# Heterogeneity of Gain-Loss Attitudes and Expectations-Based Reference Points

Pol Campos-Mercade* Lorenz Goette† Thomas Graeber‡

University of Copenhagen University of Bonn Harvard University

Alexandre Kellogg§ Charles Sprenger¶

UC San Diego Caltech

This Version: August 4, 2022

**Abstract**

Existing tests of reference-dependent preferences assume universal loss aversion. This paper examines heterogeneity in gain-loss attitudes, and explores its implications for identifying models of the reference point. In two experimental settings we measure gain-loss attitudes, and then study a canonical treatment effect which distinguishes different models of the reference point. Accounting for measurement error, we document substantial heterogeneity in gain-loss attitudes, with approximately three-quarters loss-averse subjects. We then document heterogeneous treatment effects over gain-loss attitudes consistent with formulations of expectations-based reference points. Our findings provide foundational support for reference points derived from expectations, and explain inconsistencies across prior exercises.

*JEL classification:* D81, D84, D12, D03

*Keywords*: Reference-Dependent Preferences, Rational Expectations, Personal Equilibrium, Real Effort, Expectations-Based Reference Points

*Department of Economics, University of Copenhagen; pcm@econ.ku.dk

†Department of Economics, National University of Singapore; ecslfg@nus.edu.sg

‡Harvard University, Harvard Business School, Soldiers Field, Boston, MA 02136, tgraeber@hbs.edu

§Department of Economics, University of California, San Diego; alexkellogg@ucsd.edu

¶California Institute of Technology; sprenger@caltech.edu

# 1 Introduction

Models of reference-dependent preferences are regarded as a major advance in behavioral economics, rationalizing a range of observations at odds with the canonical model of expected utility over final wealth (Camerer et al., 1997; Kahneman et al., 1990; Odean, 1998; Rabin, 2000). The predictions of any reference-dependent model hinge on two model components: the reference point governing the location around which gains and losses are encoded; and, gain-loss attitudes encapsulating how individuals weigh gains and losses relative to the reference point.

Recent tests of reference-dependent models focus on hypotheses about the location of the reference point—distinguishing backward-looking factors such as experience and status quo from forward-looking expectations-based mechanisms (Bell, 1985; Kőszegi and Rabin, 2006, 2007; Loomes and Sugden, 1986). As the reference point represents a powerful degree of freedom in application, these tests have been valuable for understanding how to discipline reference-dependent models (Abeler et al., 2011; Cerulli-Harms et al., 2019; Ericson and Fuster, 2011; Gneezy et al., 2017; Heffetz and List, 2014; Smith, 2019). Importantly, all prior exercises have been conducted under a specific *homogeneity* assumption on gain-loss attitudes: universal loss aversion, where all individuals weigh losses more severely than commensurate gains. In this manuscript, we examine the possibility that individuals are *heterogeneous* in their gain-loss attitudes—i.e., some individuals are "loss averse", weighing losses more than gains, and others are "gain-seeking", weighing gains more than losses—and explore the implications of this heterogeneity for identifying models of the reference point.

Accounting for heterogeneity in gain-loss attitudes when testing reference-dependent models is important for two reasons. First, within experimental designs used to identify expectations-based reference dependence (EBRD), different directional predictions are generated depending on whether individuals are loss-averse or gain-seeking. Gain-seeking subjects should react to key experimental treatments in exactly the opposite way as loss-averse subjects. Second, heterogeneity in gain-loss attitudes reflects an empirical realism: a recent literature has noted that even with loss aversion on average, sizable minorities of subjects in lottery choice experiments appear to be gain-seeking, apparently weighing gains more than commensurate losses (Brown et al., 2021; Chapman, Dean, et al., 2017; Erev et al., 2008; Fehr and Goette, 2007; Harinck et al., 2007; Knetsch and Wong, 2009; Nicolau, 2012; Sokol-Hessner et al., 2009; Sprenger, 2015).[1] If individuals are heterogeneous in their gain loss-attitudes and behave in theoretically predicted ways, then prior exercises have aggregated different signed effects without any way to disentangle heterogeneity in attitudes from the corresponding test of the reference point.[2] The combination of these two issues may explain the inconclusive, and at times contradictory, findings in the study of EBRD models without accounting for heterogeneity.[3]

