

Is the Endowment Effect an Expectations Effect?*

Ori Heffetz[†]

Cornell University

Johnson Graduate School of Management

John A. List

University of Chicago

Department of Economics

First Draft: January 4, 2011

This Draft: October 9, 2013

Abstract

A hallmark result within behavioral economics is that individuals' choices are affected by current endowments. A recent theory due to Köszegi and Rabin (2006) explains such endowment effect with a model of expectations-based reference-dependent preferences. Departing from past work, we conduct complementary experiments to disentangle expectations—verified probabilistic beliefs held by subjects—from other features of endowment—such as “assignment” to a good—hence allowing us to compare the effect of expectations with that of other variations. While mere assignment can affect choices, we do not find a large role in the effect for Köszegi-Rabin expectations.

KEYWORDS: endowment effect, expectations, prospect theory, experiments

JEL CLASSIFICATION: C9, D11.

*A previous version of this paper circulated under the title “Is the Endowment Effect a Reference Effect?” We thank Nicholas Barberis, Daniel Benjamin, Daniel Kahneman, Jack Knetsch, Botond Köszegi, Ted O'Donoghue, and Matthew Rabin for helpful discussions; participants in BEAM, the ESA North American Meetings, Thurgau Experimental Economics Meeting, and seminars at Hebrew University, UCSD, University of Pittsburgh, and USC for useful comments; anonymous referees for suggestions that substantially improved the paper; and Sean Garborg and David Novgorodsky for excellent research assistance.

[†]Corresponding author. E-mail: oh33@cornell.edu.

1 Introduction

A substantial body of experimental evidence has accumulated since Knetsch's (1989) first 'endowment effect' demonstration, where most subjects chose to keep a randomly-assigned owned object rather than trade it for another.¹ While numerous subsequent studies replicate the original result, recent findings show that the effect may shrink or disappear among market-experienced subjects (List, 2004), among lab subjects who are trained to trade (Engelmann and Hollard, 2010), or under specific experimental procedures (Plott and Zeiler, 2007).

This large and growing body of experimental evidence has been interpreted differently by different observers, and the search for a theory to organize it has generated much controversy.² A recent candidate theory whose potential is increasingly recognized is Kőszegi and Rabin's (2006) (henceforth KR) model of expectations-based reference-dependent preferences. The goal of our paper is to report evidence from three lab experiments aimed at assessing to what extent KR's model can explain Knetsch's (1989) findings.

In KR's model—which builds on, extends, and generalizes ideas from Kahneman and Tversky's prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991)—individuals evaluate economic outcomes not only by considering the outcomes themselves but also by comparing them with a reference point, weighing losses relative to the reference point more heavily than gains.³ Importantly, KR combine the model with a theory of reference points that are determined by expectations: “a person's reference point is the probabilistic beliefs she held in the recent past about outcomes.” The resulting expectations-based model holds promise to explain both the finding of the original effect (along with its

¹The term 'endowment effect' was coined by Thaler (1980), who predicted that a good's value to an individual would increase once it became part of her endowment.

²For a flavor of some interpretations and arguments, see e.g. the first few paragraphs in each of the papers cited above.

³These are not the only components of prospect theory. The theory also assumes, e.g., that utility is concave in gains and convex in losses, and that individuals weight probabilities. However, these other components of the theory are not necessary for predicting the endowment effect, and are not part of the KR model considered in this paper.

many replications) *and* the finding that the effect shrinks or disappears in some experimental environments. For example, under the auxiliary assumption that subjects expect to keep an endowed object *unless* the specific environment (e.g., certain experimental procedures, or subjects’ past experience as traders) leads them to expect otherwise, KR’s model is consistent with all past evidence that we are aware of.⁴ An important feature of our experiments is that we directly manipulate expectations, disentangled from endowment. This permits a clean test of the model without making auxiliary assumptions regarding the environment-dependent link between endowment and expectations.

Our first experiment has a 2×2 design, which consists of two “assignment” conditions (Coin-Mug vs. Coin-Pen) and two “expectations” conditions (Weak Expectations vs. Strong Expectations). We present 102 subjects with two goods (a mug and a pen) and ask them to toss a coin which, as the subjects subsequently learn, determines which of the two goods is assigned to them. What “assigned” means differs by expectations condition. In our Weak Expectations treatment, subjects will receive the assigned item with 1% probability, and with 99% probability they will be able to choose one of the two items. In our Strong Expectations treatment, these probabilities are reversed: subjects receive the assigned object with 99% probability, and with 1% probability they choose.

Our 2×2 design allows us to compare the effect on choice of mere assignment to an object with the effect on choice of expecting with near certainty to own the assigned object at the end of the experiment. This amounts to comparing an effect that might be part of Knetsch’s (1989) original effect but *cannot* be explained by KR, with an effect that might be part of the original effect and *can* be explained by KR. Below we refer to the former as a “non-KR-esque” effect, and to the latter as a “KR-esque” effect. To the best of our knowledge, our experiment is the first to compare the two.⁵

⁴Indeed, Knetsch and Wong (2009) explicitly make a variant of this auxiliary assumption, and conduct experiments that assess KR under this assumption. Specifically, they hypothesize that the endowment effect is triggered by endowment and ownership only in contexts where—and only to the extent that—endowment and ownership happen to influence subjects’ expectations (and they present evidence that is consistent with this interpretation).

⁵Since the non-KR-esque effect of assignment may be smaller than that of endowment—assignment is a

To the extent that endowment in past experiments affected outcomes through establishing expectations (regarding outcomes), assignment in our experiment should affect choice more in the Strong Expectations treatment than in the Weak Expectations treatment. However, while we find that assignment affects choice—subjects are more likely to choose an item when it is assigned to them by the coin-flip than when it is not (on average by 13 to 20 percentage points)—we find no evidence that the effect is larger in the Strong than in the Weak Expectations treatment. Our findings are hence new in two ways. First, we are the first to demonstrate the existence of a non-KR-esque effect not only in the absence of past confounds (e.g., those discussed by Plott and Zeiler, 2007) but, importantly, with the expectations channel carefully controlled. Second, we are the first to demonstrate the possibility that meaningful KR-esque effects are absent in spite of a wide variation in (verified) expectations regarding feasible choice sets.

This second finding—the absence of a KR-esque effect—stands in stark contrast with a closely related experiment (Ericson and Fuster, 2011) (henceforth EF) that was conducted contemporaneously with, and independently of, our experiment. While EF’s experiment has no assignment treatments—the authors simply endow 45 subjects with a mug—it has two expectations treatments that are somewhat similar to ours: a coin-flip determines the probability with which subjects will be able to trade the mug later for a pen; the probability can be either 10% (a strong expectations treatment) or 90% (a weak expectations treatment). EF find that 77% and 43% of mug owners, respectively, choose to keep the mug in the two treatments.⁶ In other words, they find that once a good is endowed, their expectations

weaker condition than endowment and implies ownership only implicitly—our assignment treatment may capture part, but not all, of the non-KR-esque effect of endowment. At the same time, as our variation in expectations (from 1% to 99% probability of being able to choose) captures almost the entire possible range, our expectations treatment is designed to capture almost all of the potential KR-esque effect of endowment. For this and other reasons discussed later in the paper, one should be cautious in interpreting this comparison. (We return to our distinction between assignment and endowment in Section 5.4.)

⁶On the other hand, EF find the opposite result—38% and 71%, respectively, choose to keep the mug in their strong and weak expectations treatments—when instead of flipping a coin in the presence of subjects they randomly assign 63 subjects to one of the two treatments in a manner not transparent to the subjects. As they observe, while the latter design is confounded along the lines suggested by Plott and Zeiler (2007)—the probability with which a subject is allowed to trade the mug may itself be perceived by subjects as conveying information regarding the mug’s value—the latter result is also informative regarding the relative

treatments do affect choice (by 34 percentage points, an estimate that lies outside of our 95% confidence intervals).

Why did these expectations treatments have the effect predicted by KR in EF’s experiment but not in ours? Since the disparate findings cannot be reconciled within the KR model, the potential explanations we investigate focus on cross-experiment differences—e.g. differences in experimental procedures and other implementation details—that from the model’s point of view are not expected to matter. In order to assess these explanations, as well as to verify that our findings replicate under modified implementation details, we conduct a second experiment (with 233 new subjects). Its design has a dual purpose: first, it attempts to eliminate some of the most salient implementation differences between our first experiment and EF’s; and second, it attempts to further explore our results, including the non-KR-esque effect.

Among our second experiment’s design modifications, we change the probabilities in the expectations treatments to 10% vs. 90% (from 1% vs. 99%). This change both makes our expectations treatment more similar to EF’s and allows us to check the robustness of our findings. We also turn our original 2×2 design into a $2 \times 2 \times 2$ design by adding an “endowment” treatment to augment the two original treatments. Our second experiment thus includes a More Endowment condition, under which a full version of our first experiment is conducted but the goods are no longer treated symmetrically, and assignment comes with strong endowment- and ownership-related language that closely follows the language used by EF. It also includes a Less Endowment condition, under which another full version of our first experiment is conducted, with assignment language that attempts to minimize even implicit hints of ownership that may be driving the non-KR-esque effect we find in our first experiment.

