Desire to Exchange Mug for Pen

Pr(choose to exchange)	(1)	(2)	(3)	(4)
Treatment H	.338** (.137)		.353** (.139)	.399** (.147)
Female		.249* (.141)	.270* (.145)	.289* (.160)
Age				.102 (.066)
November Session				-.088 (.167)
# observations	45	45	45	45

Notes: Displayed coefficients are predicted marginal effects from probit regressions (in case of dummy variables, for a discrete change from 0 to 1). Dependent variable: Indicator variable = 1 if subject wants to exchange. Standard errors in parentheses.

Level of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Subject Evaluations of the Mugs and Pens

<i>Question</i>	<i>Treatment</i>		p-value
	L	H	
1) "I like the mug better than the pen."	3.95 (1.17)	3.26 (1.25)	.061
2) "Since the beginning of the session, I have spent some time thinking about how I would use the pen."	2.95 (1.46)	3.35 (1.23)	.366
3) "Since the beginning of the session, I have spent some time thinking about how I would use the mug."	4.09 (0.87)	4.00 (1.09)	.960
4) "Since the beginning of the session, I have spent more time thinking about the mug than about the pen."	3.95 (1.13)	3.17 (1.40)	.056
5) "Which item do you think has higher retail value?"	3.50 (1.14)	3.26 (1.01)	.387

Notes: Questions 1) to 4) were asked after subjects made their decision as to whether they would like to trade, and are answered on a scale from 1 (strongly disagree) to 5 (strongly agree). Question 5) was asked as part of the debriefing questionnaire, and is answered on the following scale (with the corresponding number in brackets): "Definitely Pen" (1), "Probably Pen" (2), "About the same" (3), "Probably Mug" (4), "Definitely Mug" (5). Standard deviations in parentheses. p-values are from a two-sided nonparametric Wilcoxon-Mann-Whitney test.

A Appendix: Theory

In this appendix, we describe in more detail the model of Kőszegi and Rabin (2006) and confirm its predictions in our experiment.

A.1 General Theory

The main objective of KR (2006) is to extend Kahneman and Tversky’s (1979) prospect theory and to bring it closer to “standard” economic theory. In their formulation, the utility of a decision-maker (henceforth DM) depends both on her consumption c and on how this consumption compares to a reference level r . They assume separability between “consumption utility” $m(c)$ and “gain-loss utility” $n(c|r)$. Consumption and reference levels can have multiple dimensions, so that $c = (c_1, c_2, \dots, c_K)$ and $r = (r_1, r_2, \dots, r_K)$, and it is assumed that both forms of utility are additively separable across dimensions, such that $m(c) = \sum_{k=1}^K m_k(c_k)$ and $n(c|r) = \sum_{k=1}^K n_k(c_k|r_k)$. Total reference-dependent utility (henceforth RDU) of an outcome c given reference level r is then given by

$$u(c|r) = \underbrace{\sum_{k=1}^K m_k(c_k)}_{\text{cons. utility } m(c)} + \underbrace{\sum_{k=1}^K \mu(m_k(c_k) - m_k(r_k))}_{\text{gain-loss utility } n(c|r)},$$

where $\mu(\cdot)$ is continuous and strictly increasing, and has the properties of the Kahneman-Tversky value function: loss aversion (a kink at zero) and diminishing sensitivity. In what follows, we will use the piecewise linear specification (as in section IV of KR, 2006) $\mu(x) = \eta x$ for $x > 0$ and $\mu(x) = \eta\lambda x$ for $x \leq 0$, where $\eta > 0$ is the DM’s weight on gain-loss utility and $\lambda > 1$ her loss-aversion coefficient.¹⁹

In KR, the reference point is based on expectations of future outcomes, so that it is in fact a reference lottery given by probability measure $G : \mathbb{R}^K \rightarrow \mathbb{R}$ over consumption utility in each of the k dimensions. Thus the RDU of outcome c given expectations G is

$$U(c|G) = \int u(c|r)dG(r).$$

This formulation implies that when evaluating a dimension c_k of a consumption outcome c , the DM compares it separately to each possible value that this dimension could take according to the reference lottery, and weights it by the probability of this outcome in the reference lottery.