We implement two pre-registered experiments with a total of 1524 subjects to investigate the relevance of heterogeneous gain-loss attitudes for testing models of reference dependent preferences. Our baseline designs and treatment manipulations closely follow existing work on the two main paradigms used to test the EBRD formulation: labor supply (e.g., Abeler et al., 2011; Gneezy et al., 2017) and exchange (e.g., Cerulli-Harms et al., 2019; Ericson and Fuster, 2011; Heffetz and List, 2014). Each experiment consists of two stages. Stage 1 measures each participant's gain-loss attitudes in the specific context of the experiment. Stage 2 tests EBRD by changing subjects' expectations between a Low expectations and a High expectations condition. Under EBRD models, such manipulations change the location of the reference point, and so should change behavior. Under alternative formulations of reference points, no such effects are predicted. Hence, these designs constitute tests of the expectations-based formulation of the reference point.

[1] Chapman, Snowberg, et al. (2018) evaluate seven prior studies from lottery choice along with a prior version of this manuscript and report a weighted average of 22% gain-seeking subjects. They also document nearly 50% gain-seeking subjects in a lottery choice experiment with a representative sample.

[2] In Appendices A.4 and B.4 we demonstrate this point concretely. We show that predicted KR treatment effects are not necessarily linear in gain-loss attitudes. Hence, the average treatment effect may not coincide with the treatment effect of the average preference. Indeed, null and mis-signed average treatment effects (relative to the loss averse prediction) can easily occur with an average preference of loss aversion.

[3] While early experimental applications showed treatment effects in line with the EBRD formulation of reference points (Abeler et al., 2011; Ericson and Fuster, 2011), other exercises have shown more limited or contradictory effects (Cerulli-Harms et al., 2019; Gneezy et al., 2017; Heffetz and List, 2014; Smith, 2019).

The EBRD directional predictions for behavior change in these two leading paradigms depend on gain-loss attitudes, and are opposite for loss-averse and gain-seeking individuals. Aggregating such different-signed effects can lead heterogeneity in gain-loss attitudes to confound the test of EBRD in both settings. Our experimental innovation is the addition of Stage 1 to each design in order to measure gain-loss attitudes in each context. These measures allow us to evaluate the extent of heterogeneity in gain-loss attitudes, account for heterogeneity when testing the EBRD formulation of the reference-point, and examine the heterogeneous treatment effects over gain-loss attitudes predicted by EBRD models.

Our two studies generate two very similar results. First, we find substantial heterogeneity in gain-loss attitudes. While subjects in both studies exhibit loss aversion on average, we estimate sizable minorities of gain-seeking subjects. Even after accounting for noise and errors, both studies show around three quarters loss-averse, and one quarter gain-seeking subjects.[4] Indeed, we measure gain-loss attitudes using three different techniques—labor supply, exchange behavior, and lottery choices—and find a very similar proportion of gain-seeking individuals in all three settings. These findings reinforce prior results on heterogeneous gain-loss attitudes in lottery choice, and clarify that homogeneous loss aversion would be an incorrect assumption to maintain in tests of reference-dependent models.

Second, in each study, gain-loss attitudes from Stage 1 are highly predictive of the treatment effects observed in Stage 2. We document precisely the heterogeneous treatment effects predicted by EBRD. Loss-averse and gain-seeking subjects respond in opposite directions to the manipulation of expectations. Without accounting for heterogeneity, we would draw very different conclusions from our studies, finding more limited, or even no, aggregate support for EBRD. This represents the first experimental test of EBRD accounting for heterogeneous gain-loss attitudes, and the first experimental findings of heterogenous EBRD treatment effects over gain-loss types.

[4] Using a Multiple Price List (MPL) for lottery decisions (similar to Sprenger, 2015) at the end of the labor supply experiment, we further estimate that 25 percent of the subjects are gain-seeking for lottery decisions.