Among our second experiment’s findings, in the More Endowment condition we essentially replicate the findings from our first experiment: while we find that assignment affects

strength of the expectations channel as a potential driver of the effect.

choice—subjects are more likely to choose an item when it is assigned to them by the coin-flip than when it is not (this time by 20 to 22 percentage points on average)—we again find no evidence that the effect is larger in the Strong than in the Weak Expectations treatment. This replication, under conditions that include the dramatic increase in endowment-related language, the change in probabilities, and other implementation changes (including slightly different subject population and items), rules out important potential explanations regarding the differences in the role of expectations across our and EF’s results. In the Less Endowment condition, which can be interpreted as exploring the lower boundary of the non-KR-esque effect of assignment, we find that the effect disappears (and perhaps even reverses) when the assignment language becomes less suggestive of ownership than in our first experiment.

Finally, to further explore the robustness of our main findings, we conduct a third experiment with an additional 225 subjects. Its design replicates the More Endowment condition from the second experiment, but in addition to flipping a coin (to determine which item is assigned to them) subjects also flip a plastic chip that determines, in a transparent way, which expectations treatment (10% vs. 90%) they are randomized into. This experiment, which may be our most similar to EF’s, again essentially replicates the findings from our first experiment: we find a 12-percentage-point average effect of assignment on choice that is not larger in the Strong than in the Weak Expectations treatment.

Overall, while our data cannot reject small but positive KR-esque effects of our expectations treatments on choice, our 95% confidence intervals are sufficiently narrow to reject large effects. For example, pooling the 342 More Endowment subjects across our second and third experiments, we can reject KR-esque effects that are larger than 3 and 14 percentage points, respectively, among our coin-mug and coin-pen subjects.⁷ In particular, our con-

⁷The 95% confidence intervals around a possible KR-esque effect among our coin-mug and coin-pen subjects in Experiment 3 are, respectively, $[-27\%, 10\%]$ and $[-17\%, 18\%]$ (we reorient the signs so that a positive number represents an effect in the direction predicted by KR). In our second experiment’s More Endowment condition, the respective confidence intervals are $[-40\%, 6\%]$ and $[-28\%, 23\%]$. Pooling the data across these two conditions (which are indistinguishable apart from the differences regarding transparent randomization of expectations), the two confidence intervals are $[-26\%, 3\%]$ and $[-15\%, 14\%]$. Further pooling the data to include Experiment 1’s subjects or, indeed, all 560 subjects across our three experiments (280 coin-mug and 280 coin-pen subjects) makes these confidence intervals still narrower. (Experiment 1’s

fidence intervals—from each of our experiments separately and from pooled data—always reject a 34-percentage-point effect like that found by EF.

More generally, our results suggest that while mere assignment can, under some conditions, have an effect on choice that is not easily explainable by KR, expectations per se may fail to generate the effect predicted by KR. This may in turn suggest that for a viable explanation of Knetsch’s original result, expectations-based models may be too narrow—as they leave out a potentially important part of the explanation—and, at the same time, may be too broad—as their notion of expectations may not be sufficiently refined. While our data demonstrate that having formed correct probabilistic beliefs may in itself not be enough for meaningfully affecting behavior, our findings are at the same time fully consistent with the notion—as well as with past evidence—that specific manipulations of expectations *can* have significant effects, e.g., by affecting the time individuals spend thinking about different outcomes, or even simply by being made artificially salient.

The remainder of this paper proceeds as follows. Section 2 describes the design of our first experiment and discusses the theoretical framework that motivates it. Section 3 presents the results from our first experiment. Section 4 describes the motivation and design of our second experiment, and Section 5 discusses its results. Section 6 discusses our third experiment. Section 7 concludes.

2 Experiment 1: Assignment and Expectations

In this section we describe our first experiment, solve KR’s model in the context of our experiment, and derive experimental predictions. We start with a brief outline of a simple version of the model.⁸

own confidence intervals are $[-33\%, 9\%]$ and $[-26\%, 33\%]$.) These and other confidence intervals are readily calculated from the information reported in Tables 1–3 below.

⁸KR’s model could be viewed as consisting of three components: (a) prospect theory’s value function; (b) a reference point that is determined by expectations; and (c) a theory of rational expectations. KR note that the last component—the assumption that the expectations that determine the reference point are “fully rational” (in a sense spelled out in their model)—could be replaced by “any theory of how these expectations are formed. But as a disciplined and largely realistic first pass,” KR “assume that expectations are fully

2.1 The KR Model

In KR’s model, a consumer’s utility depends not only on her K -dimensional consumption vector \mathbf{c} but also on a reference vector \mathbf{r} . Her overall utility,

$$u(\mathbf{c}|\mathbf{r}) = \sum_k m_k(c_k) + \sum_k \mu(m_k(c_k) - m_k(r_k)),$$

consists of two components, both separable across dimensions. The first, “consumption utility,” corresponds to standard, ‘classic,’ utility. The second, “gain-loss utility,” corresponds to prospect theory’s reference-dependent utility. The value function μ satisfies $\mu(x) = \eta x$ for $x > 0$, and $\mu(x) = \eta\lambda x$ for $x \leq 0$. The parameter $\eta > 0$ is the weight an individual attaches to gain-loss utility, and $\lambda > 1$ is her “coefficient of loss-aversion.” Hence λ is a measure of prospect theory’s famous “kink”: the pain from a(n arbitrarily small) loss is greater than the pleasure from a gain of equal size. The model allows for both \mathbf{c} and \mathbf{r} to be stochastic, and assumes that individuals maximize expected utility.

As noted above, in KR’s version of prospect theory, the reference vector \mathbf{r} results from expectations: “Specifically, a person’s reference point is her probabilistic beliefs about the relevant consumption outcome held between the time she first focused on the decision determining the outcome and shortly before consumption occurs” (Kőszegi and Rabin, 2006).

The reference \mathbf{r} is determined endogenously, as what KR term a *preferred personal equilibrium* (PPE). A PPE is a probability distribution over consumption outcomes that satisfies the following two conditions. First, it is a *personal equilibrium* (PE), which is a rational expectations equilibrium in the following sense. Given a consumer’s expectations regarding the state of the world—represented by a probability distribution over *choice sets*—she forms expectations regarding choice outcomes—a probability distribution over *consumption vectors*. These expectations over outcomes are rational in that they are consistent: a consumer who holds them as her reference will indeed find that following through, by making the ex ante

rational.” Notice that the present paper is not aimed at asking to what extent the different components of KR’s theory explain the endowment effect, but rather to what extent the *combined* theory explains behavior.

expected choices, maximizes her utility.

Second, a PPE is a preferred PE: when more than one PE exists, a PPE is the one that maximizes ex ante expected utility. In other words, when the consumer can form more than one set of expectations regarding outcomes which, once her reference, is consistent with optimal choices ex post—she will choose as her reference point the ex ante preferred one. For a formal exposition and a detailed discussion see Kőszegi and Rabin (2006).

2.2 Experimental Design

We now apply KR’s model to our first experiment. The main contribution of our experiment is that it disentangles assignment from expectations: our 2×2 design separately varies a subject’s assignment to an item and her expectations regarding how much assignment will matter for the choice sets she will face. We do the former by a coin flip and the latter by directly and explicitly informing her about the relevant probabilities. With no ambiguity regarding her choice-set expectations—which we verify with a quiz (see below)—we can solve the model and derive testable predictions.

Our first experiment involves two consumption goods, a mug and a pen, which we denote below c_i and c_j . It is conducted at a large university in the Midwest. Each experimental subject is seated at a table, on which the two goods are located, along with a set of printed instructions, a printed survey, a coin, an envelope, and something to write with (for a photo of the experimental setup and for an example of the experimental instructions see Appendix A).

On the first page of the instructions subjects are asked to flip a coin and to choose a number between 1 and 100. The experimental procedure is explained on the second page, which subjects are allowed to see only after they marked down on the first page both their coin-flip outcome and their choice of a number. In our Strong Expectations treatment, the second page opens with the text:

In front of you are two items. You will get one of them as a gift to take home. Whether or not you can choose your gift is determined at random, as explained below in detail.

In brief, there is a 1% probability that you will be able to choose which item you take home. However, there is a 99% probability that you will NOT be able to choose, and that your gift will automatically be the pen if you flipped “heads” and automatically be the mug if you flipped “tails.”

Feel free to inspect the items but please return them both to their places before we continue.