But what exactly determines the reference lottery? KR employ the concept of *personal equilibrium* (PE), which requires the following:²⁰

- the DM has a plan for every contingency that she can possibly face in a given decision

¹⁹This is not restrictive in our situation with binary outcomes.

²⁰Kőszegi and Rabin (2007) refer to this equilibrium concept, which applies in situations in which decisions are only made once the reference lottery is fixed by expectations (as in our application), as “unacclimating personal equilibrium.” In contrast, if the DM could determine the reference lottery by her choice, the “choice-acclimating personal equilibrium” concept would apply.

situation;

- her reference lottery is based on her expectation to put these plans into action once she knows which contingency applies; and
- her actions maximize expected RDU, given her reference lottery.

The timing of a decision situation in which PE applies is as follows:

- $t = -1$: DM starts focusing on the decision; knows all the possible choice sets she might face at $t = 1$ and the probabilities with which they occur;
- $t = 0$: DM makes plans for each possible choice set, thereby setting her reference lottery;
- $t = 1$: ‘Nature’ determines the choice set the DM actually faces. The DM follows through on her plan for this choice set, which then yields either a deterministic or a probabilistic outcome.

The PE concept requires the DM’s actions at $t = 1$ to be optimal given her beliefs at $t = 0$, beliefs formed at $t = 0$ to be rational given her plans, and her plans to be “credible” in the sense that at $t = 1$ the DM has no incentive to deviate from them (akin to subgame perfection in standard multi-player game theory).

KR emphasize that it is possible for multiple PEs to exist in a given situation. An example: assume the consumption utilities resulting from actions x and y are not too different. Then, one equilibrium may be for the DM to plan on x and indeed do x , but another one may be for her to plan on y and do y (because if she planned on y and ended up doing x , she would feel a loss from not getting the consequence of y , which may outweigh a possible consumption utility advantage of x). In such a case KR’s preferred way of equilibrium selection is to take the PE that gives the DM the highest expected RDU, the *preferred personal equilibrium* (PPE).

A.2 Derivation of KR Portion of Proposition 1

Denote subject i ’s consumption utility from the mug by u_{mug} and the consumption utility from the pen by u_{pen} .²¹ The treatment variable is p , the probability with which the subject will be able to exchange the mug (which she is endowed with) against the pen. Each subject will have to indicate whether she wants to exchange the mug for the pen (but her decision only applies with probability p ; with probability $1 - p$, she keeps the mug independently of her decision). We will refer to her two possible actions as “exchange” and “keep.”

The subjects know the choice sets that they may be facing, but are unable to commit to a decision until the end of the session. In terms of the timeline from the previous subsection, the moment when we announce to subjects that they own the mug and may have the possibility to

²¹Throughout, we assume $u_{pen}, u_{mug} > 0$, i.e. that subjects prefer both receiving the mug for sure and receiving the pen for sure to receiving nothing.

exchange the mug against a pen comprises $t = -1$. $t = 0$ is then the timespan during which the subjects plan what they are going to do at $t = 1$, and this plan determines their reference lottery. The moment when they make their decision is $t = 1$. Note that we elicit a decision from all subjects, independently of whether their decision to exchange or not will in fact matter. In terms of the KR theory, such a situation (occurring at $t = 1$) is no different from the alternative in which we first roll the die and then only elicit a decision if the appropriate number comes up. At this point, the plan is made, and whether we elicit it as a plan or as an actual choice should not make a difference.