Our empirical results show both the heterogeneity in gain-loss attitudes and the corresponding heterogeneous treatment effects predicted by the theory in the two primary experimental environments in which the theory has been tested. This provides a credible foundation for the EBRD formulation of the reference point. Our findings indicate that mixed evidence on EBRD is likely not driven by a failure of the expectations-based formulation of reference points, but rather by a failure of the second component of the joint hypothesis inherent to prior tests: that gain-loss attitudes are both universal and loss averse. Without accounting for heterogeneous gain-loss attitudes, prior tests may suffer from both aggregation and power issues: the average treatment effect need not be the treatment effect of the average individual (which we discuss in detail in Appendices A.4 and B.4), and potentially muted theoretical average effects can require larger sample sizes for appropriately-powered experiments. In a simple and reproducible way, we show that the predictions of EBRD are reliably recovered once one accounts for heterogeneity in gain-loss attitudes.

A further contribution of our work is to add several large sample observations on the heterogeneity of gain-loss attitudes to a growing literature on the topic (Brown et al., 2021; Chapman, Snowberg, et al., 2018). Chapman, Snowberg, et al. (2018) indicate eight prior studies with a documented distribution of gain-loss attitudes, only one of which is measured outside of lottery choice (a prior version of this paper). Ours are the first findings to document the distribution of gain-loss attitudes in labor supply and exchange settings, and the predictive validity of resulting individual measures. In both experiments, we document an average attitude of loss aversion, but a sizable minority of subjects exhibits gain-seeking attitudes.

This paper highlights the need to account for heterogeneity in gain-loss attitudes in order to use and test models of reference-dependence. Besides expectations-based models, our results also have implications for other applications of gain-loss attitudes, including Rabin's paradox (Rabin, 2000), insurance for small losses (Slovic et al., 1977), and preferences for bunched resolution of uncertainty (Kőszegi and Rabin, 2009). The explanations for these phenomena rely on universal loss aversion. Admitting heterogeneity in gain-loss attitudes will lead to more nuanced predictions in each of these settings. Future work on these phenomena is now equipped with a methodology for investigating and controlling for the influence of heterogeneity in gain-loss attitudes.

The manuscript proceeds as follows. In Section 2, we discuss our two-stage labor supply experiment ($N = 500$), building upon the original designs of Abeler et al. (2011) and Gneezy et al. (2017). In Section 3, we discuss our two-stage exchange experiment ($N = 1024$), building on the designs of Cerulli-Harms et al. (2019), Ericson and Fuster (2011), and Heffetz and List (2014). Section 4 provides additional discussion and concludes.

# 2 Labor supply experiment

## 2.1 Experimental design

The labor supply experiment consists of two stages. In Stage 1, we present subjects with a number of decisions that elicit how much effort they are willing to provide at various piece rates, both fixed and uncertain. The objective is to recover each individual's gain-loss attitudes. In Stage 2, we present subjects with a set of choices that manipulate the implied expectations-based reference point while holding other potential reference points constant, constituting a test of the EBRD formulation.

**Stage 1: Measuring Gain-Loss Attitudes.** Subjects were informed about the experiment's various parts and the task they would be asked to complete—transcribing a row of blurry Greek text.[5] They went on to complete two practice tasks to familiarize themselves with the process.

Next, subjects used a slider to indicate how many of these transcription tasks they were willing to complete at a given piece rate. They were shown the earnings associated with a

[5] The task is borrowed from Augenblick and Rabin (2018).

given number of tasks, as well as an estimate of the corresponding completion time. Each piece rate offering was either fixed, e.g., $w = \$0.20$ per completed task, or uncertain, e.g., a 50% chance of $w_h = \$0.30$ per task and a 50% chance of $w_l = \$0.10$ per task. Subjects made decisions for a total of 30 piece rates, 10 of which were fixed. Each uncertain piece rate was linked to a fixed piece rate with the same mean, i.e., $0.5w_h + 0.5w_l = w$. We rely on these two types of piece rates to identify gain-loss attitudes for each individual accounting for auxiliary parameters such as the shape of their cost function.

On each decision screen, subjects made choices for five different piece rates. On a given decision screen, all offered piece rates were fixed, or all were uncertain. Subjects completed a total of six decision screens which appeared in random order. Similar to Augenblick and Rabin (2018), we selected (expected) piece rates between \$0.05/task and \$0.3/task (an hourly wage rate between approximately \$4.00 and \$26.00, according to their average time of completion).