The text in our complementary Weak Expectations treatment is identical, but “1%” is replaced with “99%,” and “99%” is replaced with “1%.” The reminder of the second page of the instructions explains the experimental protocol in detail, including explaining how the 1% vs. 99% randomization will be carried out (the sealed envelope next to the subject contains a number between 1 and 100, which has a 1% probability of matching the number the subject wrote on the first instructions page). In addition, the explanation above regarding probabilities is repeated, using alternative—and perhaps more intuitive—descriptions. For example, in the Strong Expectations treatment:

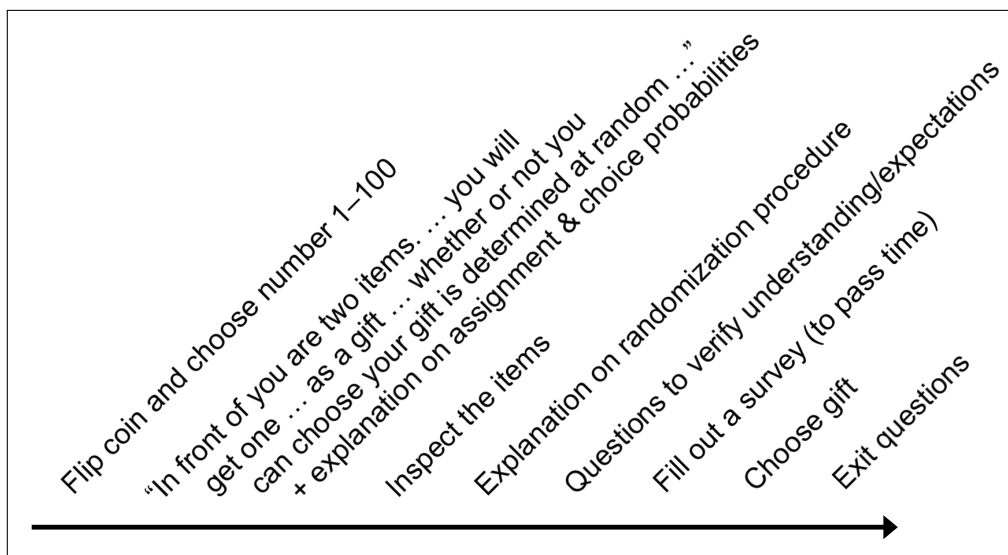
Notice that you have a 1% chance (or 1/100) to be able to choose your gift at the end of the survey. In other words, there is a very high probability that you will take home the gift determined by the coin-flip, regardless of which gift you choose.

Subjects are then asked if they have any questions. After all questions are answered by the experimenter, they proceed to the third page, which includes two quiz questions to verify that they indeed understood the instructions and formed correct expectations regarding the probabilities of future consumption of either good given their expected choice. The quiz is identical across treatments, and is constructed in a way that leaves little chance for answering it correctly without full understanding of the exact probabilities with which each potential outcome can occur.

Once subjects finish the quiz, the experimenter checks their answers, and they then proceed to fill out a survey (to pass time).⁹ Having completed the survey, they are asked to

⁹Subjects who did not answer both quiz questions correctly are informed that they have incorrect answers

make their choice of an item. Their choice is our outcome of interest.¹⁰ After having made this choice, on the last page of the instructions subjects are asked exit questions and the 1% vs. 99% uncertainty is resolved. The timeline below summarizes the different stages of the experiment.



Experiment 1: Timeline

This experimental design has a few important features. First, the goods are treated symmetrically, and the instructions contain no biased language like “keep” or “trade” that might affect subjects’ choices. Second, the coin-flip guarantees that subjects are aware that their default gift resulted from a random 50-50 draw. It rules out the concern that a good’s assignment as a default gift may be interpreted by subjects as informative, e.g. regarding

and are asked to re-read the instructions and try again. If they again have mistakes, the experimenter goes over the instructions and quiz with them for a third time, and then asks them to proceed regardless of their answers. Importantly, we record the number of attempts it took each subject to answer the quiz correctly, and use it as a measure of how confident we are that a subject formed expectations as intended.

¹⁰Notice that we record gift choices before the 1% vs. 99% uncertainty is resolved, hence before subjects find out whether they will get their choice or the gift assigned by the coin-flip. This allows us to elicit real-stake choices from all subjects (rather than from only the fraction of them who happens to end up in the choice condition). On the other hand, with this design we cannot test whether subjects’ choices once they know which condition they face are consistent with their ex ante choice-expectations. It would be interesting to test this aspect of the model in the future, for example by modifying the design to record choice only after the (choice vs. assigned gift) uncertainty is resolved. Notice however that this will require $1/q$ times more subjects, where $q = 0.01$ in one treatment, and $q = 0.99$ in the other. Finally, notice that while our current method of eliciting choice could be interpreted as the ‘strategy method,’ it could alternatively be interpreted as a standard fully-incentivized elicitation of choice between a certain outcome and a lottery.

the goods' values (a quality signal) or regarding the "right" choice behavior expected by the experimenter (demand effects). These concerns, which are raised by Plott and Zeiler (2007), cannot arise in our experiment. This in turn guarantees that an effect of coin-flip assignment on choice in our experiment cannot be explained by the standard neoclassical model.

Third, our procedure is explained to subjects at the outset, and so they are never surprised by facing a choice they did not previously realize they might have to make. This feature is absent in past variations of Knetsch's (1989) original experiment, where subjects are first endowed with a good, and only later learn that in fact they can trade it. In other words, subjects in our experiment explicitly learn at the outset the complete probability distributions regarding outcomes (conditional on choices), and we have no need to speculate—as, e.g., Knetsch and Wong (2009) do—regarding their expectations. Moreover, we have direct evidence on these expectations in subjects' responses to the quiz questions.

2.3 Solving the Subject's Consumer Problem

We now solve the consumer problem faced by subjects. Assume w.l.g. that a subject's default gift, as determined by the coin-flip, is c_1 ; the alternative gift is c_2 . Depending on experimental treatment, the subject can choose her gift with probability $q \in \{0.01, 0.99\}$. Also assume w.l.g. that $m_1(0) = m_2(0) = 0$. If the subject expects to keep c_1 , she can do so regardless of the envelope draw. Her reference consumption is then c_1 regardless of q . If she indeed chooses to keep c_1 when asked for her choice later, her reference coincides with her actual consumption, and her utility—expected as well as realized—is just $m_1(c_1)$, with no gain-loss terms. On the other hand, if she deviates from her reference consumption and chooses c_2 , her utility will be $m_1(c_1)$ with probability $1 - q$ and $m_2(c_2) + \eta m_2(c_2) - \eta \lambda m_1(c_1)$ with probability q . It is thus straightforward to show that given her expectation to keep c_1 (e.g. when answering the quiz), she will indeed choose to keep c_1 (when choice is elicited) as long as

$$\frac{m_1(c_1)}{m_2(c_2)} \geq \frac{1 + \eta}{1 + \eta \lambda}. \quad (1)$$

In other words, choosing c_1 is consistent with a PE as long as (1) holds.

Alternatively, the subject may expect (when answering the quiz) to choose c_2 when choice is elicited later. In that case, she expects to consume c_1 with probability $1 - q$ and c_2 with probability q . Given such expectations and following a line of reasoning similar to the one above, it can be shown that she will indeed choose c_2 later as long as

$$\frac{m_1(c_1)}{m_2(c_2)} \leq \frac{1 + (1 - q + q\lambda)\eta}{1 + ((1 - q)\lambda + q)\eta}. \quad (2)$$

Thus, when (2) holds, choosing c_2 is consistent with a PE.

The model's predictions can now be analyzed for any $0 \leq q \leq 1$. However, for our purposes it is sufficient to point out the following approximate results.¹¹ For q close to 0, there is a unique choice consistent with a PE for $\frac{m_1(c_1)}{m_2(c_2)}$ on either side of $\frac{1+\eta}{1+\eta\lambda}$: if the former ratio is greater than the latter, a subject will choose c_1 ; if it is smaller, she will choose c_2 . As q grows, there is a growing region for $\frac{m_1(c_1)}{m_2(c_2)}$ to the right of $\frac{1+\eta}{1+\eta\lambda}$ where choosing either c_1 or c_2 can be consistent with a PE, and for q close to 1 this region stretches all the way to $\frac{1+\eta\lambda}{1+\eta}$. It can then be shown—by comparing the expected utility associated with one PE with that associated with the other—that for $q \approx 1$ a PPE is only consistent with the subject choosing c_1 if $\frac{m_1(c_1)}{m_2(c_2)} \geq 1$ and c_2 otherwise.

2.4 Predictions

In the $q = 0.01$ (Strong Expectations) treatment, subjects expect to consume the default c_1 with very high probability. According to KR, choosing c_2 would hence involve a relatively large loss term. For example, if $\eta = 1$ and $\lambda = 3$, subjects are predicted to choose c_1 as long as $m_1(c_1) \geq \frac{1}{2}m_2(c_2)$. This is a rather weak condition as long as subjects perceive

¹¹The results we state here are only approximately true in the following sense: they are exactly true for $q = 0$ and for $q = 1$, but become only approximations for q in the vicinity of these extreme values. Replacing these approximations with exact mathematical expressions unnecessarily complicates the presentation while having negligible effect on the relevant empirical predictions.

the two goods to be of roughly comparable value for them.¹² Importantly, interpreting past variations of Knetsch’s (1989) experiment within KR’s framework would involve an identical analysis (and condition); but in past variations, one has to *speculate* about the (supposedly very small) value of q that subjects implicitly infer when they are endowed with a good (with “endowment” implemented in different ways), while in our experiment, subjects explicitly learn q at the outset.