We now examine how the conditions on u_{mug} and u_{pen} for a subject to choose “exchange” in a PE or (more restrictively) the PPE vary with our treatment variable p .²²

If a subject plans to choose “exchange” if given the choice, then she expects $r_{end\ w/pen}$ to occur with probability p and $r_{end\ w/mug}$ to occur with probability $1 - p$. Her utility of outcome c , given her expectations, is then

$$U(c|\text{“exchange”}) = pu(c|r_{end\ w/pen}) + (1 - p)u(c|r_{end\ w/mug}).$$

If she follows through with her plan, her gain-loss utility in the different outcome cases is as follows:

- if her decision applies (which happens with probability p): $p \cdot 0 + (1 - p)\eta(u_{pen} - \lambda u_{mug})$
- if it does not apply (which happens with probability $1 - p$): $p\eta(u_{mug} - \lambda u_{pen}) + (1 - p) \cdot 0$

Therefore, her expected utility from sticking to her plan to say “exchange” is given by²³

$$\begin{aligned} EU(\text{“exchange”}|\text{“exchange”}) &= p[U(\{1\ \text{pen}, 0\ \text{mugs}\}|\text{“exchange”})] \\ &\quad + (1 - p)[U(\{0\ \text{pens}, 1\ \text{mug}\}|\text{“exchange”})] \\ &= p[u_{pen} + (1 - p)\eta(u_{pen} - \lambda u_{mug})] \\ &\quad + (1 - p)[u_{mug} + p\eta(u_{mug} - \lambda u_{pen})] \end{aligned}$$

Meanwhile, if having planned to exchange, she deviates and chooses to say “keep” instead, her gain-loss utility would equal $p\eta(u_{mug} - \lambda u_{pen})$ for sure, and her total expected utility would be equal to

$$EU(\text{“keep”}|\text{“exchange”}) = u_{mug} + p\eta(u_{mug} - \lambda u_{pen}).$$

Choosing to exchange is then a PE if and only if $EU(\text{“exchange”}|\text{“exchange”}) \geq EU(\text{“keep”}|\text{“exchange”})$, which boils down to the condition

$$u_{pen} \geq u_{mug} \underbrace{\frac{1 + \eta(\lambda + p(1 - \lambda))}{1 + \eta(1 - p(1 - \lambda))}}_{\equiv X(p)}$$

²²We limit our focus to pure-strategy equilibria.

²³The notation $U(\text{“}x\text{”}|\text{“}y\text{”})$ means “utility from action x if the reference lottery is (the result of plan) y .”

As $\lambda > 1$ and $\eta > 0$ we have that

$$X'(p) < 0, \quad X(0) = \frac{1 + \eta\lambda}{1 + \eta} > 1, \quad X\left(\frac{1}{2}\right) = 1, \quad \text{and} \quad X(1) = \frac{1 + \eta}{1 + \eta\lambda} < 1.$$

In words, this means that if the probability p of getting the possibility to exchange the pen for the mug is low, the consumption utility a subject gets from the pen has to “much higher” for “exchange” to be a personal equilibrium. On the other hand, if p is high ($p > \frac{1}{2}$), “exchange” may be a PE even if $u_{pen} < u_{mug}$. However, in such a case, “keep” may be a PE as well for certain (u_{mug}, u_{pen}) combinations, and it may indeed be the PPE. We have $EU(\text{“keep”}|\text{“keep”}) = u_{mug}$ (there is no gain-loss utility when the subject is certain to keep the mug). Thus, “exchange” would be the PPE if and only if $EU(\text{“exchange”}|\text{“exchange”}) \geq EU(\text{“keep”}|\text{“keep”})$, or

$$p(u_{pen} - u_{mug}) + \eta[p(1 - p)(u_{pen} - \lambda u_{mug}) + (1 - p)p(u_{mug} - \lambda u_{pen})] \geq 0,$$

which simplifies to

$$(u_{pen} - u_{mug}) + \eta(1 - p)(1 - \lambda)(u_{pen} + u_{mug}) \geq 0.$$

For $p < 1$, “exchange” can be the PPE only if the pen gives higher consumption utility ($u_{pen} > u_{mug}$), as the second term is negative due to loss aversion ($\lambda > 1$). Crucial for our hypothesis, note that as p is increased, “exchange” is the PPE for more pairs of u_{pen} and u_{mug} (the required consumption utility surplus from the pen becomes smaller), as increasing p makes the second term less negative. Finally, note that the more loss averse individuals are, the less likely it is that “exchange” is the PPE, as increasing λ amplifies the gain-loss utility from sometimes being disappointed about being unable to exchange.