In the $q = 0.01$ (Strong Expectations) treatment, therefore, according to KR and under the above assumptions, coin-flip assignment should affect choice. On the other hand, in the $q = 0.99$ (Weak Expectations) treatment, assignment is not expected according to KR to affect choice, and half the subjects are expected to choose c_2 —the good that was not assigned to them by the coin-flip. (Remember that the standard neoclassical model—i.e., the special case where $\eta = 0$ or $\lambda = 1$ —predicts half the subjects to choose c_2 regardless of treatment.)

To summarize: (i) finding no effect of coin-flip assignment on choice under any of the four experimental conditions is inconsistent with KR’s predictions as long as $m_1(c_1)$ and $m_2(c_2)$ are not very different and as long as η and λ are such that KR is not effectively reduced to the standard model. (ii) Finding coin-flip assignment effect for $q = 0.01$ but not for $q = 0.99$ is consistent with KR (and cannot be explained by the standard model). (iii) Finding coin-flip assignment effect under both expectations treatments would require an explanation—a theory of attachment or a psychology of coin-flip assignment—that KR does not provide (regardless of parameter values), though it can be combined with KR. In that case, one would refer to the smaller of the two effects as a baseline non-KR-esque effect, subtract it from both effects, and compare the residuals across treatments (in the spirit of (i) and (ii) above); if the effect is larger for $q = 0.01$ than for $q = 0.99$, KR might explain some—though not all—of the effect, and one could compare the magnitude of the non-KR-esque effect with that of the KR-esque effect. (iv) Finally, we are aware of no theory (or

¹²Remember that on average, for half the subjects c_1 is the mug and for the other half it is the pen. They cost roughly the same and, crucially, were chosen to be similar to those used in other endowment effect experiments (where the endowment of one of them had a large effect on choices). Also remember that for KR to apply, the goods have to be for final consumption.

combination of theories) that could explain finding an effect for $q = 0.99$ but not for $q = 0.01$ (or, more generally, finding a larger effect for the former treatment than for the latter).

3 Experiment 1: Results

102 subjects participated in our first experiment, which was conducted during September and October 2009. Subjects were mostly undergraduate students, recruited on campus (see Appendix A for an example recruiting flyer). The recruiting materials advertised a show-up fee as compensation, with no mention of the mug, pen, or any take-home gift. While recruiting could have been easier if we had advertised a gift, refraining from doing so was important for avoiding subjects' early formation of expectations regarding ownership of the items, both prior to the experiment and upon entering the lab and seeing the items displayed on the table.¹³

We first describe subjects' replies to the quiz questions, and then we analyze their replies to the choice question.

3.1 Expectations

The quiz page on the experimental instructions includes two questions. Question 1 is reproduced below.

¹³We conducted 31 sessions, with a single experimenter and between 1 and 9 subjects (median = 3) per session. Median estimated session duration = 29 minutes. Randomization into expectations condition, unbeknown to subjects, was done at the individual level. Our first 21 subjects were recruited for "a research study" lasting "approximately 30 minutes" with a \$5 show-up fee as compensation. We then made the following (theory-neutral) modifications to our materials (the modified materials applied to the remainder 81 subjects): (a) subjects were recruited for a study lasting "less than 40 minutes" with a \$10 show-up fee; (b) we switched the order of the mug and pen, both on the instructions (so that "pen" appears first) and in display (so that the pen is placed on the left); (c) we added to the first paragraph of the instructions a sentence explicitly instructing subjects to turn a page only when they have completed it, and to not change answers once completed; and (d) we added on the last page of the instructions (after choice is made) two exploratory, hypothetical WTP questions (that we do not use in this paper). The example materials in Appendix A reflect these modifications.

1. With 1% probability, the number I wrote down will turn out the same as the number in the envelope. In that case:

(Please check one box. If you check the bottom box, please also fill out the blank space.)

- ☐ My gift will be the one I choose, regardless of the coin-flip.
- ☐ My gift will be _____, as determined by the coin-flip, regardless of my choice.

Question 2 is identical to question 1 in all but the first two lines, which in question 2 read:

With 99% probability, the number I wrote down will not turn out the same as the number in the envelope. In that case:

Of our 102 subjects, 94 subjects (92%) answered question 1 correctly in their first attempt. The eight subjects who did not were asked to re-read the instructions and try again. Seven of them answered the question correctly in their second attempt. The experimenter explained the instructions and the quiz questions to the one subject who did not, but we have no evidence that that subject indeed understood the instructions and hence we have no evidence regarding that subject's expectations. In question 2, 100 subjects (98%) answered correctly in their first attempt, and the two subjects who did not also did not answer it correctly in their second attempt. The experimenter then explained to them the instructions and the quiz. Overall, 94 subjects (92%) answered both questions correctly on their first attempt. It seems reasonable to assume that these subjects fully understood the instructions, and formed expectations as intended by the experimental design.¹⁴ We present choice results for the entire population below, and for these 94 subjects in the appendix.

¹⁴Notice that psychologically, fully understanding the relevant probabilities may not be equivalent to forming expectations. However, as is standard in economics, KR *define* the latter as the former. Also, remember that the expectations that our quiz verifies are regarding choice sets, and regarding final consumption conditional on choice. Since our quiz does not elicit planned choice, we cannot directly and fully verify subjects' reference (i.e., expected) consumption vector. This point is related to footnote 8 above: our experimental design relates to the KR model in its entirety rather than to its individual components. We return to these points in the next section.

3.2 Choice

Table 1 reports results for our entire population of 102 subjects. Its leftmost column, titled “All,” suggests that overall, across the two expectations treatments, coin-flip assignment may affect subjects’ choices. The first row shows that 60 subjects’ coin-flips assigned them with a mug and 42 subjects’ coin-flips assigned them with a pen, and the second row shows that of the former, 48 subjects chose a mug as their gift, while 28 of the latter did so. The respective proportions—80% and 67%—are presented in the third row. Finally, the bottom row reports the result of a two-sample two-sided test of equality of proportions. With p -value = 0.13 we cannot quite reject the null that the difference (of 13%) in proportion of mug choosers between coin-mug and coin-pen subjects is statistically indistinguishable from zero. (Obviously, we have more power to reject a one-sided null, which, arguably, is the relevant null to test; see e.g. Plott and Zeiler’s (2007, Table 1) tests, all of which are one-sided.)

[Table 1 about here.]

The next two columns of Table 1 report results by expectations treatment. They show that the difference in choice proportions across coin-flip-assignment conditions (13%) cannot be explained by expectations: the difference in proportions in the Strong Expectations treatment (8%) is in fact roughly half that in the Weak Expectations treatment (17%). Indeed, the rightmost column shows that while in the Strong Expectations treatment 3% less coin-pen subjects chose mug than in the Weak Expectations treatment—a small difference in the direction predicted by KR—the difference among coin-mug subjects—which according to KR should be negative—is positive (12%). As the bottom row shows, none of these latter differences is statistically significant.

Appendix Table A1 reports a version of Table 1 that is based on only the sub-sample of 94 subjects who understood the instructions on first reading, formed expectations as intended, and correctly answered the quiz—our manipulation check—on the first attempt. Among these subjects coin-flip assignment seems to have affected choice more strongly than among

the entire population: the overall difference in proportions between coin-mug subjects who chose a mug and coin-pen subjects who chose a mug grows to 20%, and it is statistically different from zero with $p\text{-value} = 0.03$. However, as in Table 1, the difference in proportions is not larger in the Strong Expectations treatment (17%) than in the Weak Expectations treatment (19%) and hence cannot be explained by expectations. The rightmost column again shows that while the proportion of mug choosers among coin-pen subjects is larger (8% difference) in the Weak than in the Strong Expectations treatment—i.e., it is in the direction predicted by KR—it is also larger (9%) among coin-mug subjects—opposite to KR’s prediction.

To summarize, empirical results in Tables 1 and A1 suggest that while coin-flip assignment could affect subjects’ choices, it does not do so more in the Strong Expectations treatment. While the former cannot be explained by the standard neoclassical model, combined with the latter it cannot be explained by KR either.¹⁵ Our interpretation of the results from our first experiment is hence that (a) we find a possible non-KR-esque effect of coin-flip assignment on choice;¹⁶ and (b) we find no evidence of an additional, KR-esque effect of expectations on choice on top of this non-KR-esque effect.¹⁷

¹⁵KR could explain our findings if subjects exhibit a level of probability weighting so extreme that they regard a 1% chance and a 99% chance as roughly equally probable. Naturally, such interpretation would empty KR from much of its empirical content. Alternatively, a version of prospect theory that posed that the reference point is determined by our coin-flip assignment could, trivially, also explain our findings. However, for such a theory to be both generally useful and falsifiable, it would have to explicitly stipulate what in general does—and, importantly, what does *not*—affect the reference point. The often loosely-used notion “status quo,” for example, cannot explain our findings in any meaningful way without this crucial explicitness.

¹⁶ In a previous version of this paper (Heffetz and List, 2011) we report the results of a follow-up experiment that was designed as an attempt to further disentangle our experimental procedures from expectations. In that experiment we find no strong evidence that any of the treatments matters much, but due to a complicated design and to a social component that is perhaps introduced, these findings may be hard to interpret. See Heffetz and List (2011) for details (including experimental instruments and detailed result tables).

¹⁷One technical concern regarding Tables 1 and A1 is related to the fact that, as seen in the Weak Expectations column in either table, twice as many subjects in that treatment were assigned by their coin-flip to the mug as those assigned to the pen. However, we believe that this unbalance reflects natural variation rather than reflecting e.g. that subjects somehow found a way to affect (or to cheat about) their coin-flip. Our conclusion is based on the following observations. First, the experimenter was always present in the vicinity of subjects (although, admittedly, not always directly looking). Second, more importantly, when subjects flip the coin (on the first page of the instructions), they do not know yet which coin-flip outcome would assign them to which item. Third, and most importantly, unbalanced coin-flips are only found in the Weak Expectations treatment, where the coin-flip has virtually no effect on subjects’ choice

4 Experiment 2: Assignment, Expectations, and Endowment

As aforementioned, concurrently to our research Ericson and Fuster (2011) conducted an experiment in the spirit of our first experiment. How do our results compare with theirs?

First, regarding the possible non-KR-esque effect we find, EF’s experiment is silent: while our 2×2 experiment varies both assignment and expectations, EF’s experiment assigns all subjects to the same item (by endowing all with a mug), and varies only expectations. Indeed, to the best of our knowledge, our non-KR-esque effect—if shown to replicate—would constitute a new result: it would be the first demonstration that assignment to an item can affect choice even when all previously suggested confounds from the perspective of the neoclassical model have been eliminated and, importantly, even when subjects demonstrably know that assignment has virtually no effect on expected ownership. One of the purposes of our second experiment is hence to replicate this new potential result.

Second, regarding our failure to find a meaningful KR-esque effect of expectations, our results—again, assuming they replicate in larger samples—contrast with those of EF, who find that assignment has a significantly larger effect on choice (by 34 percentage points) in their strong compared with their weak expectations treatment. In the rest of this section we discuss potential explanations for the different results, and explain how our second experiment is designed to further explore some of these explanations. For an example of the master document from which our second experiment’s instructions were created (as well as a photo of the experimental setup) see Appendix B.¹⁸

set (it could limit their choice set with only 1% probability) and, according to both the neoclassical model and KR, should have virtually no effect on outcomes. Reassuringly, in the Strong Expectations treatment, where the coin-flip strongly affects subjects’ outcomes (by simply eliminating choice and determining their gift with 99% probability), coin-flip assignments are perfectly balanced. Pooling all 102 subjects across both treatments, a two-sided binomial probability test of balanced coin assignment yields an exact p -value of 0.09.

¹⁸As explained below, yellow and blue highlights in the master document each represent a different endowment treatment.

4.1 Experimental Design

KR’s theory alone cannot explain the different findings across our experiment and EF’s. Indeed, from the theory’s perspective, once coin-flip assignment has been determined in our experiment, the two experiments are equivalent with one exception: while our expectations variation includes 1% and 99% treatments, EF’s includes 10% and 90% treatments—but according to the theory, this difference should lead to a *larger* effect of expectations in our experiment. In other words, from the theory’s perspective, our and EF’s experiments are two equivalently-clean attempts at demonstrating the theory’s potential to explain at least some of Knetsch’s (1989) effect. That EF find that expectations could explain some of the effect while we find no such evidence is, from the theory’s perspective, puzzling.

At the same time, our first experiment and EF’s experiment naturally differ in many implementation details. And while KR’s theory does not predict them to matter, one might worry that—due to explanations that go beyond KR—they do matter. Below we list what we hypothesized could be six such explanations.

(1) A potential concern regarding the design of our first experiment is that a 1%-probability event may be perceived by subjects as so unlikely to occur that they do not choose carefully. While this possibility cannot easily explain our findings, it motivated us to change the probabilities in our second experiment from 1% vs. 99% to 10% vs. 90%, as in EF’s experiment.

(2) While in our first experiment assignment to an item does not use explicit endowment language, EF endow their subjects with a mug using expressions like “the mug in front of you is yours, you own it,” “you may have the option to exchange your mug for the pen,” “[y]ou own this item for real,” “keep the item you currently own,” etc. Since we hypothesized that the effect on choice of an expectations manipulation might depend on whether or not expectations are accompanied by a strong sense of endowment and ownership, our second experiment adds an endowment manipulation to our first experiment, turning it from a 2×2 - to a $2 \times 2 \times 2$ -design experiment.

The master document in Appendix B contains the full instructions of a “More Endowment” treatment (by reading only text with blue highlights or no highlights) and the corresponding “Less Endowment” treatment (yellow highlights and no highlights). Each of the two new endowment treatments consists of a full version of our first experiment, with or without explicit endowment language. Thus, for example, where the instructions in the Less Endowment treatment read “there is a 10% probability that you will be able to choose which item you take home,” in the More Endowment treatment they read “The item you own is yours to keep. You own it for real, not just for the purpose of the study. There is a 10% probability that you will be able to exchange it for the other item if you want to.” More generally, the language in our More Endowment treatment closely follows the endowment-laden language in EF. This includes, for example, asking subjects to write “keep” or “trade” to indicate their choice (as opposed to writing the more symmetrical “pen” or “mug” in the Less Endowment treatment).

Moreover, while our first experiment already carefully avoids explicit endowment language, the Less Endowment treatment in our second experiment is rewritten to also reduce language that might *implicitly* lead subjects to think about endowment or ownership. Thus, for example, while subjects in our first experiment encounter the word “gift” nineteen times by the time they indicate their choice, in our second experiment’s Less Endowment treatment they encounter it only six times.¹⁹

(3) As EF mention in their paper, one psychological mechanism that may explain their findings runs from a subject’s expectations to own an item with high probability to the

¹⁹The fact that the word “gift” is not entirely eliminated from the instructions merely reflects our attempt to keep the instructions’ language intuitive and easy to understand. More generally, our effort at reducing implicit endowment language in the instructions is far from theory-based and should be treated as exploratory. It is intentionally restrained by deliberately avoiding replacement language in cases where we feared the resulting “cleaned up” instructions would feel too contrived. We carefully refer to our endowment treatments as More vs. Less Endowment (rather than, say, Full Endowment vs. No Endowment) to reflect this fact. Finally, note that while the word “gift” in our instructions may be perceived by subjects as implying endowment or ownership, we are careful to never use it in a way that could lead to the information or demand effects discussed by Plott and Zeiler (2007). Indeed, where “gift” is used, our instructions explicitly invite subjects to “choose your gift” and to “indicate which gift, the pen or the mug, you would like to receive” (and, crucially, subjects always know that mug/pen assignment is determined at random).

subject spending time thinking about the item, which in turn affects the subject's reference point. Consistent with such an interpretation, they report that compared with subjects who were endowed with a mug and expected with high probability to be able to trade it later for a pen, subjects who expected a later trade option with low probability “more strongly agree” (see details below) that they spent more time thinking about the mug than about the pen. This interpretation of the data could indeed explain the difference between our results and theirs if the strength of the link from expectations treatments to the time spent thinking about outcomes varies across the experiments. Indeed, if expectations affect outcomes due to the time spent thinking about the outcomes, then the implementation details of how expectations are manipulated in each specific experiment (through repeated explanations, reminders to subjects, visual cues, etc.) might be what drives each experiment's findings. But then, experiments that are viewed as tests of KR could in reality be testing special cases of “time spent thinking” theories.

To explore this possibility, we elicit in our second experiment data comparable to EF's time-spent-thinking data. After subjects make their choice, they are instructed as follows:

Before opening your envelope, please carefully read and think about each of the following statements. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement. **Notice that the scale now goes from 1 to 7.**

Disagree strongly	Disagree	Disagree a little	Neither agree nor disagree	Agree a little	Agree	Agree strongly
1	2	3	4	5	6	7

- ___ a. I like the pen better than the mug.
- ___ b. During the session, I have spent some time thinking about how I would use the pen.
- ___ c. During the session, I have spent some time thinking about how I would use the mug.
- ___ d. During the session, I have spent more time thinking about the pen than about the mug.
- ___ e. During the session, I expected the pen to be the item I take home.
- ___ f. During the session, I expected the mug to be the item I take home.
- ___ g. During the session, I expected the pen more than the mug to be the item I take home.
- ___ h. During the session, I felt that I owned the pen. I felt that it was already mine.
- ___ i. During the session, I felt that I owned the mug. I felt that it was already mine.
- ___ j. During the session, I felt that I owned the pen more than I felt that I owned the mug.

Statements *a–d* above are almost identical to EF’s statements 1–4, with statements *b–d* yielding comparable data on subjects’ time spent thinking about the items.²⁰ We added to the latter three statements two additional sets of three statements that mimic their structure and elicit subjects’ reports regarding which item they expected to take home (statements *e–g*) and regarding feeling of ownership towards the items (statements *h–j*). These additional six statements serve two purposes. First, they are aimed to provide additional evidence on the effects of our treatments on expectations and on feeling of endowment/ownership (see point (4) below). Second, importantly, the two sets of statements are aimed to be comparable to the set of time statements *b–d*, thus providing evidence on our treatments’ effects on time spent thinking, on expectations, and on ownership feeling *relative* to each other.

(4) One may question whether understanding probabilities (e.g. as demonstrated by answering our quiz questions) can indeed be identified with the psychologist’s notion of forming expectations. We again note that since our experiments are designed in the framework of KR’s model, our implementation of “expectations” closely follows their definition, which equates them with beliefs regarding probabilities. However, if one draws a distinction between “understanding probabilities” and “expecting” one might worry that, for whatever reason, while under EF’s experimental procedures probabilities and expectations coincide, under our procedures they do not. Statements *e–g* above help us explore the possibility that while our subjects understand the instructions, they do not report an effect on their expectations. In addition, these statements are useful in bridging the gap between the expectations verified by our pre-choice quiz—expectations regarding choice sets and regarding final consumption conditional on choice—and the expectations that are the reference in KR’s model—(unconditional) expectations regarding the item that subjects will take home.

(5) Relatedly, as KR repeatedly note, their theory is somewhat vague on the timing of expectations formation: “Psychological and economic judgment is needed, for instance, in

²⁰EF’s statements 1–4 are: “I like the mug better than the pen.” “Since the beginning of the session, I have spent some time thinking about how I would use the pen.” “ Since the beginning of the session, I have spent some time thinking about how I would use the mug.” and “Since the beginning of the session, I have spent more time thinking about the mug than about the pen.” Their response scale ranges from 1 to 5.

choosing the appropriate notion of ‘recent expectations.’” While the timing of expectations formation seems roughly comparable across our experiment and EF’s, it is perhaps even more so in our second experiment, where the time-filler survey is shorter (it is a one-page “Big Five” personality traits questionnaire (John and Srivastava, 1999) similar to what EF use).

(6) Finally, and also related to the last two points, recall that in KR’s model, once PPE expectations are formed, individuals’ choices never deviate from equilibrium choices. More generally, consumers in rational choice models are assumed to never ignore information—information regarding probability distributions being only one example—when making their choices. However, in reality (and specifically, in the reality of experiments) it is likely that expectations and other types of information affect choices only to the extent that individuals pay attention to them at the moment of making the choice. This is simply another way of stating the perhaps trivial observation that in the case of informational manipulations, effect size depends not only on the size of the manipulation (e.g., 10% vs. 90% probability) but also on its salience (e.g., how many times the information is repeated to subjects, how it is emphasized visually on the page (or on the screen), how it is framed, etc.). One could always worry that while our subjects form expectations as intended by our treatments, they later forget them or simply do not pay enough attention to them. Alternatively, one could worry that EF’s subjects are overwhelmed by expectations cues so salient and frequent that they forget everything else and are effectively forced to pay attention *only* to expectations.

In principle, these concerns are hard to address without a theory of attention or of saliency of informational manipulations. It is quite possible that EF’s manipulation is more salient, e.g. due to their use of computer screens (which are perhaps more attention-grabbing than our paper instructions); their use of pictures for illustrating probabilities; the many repetitions in their instructions; etc. Indeed, if these differences between the experiments explain the different results, then it is not clear how useful KR’s model is without specifying the implementation details that are required for expectations to affect choice.

At the same time, our first experiment is careful, for example, to not emphasize assignment more than expectations. So while we cannot rule out that EF’s expectations manipulation is “stronger” than ours in some sense that is not specified by the theory, our expectations manipulation seems roughly as strong as our assignment manipulation—which in turn seems strong enough to affect choice.

As a further check on the relative saliency of our expectations vs. assignment manipulations, in our second experiment’s instructions we avoid mentioning the coin-flip from the moment a respondent finished answering the quiz questions until after she chooses an item. Furthermore, we add a reminder, just before a subject indicates her choice, that her choice is real and that she should choose according to her preferences.²¹ Finally, remember that in our second experiment we elicit subjects’ own statements *e-g* regarding their expectations. These statements, which appear immediately after subjects indicate their choice, provide further evidence regarding the strength and persistence of our expectations manipulation.

4.2 Predictions

In this $2 \{\text{Coin-Mug vs. Coin-Pen}\} \times 2 \{\text{Strong Expectations vs. Weak Expectations}\} \times 2 \{\text{More Endowment vs. Less Endowment}\}$ design experiment, the standard neoclassical model still does not predict an effect on choice of any of the treatments, while KR (with η sufficiently greater than 0, λ sufficiently greater than 1, and $m_1(c_1)$ sufficiently close to $m_2(c_2)$) predicts that coin-flip assignment should affect choice more in Strong than in Weak Expectations cells, regardless of endowment condition. Finding such a KR-esque effect of

²¹Specifically, in our first experiment the paragraph immediately before a subject is asked to make her choice reads:

Please indicate which gift, the pen or the mug, you would like to receive regardless of the item that would be assigned to you by the coin flip. In other words, choose the item you want, which may or may not be the same as the item you would get based on the coin flip in the beginning.

In our second experiment, we drop the part of the text that starts with the word “regardless” and ends at the end of the paragraph above, replacing it with a full stop followed by an underlined sentence that reads:

Remember that at the end of the session you will actually take home with you one of the items; fill in the box below according to the item you prefer.

expectations in our second experiment would suggest that its absence in our first experiment might be due to implementation details such as the 1%-vs.-99% probabilities, the timing of expectations formation, or a low saliency of expectations relative to assignment by the time subjects make a choice (see explanations (1), (5), and (6) in the previous subsection); while finding the effect only in the More Endowment condition would suggest a necessary interaction between expectations and endowment (explanation (2) above).

Additionally, the ten statements $a-j$ shed light on whether our expectations manipulation affects the time subjects spend thinking about the items, and on whether it affects subjects' own perceptions of their expectations (explanations (3) and (4)).

Finally, recall that the potential effect found in our first experiment, of assignment on choice regardless of expectations condition, does not lend itself readily to theory-based hypotheses (see footnote 15 above). That said, the variation introduced in our second experiment between the More and Less Endowment treatments explores this non-KR-esque effect both under more explicit and under less implicit references to ownership and endowment relative to our first experiment.

5 Experiment 2: Results

Our second experiment was conducted during May 2011 at a large university in the Northeast. 233 subjects participated. They were mostly undergraduate students, recruited through an online SONA system (see Appendix B for a recruiting screenshot). The recruiting screen advertised show-up fee or course credit as compensation, with no mention of the items or a gift.²²

We again start with subjects' replies to the quiz questions and then move to analyze their choices. We close with analyzing their post-choice statements.

²²We conducted 25 sessions, with a single experimenter and between 2 and 12 subjects (median = 10) per session. Median estimated session duration = 22 minutes. Subjects were recruited to take "an in-person study on economic behavior and earn \$10 OR 1 credit. ... Duration 40 minutes." Randomization into expectations-and-endowment condition, unbeknown to subjects, was done at the individual level.

5.1 Expectations

The two quiz questions in our second experiment are identical to those in our first experiment, with two modifications that reflect the new experimental design: (a) the probabilities change to 10% and 90%; and (b) the word “gift” is eliminated (the language “My gift will be the one I choose” is replaced with “I will take home the item I choose;” see Appendix B).

Of our 233 subjects, 208 (89%) answered question 1 correctly on first attempt, and 16 of the remaining 25 answered it correctly on second attempt. In question 2, 212 subjects (91%) answered correctly on first attempt, and 13 of the remaining 21 answered it correctly on second attempt. Overall, 200 subjects (86%) answered both questions correctly on first attempt. We again present our results for the entire population below, and for these 200 subjects in the appendix.

5.2 Choice

Tables 2 and A2 present our results. Each of the tables consists of two panels, corresponding to the two endowment treatments (titled “More Endowment” and “Less Endowment”). Each panel has a structure identical to that of Table 1.

[Table 2 about here.]

Overall, results in the More Endowment panel in Tables 2 and A2 suggest that our main findings from Experiment 1 hold up statistically, and are robust to re-setting expectation probabilities at 10%-90%, adding strong endowment language to the instructions, and otherwise modifying the design as discussed above. Qualitatively, the results in this panel in the two tables replicate those in Tables 1 and A1, and provide additional support to the two conclusions above. First, the leftmost column shows that overall, coin-flip assignment affects choice: the difference between the proportion of mug choosers among coin-mug vs. coin-pen subjects is 22% ($p = 0.01$) in Table 2 and 20% ($p = 0.04$) in Table A2. Second, the next two columns show no evidence that the difference is larger in the Strong than in the

Weak Expectations treatment. Indeed, as in Experiment 1 and contrary to KR’s prediction, its point estimate is *smaller*: in Table 2 it is 11% in the Strong Expectations treatment, compared with 31% in the Weak Expectations treatment; in Table A2, it is 18% compared with 23%. Finally, as the rightmost column shows, the differences in proportion of mug choosers across expectations conditions are opposite in sign to KR’s prediction among both coin-mug and coin-pen subjects.

This replication of a coin-flip assignment effect without a KR-esque expectations effect reinforces our interpretation of the findings from our first experiment. For example, it provides additional evidence that the absence of the latter effect cannot be explained by how the specific items in our experiments enter preferences: in terms of the utility function described in Section 2, the strong coin-flip effect suggests that for many subjects, $m_1(c_1)$ and $m_2(c_2)$ cannot be too far apart.

Finally, the bottom panel in Tables 2 and A2 reveals that the above effect of coin-flip assignment on choice does not replicate in the Less Endowment treatment. The leftmost column in either table shows a *negative* effect (not statistically significant, at -9% and -8% in Tables 2 and A2), and the next two columns show that while there is virtually no effect of assignment on choice in the Strong Expectations treatment (1% and 2% in Tables 2 and A2), there is a negative and rather large effect in the Weak Expectations treatment (-18% and -19% , which could become statistically significant with more power). While we are careful not to over-interpret these findings, we note that they are consistent with the (perhaps trivial) notion that the psychological effects of assignment are sensitive to language, and that “cleaning up” the language in our instructions in certain ways may eliminate or even reverse them. We also note that, as the rightmost column in either table shows, the differences in proportion of mug choosers across expectations conditions are in the direction of KR’s prediction among both coin-mug and coin-pen subjects (5% and 14% – 16% , respectively, not statistically significant).

5.3 Statements

Appendix D (with its Figures A1 and A2) reports and discusses our findings on responses to the ten post-choice statements. Our main conclusion is that while we find small and insignificant cross-treatment variation in mean responses to statements $a-d$, we find larger and often statistically significant differences—in the hypothesized direction—across treatments in statements $e-j$. Recall that while statements $a-d$ are designed to closely reproduce EF’s four statements regarding liking or spending more time thinking about one of the items versus the other, the new statements $e-j$ are designed to provide further information about subjects’ self-reported expectations or feelings of ownership during the experiment. Hence, while the effects that we find on responses to the latter provide additional evidence that our assignment, expectations, and endowment treatments generally “worked” as intended (namely, they generally generated expectations and feelings of ownership as intended), the absence of such effects on responses to the former suggests that—in contrast with EF’s findings from similar statements—our treatments did not additionally affect how much subjects liked the items or how much time they spent thinking about them.

We make three additional observations regarding the statement data reported and discussed in the appendix. First, the data suggest that even relatively low probabilities (of 10%) might sometimes be enough for a statistically significant effect of coin-flip assignment on respondents’ self-reported expectations. Second, they suggest that our expectations treatments may have affected the self-reported expectations of coin-pen subjects more than those of coin-mug subjects. Third, we find that while our endowment treatments are successful in generating feelings of ownership, the effects are consistently larger under Strong (vs. Weak) Expectations, suggesting—perhaps not surprisingly—that our expectations treatments, in themselves, affect feelings of ownership. This in turn highlights the issue that one may not be able to keep subjects’ feelings of ownership (or of endowment) perfectly constant while manipulating expectations; it therefore suggests a possible causal link from expectations to feelings of ownership—which reverses the causal direction implied by interpreting the original

endowment effect as an expectations effect.

5.4 Interpretation

In terms of choice, we qualitatively replicate, under More Endowment, our first experiment’s findings with different probabilities, explicit endowment language, a shorter time-filler survey, and no explicit reference to coin-flip assignment between quiz and choice. While this replication does not separately quantify the impact of each of these design changes but only the sum of their impacts, it suggests that points (1), (2), and (5) in Section 4.1 (regarding probabilities, endowment, and timing) are not likely to jointly explain either our original findings or why they differ from EF’s. As to point (6) (regarding the salience of informational manipulations), while it may well explain why our results are different from EF’s, our replication provides rather strong evidence regarding the relative magnitude of assignment effects vs. expectations effects—evidence that EF do not have. At the same time, our choice findings under Less Endowment demonstrate the sensitivity of assignment effects to language.

In terms of subjects’ post-choice statements, we find that while our expectations manipulation affects at least some subjects’ reports regarding their (take-home-item) expectations, it does not make subjects spend more time thinking about (or liking) one of the items. This suggests that even when subjects perceive our informational manipulation of probabilities as affecting their final-consumption expectations—rendering point (4) less of a concern—manipulating these expectations might not in itself be sufficient for having large effects on behavior, as discussed in point (3).

Moving beyond KR-esque expectations, we note that the evidence from both choice and statements seems to support the new distinction we make in this paper between assignment—the mere process of assigning subjects a default item—and endowment—the process of inducing feelings of ownership of the assigned item. Our findings above—both that the effect of assignment on choice could disappear (and perhaps even reverse) in the absence of en-

dowment/ownership language, and that our endowment and assignment treatments act differently in affecting respondents’ self-reports regarding ownership—suggest that assignment and endowment are indeed far from the same thing and can, at least to some extent, be conceptually distinguished and experimentally disentangled.²³

6 Experiment 3: Assignment and Expectations, Transparently Randomized

As discussed above, Experiment 2 was designed to explore what we suspected could be the most important differences between our and EF’s baseline. Yet, there is another potentially important difference that we left unexplored. Recall that after being endowed with a mug, EF’s subjects flip a coin to determine their 10%/90% expectations condition. In our experiments, in contrast, subjects flip a coin to determine their mug/pen assignment condition, but their expectations condition is not transparently randomized (it is predetermined on their printed instructions). Could it be that our subjects, not knowing that their expectations condition was determined at random, mistakenly infer that, e.g., a low probability to be able to trade may signal that trading is desirable?²⁴

To explore the robustness of our results to transparent randomization of the expectations condition, we conducted a new experiment, which we denote as Experiment 3. For its

²³While to the best of our knowledge we are the first to define (and experimentally manipulate) “assignment” in this context, the term “endowment” has come to mean different things to different people in the endowment effect literature. In the early literature, “endowment” seemed to mean giving someone an object. When research started emerging with altered techniques that eliminated the endowment effect, “endowment” started being used to mean a psychological state of feeling ownership of an object (and the discussion focused on how experimental procedures might alter this psychological state).

²⁴Our coin-flip assignment rules out any inference from expectations condition to the desirability (or value) of the *items* (i.e., of the mug vs. the pen). Hence, to the extent that subjects in our experiments make any (mistaken) inference from their expectations condition, it can only be regarding the desirability of *exchange itself* (regardless of the exchanged items). Since we did not consider this possibility, we did not try to rule it out in Experiments 1 and 2, opting instead for a simpler design. In contrast, in EF’s experiment assignment is not randomized, and hence in the absence of transparent randomization of expectations condition, subjects could have made inferences from expectations condition to the value of the items themselves; see footnote 6 above.

experimental instructions and a photo of the experimental setup see Appendix C. From a subject-experience point of view, the experiment is identical to the More Endowment condition in Experiment 2 except that the randomization into expectations condition is both transparent and made clear to subjects throughout the experiment. Specifically, in Experiment 3 subjects flip, on the first page of the instructions, both a coin—determining, as before, their mug/pen assignment—and a plastic chip—determining, in addition, their 10%/90% expectations condition. Having flipped the coin and chip, the experiment proceeds as before, but every time a reference is made to the 10%/90% condition, the instructions include an explicitly conditional statement of the form: “If you flipped ‘10% Exchange Option’ *<Strong Expectations instruction here>*. If you flipped ‘90% Exchange Option’ *<Weak Expectations instruction here>*.”

Experiment 3 was conducted in April 2013 at the same on-campus lab where Experiment 2 was conducted two years earlier. It used the same recruiting system and recruiting information (study duration, compensation), the same general subject pool, and the same experimental setup (including the same mugs and pens, but not the same experimenter). 225 subjects participated (one additional subject withdrew from experiment midway).²⁵

Unlike our previous experiments, respondents found the two quiz questions—which remained identical to those in Experiment 2—harder to answer on first attempt: of the 225 subjects, only 144 (64%), 162 (72%), and 107 (48%), respectively, answered question 1, question 2, and both, correctly on first attempt. While subjects’ comments and questions during the lab sessions, as well as their first-attempt quiz response patterns, suggest that many who failed to answer correctly on first attempt misread the *quiz questions* but understood the experimental instructions, we cannot rule out that the transparent double-randomization (coin and chip) made the instructions harder to follow. Reassuringly, by second attempt, 218 (97%), 220 (98%), and 213 (95%) of the 225 subjects correctly answered question 1, question 2, and both. These are essentially the same percentages as in Experiment 2 (96%,

²⁵We conducted 25 sessions, with a single experimenter and between 4 and 12 subjects (median = 10) per session. Median estimated session duration = 24 minutes.

97%, and 96%, respectively).

Table 3 reports results for our entire population of 225 subjects. (For completeness, Tables A3 and A4 in the appendix present results for the 107 and 213 first- and second-attempt subjects.) Qualitatively, Table 3 replicates Table 1 and the More Endowment panel of Table 2, and is consistent with the two general conclusions discussed above. First, in the leftmost column, coin-flip assignment affects choice: the overall difference between the proportion of mug choosers among coin-mug vs. coin-pen subjects is 12% ($p = 0.06$). Second, in the next two columns we find no evidence that the difference is larger in the Strong Expectations treatment (diff = 8%, $p = 0.40$) than in the Weak Expectations treatment (diff = 16%, $p = 0.08$), and the rightmost column shows statistically insignificant expectations effects: among coin-mug subjects the effect (8%, $p = 0.39$) is opposite to KR’s prediction, and among coin-pen subjects it is nearly zero (1%, $p = 0.91$). (Table A3’s estimates move around and none of the differences are statistically significant; Table A4 closely replicates Table 3.)

[Table 3 about here.]

Finally, Figure A3 in the appendix reports results for Experiment 3’s ten post-choice statements. To facilitate comparison with Experiment 2’s statement results, the figure replicates the structure of Figure A1 (leaving the Less Endowment part of the graphs blank). The results are generally similar across the two figures, but the expectations treatments in Experiment 3 seem to have cleaner effects on self-reported expectations (statements *e–g*): under Weak Expectations, coin-mug and coin-pen subjects are closer together than in Experiment 2; under Strong Expectations, they are farther apart.

We conclude that our two main findings are robust to transparent, simultaneous randomization of assignment and expectations.

7 Conclusion

The three experiments reported in this paper replace Knetsch’s (1989) endowment treatment with an assignment treatment and an expectations treatment, accompanied by more or by less endowment language. We find, first, that assignment can affect choice in a way that is not easily explained by KR (nor by the standard model); and second, we find no evidence of a meaningful KR-esque effect of expectations on choice.

To the best of our knowledge, both findings are new. The former suggests that there is something about assignment that deserves economists’ attention and that—depending on language—can affect behavior even when past explanations are controlled for, including information confounds, demand effects and, importantly, the expectations channel. We view this as an important area for future research.

As to our latter finding, its contrast with EF’s finding—of a sizable effect that we can easily reject—cannot be reconciled within the KR model, even when the model is extended to include certain basic requirements regarding implementation. This suggests that the notion of expectations employed by economists might be too crude to account for observed differences in the effects of “expectations manipulations” on behavior across experiments. For example, if the different findings are driven by different intensities of the expectations manipulations then, relative to Knetsch’s original effect, either EF’s manipulation is interpreted as excessively strong—with unintended effects that are driven by more than KR’s notion of probabilistic beliefs—or our manipulation is viewed as excessively weak—but then verified probabilistic beliefs (accompanied by self-reports) are not sufficient for establishing KR-esque expectations. Furthermore, the general finding in our second and third experiments—that our treatments affect respondents’ statements regarding expectations and endowment but have no comparable effect on statements regarding liking or spending time thinking about the objects—provides a direction in which different lab manipulations of expectations can result in different effects on subjects’ perceptions. This finding suggests an important direction in which our conceptualization of expectations could be refined.

While expectations regarding outcomes seem to matter for individuals’ choices in certain contexts, we fail to find supporting evidence in the context of Knetsch’s (1989) endowment effect.²⁶ We trust that future work—both theoretical and experimental—will think carefully about the source and nature of expectations. While in economists’ models expectations are often equivalent to beliefs about probability distributions, our findings suggest that the context in which expectations are established matters a great deal. Beyond these experiments and beyond the points we emphasize in this paper—e.g. regarding the salience of informational manipulations—one might further wonder whether expectations exogenously dictated in an experiment affect choices in the same way that naturally occurring expectations do; whether expectations have different effects when they are perceived as, e.g., unfair; or whether (and how) aspirations and expectations interact. These all serve as important research avenues going forward.

References

- Abeler, Johannes, Armin Falk, Lorenz Götte, and David Huffman. 2011. “Reference Points and Effort Provision.” *American Economic Review* 101(2), 470–492.
- Engelmann, Dirk and Guillaume Hollard. 2010. “Reconsidering the Effect of Market Experience on the “Endowment Effect.”” *Econometrica* 78(6), 2005–2019.
- Ericson, Keith M. Marzilli and Andreas Fuster. 2011. “Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments.” *Quarterly Journal of Economics* 126(4), 1879–1907.
- Gill, David and Victoria Prowse. 2012. “A Structural Analysis of Disappointment Aversion in a Real Effort Competition.” *American Economic Review* 102(1), 469–503.

²⁶To the best of our knowledge, our and EF’s studies are the first direct tests of KR’s model in the context of Knetsch’s (1989) original experiment. Recent experimental work that tests KR’s model in other contexts include Abeler et al. (2011) and Gill and Prowse (2012), who test versions of the model’s predictions for effort provision; Smith (2008), who—in a pioneering study that, to the best of our knowledge, is the first experimental test of KR—tests for a WTP-WTA gap; and Ericson and Fuster’s (2011) WTA experiment.

- Heffetz, Ori and John A. List. 2011. "Is the Endowment Effect a Reference Effect?" NBER Working Paper 16715.
- John, Oliver P., and Sanjay Srivastava. 1999. "The Big Five trait taxonomy: History, measurement, and theoretical perspectives." In Lawrence A. Pervin and Oliver P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York: Guilford.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 48, 263–291.
- Knetsch, Jack L. 1989. "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *American Economic Review* 79(5), 1277–1284.
- Knetsch, Jack L. and Wei-Kang Wong. 2009. "The endowment effect and the reference state: Evidence and manipulations." *Journal of Economic Behavior & Organization* 71, 407–413.
- Kőszegi, Botond and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121(4), 1133–1165.
- List, John A. 2004. "Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace." *Econometrica* 72(2), 615–625.
- Plott, Charles R. and Kathryn Zeiler. 2007. "Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory." *American Economic Review* 97(4), 1449–1466.
- Smith, Alec. 2008. "Lagged Beliefs and Reference-Dependent Utility." University of Arizona working paper #08-03.
- Thaler, Richard. 1980. "Toward a positive theory of consumer choice." *Journal of Economic Behavior & Organization* 1(1), 39–60.
- Tversky, Amos, and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *Quarterly Journal of Economics* 106, 1039–1061.

Table 1: Choice by Coin-flip in Experiment 1 (All 102 Subjects)

	All	Weak Expectations ($q = 99\%$)	Strong Expectations ($q = 1\%$)	Difference Weak – Strong
(# coin-mug, # coin-pen)	(60, 42)	(34, 16)	(26, 26)	
(# coin-mug who chose mug, # coin-pen who chose mug)	(48, 28)	(29, 11)	(19, 17)	
(% coin-mug who chose mug, % coin-pen who chose mug)	(80%, 67%) diff = 13%	(85%, 69%) diff = 17%	(73%, 65%) diff = 8%	(12%, 3%)
Result	$p = 0.13$	$p = 0.17$	$p = 0.55$	$p = 0.24,$ $p = 0.82$

Notes: Percentages are rounded to the nearest whole number. All p -values are from two-sample two-sided tests of equality of proportions.

Table 2: Choice by Coin-Flip in Experiment 2 (All 233 Subjects)

	All	Weak Expectations ($q = 90\%$)	Strong Expectations ($q = 10\%$)	Difference Weak – Strong
More Endowment				
(# coin-mug, # coin-pen)	(57, 60)	(32, 26)	(25, 34)	
(# coin-mug who chose mug, # coin-pen who chose mug)	(42, 31)	(26, 13)	(16, 18)	
(% coin-mug who chose mug, % coin-pen who chose mug)	(74%, 52%) diff = 22%	(81%, 50%) diff = 31%	(64%, 53%) diff = 11%	(17%, –3%)
Result	$p = 0.01$	$p = 0.01$	$p = 0.40$	$p = 0.14,$ $p = 0.82$
Less Endowment				
(# coin-mug, # coin-pen)	(63, 53)	(32, 26)	(31, 27)	
(# coin-mug who chose mug, # coin-pen who chose mug)	(41, 39)	(20, 21)	(21, 18)	
(% coin-mug who chose mug, % coin-pen who chose mug)	(65%, 74%) diff = –9%	(63%, 81%) diff = –18%	(68%, 67%) diff = 1%	(–5%, 14%)
Result	$p = 0.32$	$p = 0.13$	$p = 0.93$	$p = 0.66,$ $p = 0.24$

Notes: Percentages are rounded to the nearest whole number. All p -values are from two-sample two-sided tests of equality of proportions.

Table 3: Choice by Coin-flip in Experiment 3 (All 225 Subjects)

	All	Weak Expectations ($q = 90\%$)	Strong Expectations ($q = 10\%$)	Difference Weak – Strong
(# coin-mug, # coin-pen)	(100, 125)	(56, 61)	(44, 64)	
(# coin-mug who chose mug, # coin-pen who chose mug)	(66, 67)	(39, 33)	(27, 34)	
(% coin-mug who chose mug, % coin-pen who chose mug)	(66%, 54%) diff = 12%	(70%, 54%) diff = 16%	(61%, 53%) diff = 8%	(8%, 1%)
Result	$p = 0.06$	$p = 0.08$	$p = 0.40$	$p = 0.39,$ $p = 0.91$

Notes: Percentages are rounded to the nearest whole number. All p -values are from two-sample two-sided tests of equality of proportions.