**Stage 2: Experimental manipulation of expectations.** After completing the Stage 1 choices, we informed subjects that they would make two additional effort decisions with slightly different earnings structures. In these additional decisions, subjects were informed that there would be a 50% chance of a per task piece rate of \$0.20, a $p$% chance that a fixed payment \$20 would be paid regardless of the number of completed tasks, and a $q$% chance that a fixed payment \$0 would be paid regardless of the number of completed tasks.[6] Subjects chose a number of tasks to complete in two conditions: Condition Low, where $p = 0.05$ and $q = 0.45$; and Condition High where $p = 0.45$ and $q = 0.05$. Each subject made both decisions in different screens, which were displayed in random order.

In both conditions subjects received a piece rate with 50% chance. With complementary chance, their earnings were unrelated to the number of tasks completed, and were either Low or High in expectation across the two conditions.[7] Within EBRD models, the Low and

[6] These instructions remained purposefully vague about the amounts of money involved as well as any variation over the two choices because our aim was to obtain within-individual comparisons.

[7] This structure allows us to study both within-subject treatment effects by comparing a given subject's answers across conditions and between-subject treatment effects by restricting the sample to only the first

High conditions induce different expectations of earnings and so induce different reference points. This, in turn, leads to different willingness to work across the two conditions. In the neoclassical model and in models with backwards looking reference points, this manipulation should have no effect on optimal choice.

**Lottery elicitation, incentives, and questionnaire.** Following the real-effort decisions, subjects evaluated two risky lotteries using Multiple Price Lists (MPLs), a common elicitation technique to measure gain-loss attitudes in the monetary lottery domain. Subjects made a total of 42 monetary lottery choices in two probability equivalent tasks (following Sprenger, 2015) in which we held fixed a sure payoff of \$5 [\$0] and offered the lottery $(p, \$10; 0)$ [or $(p, \$3; -\$3.5)$] with $p$ ranging from 0% to 100% in increments of 5% as the alternative.[8]

Both the labor supply and lottery choices were incentivized. The experimental earnings were based on one of the 32 effort choices or the 42 monetary lottery choices, with each choice having the same chance of being randomly selected to be the *decision-that-counts*. Regardless of which decision or how many tasks were selected, each subject had to complete a minimum of 10 transcriptions. If the decision-that-counts was one of the monetary lottery choices, the computer generated a random number and determined the outcome of the lottery, and the subjects received their payment upon completion of the mandatory tasks and an ensuing survey. If one of the effort decisions was selected for payment, subjects first completed the mandatory 10 tasks and then the additional number they indicated in that decision; if the relevant rate was stochastic, uncertainty in wages was not resolved until after they had completed all of the additional tasks.[9]

condition subjects saw. We pre-registered predictions about within-subject treatment effects. Appendix Table A3 provides the between-subject results for comparison. Quite similar results are obtained regardless of the method of analysis.

[8] Assuming subjects have monotonic preferences over money—e.g., they prefer \$5 for sure to a 0% chance of \$10 and prefer a 100% chance of \$10 to \$5 for sure—the $p$ at which they switch from preferring one option to another informs us about their gain-loss attitudes.

[9] All subjects had been informed of this procedure in the instructions.

After all the tasks were completed, subjects were presented with a series of Raven's matrices (John Raven and Jean Raven, 2003) to obtain a measure of cognitive skill, followed by a demographic survey (gender, major, age, parental income, and risk attitudes).

**Procedures and pre-registration.** Our sample for the labor supply experiment consists of 500 subjects recruited through the UC San Diego Economics Laboratory. The experiment was pre-registered at the AEA RCT Registry (Campos-Mercade et al. 2021, AEARCTR-0007277) and conducted between April and July 2021. On average, subjects earned \$15.5. The experiment was implemented in *oTree* (Chen et al., 2016). A full set of decision screenshots is provided in Appendix C.

## 2.2 Identifying Gain-Loss Attitudes and Heterogeneous Theoretical Predictions

We derive theoretical predictions of the leading Kőszegi and Rabin (2006, 2007) EBRD model in the labor supply context.[10] We assume that subjects' utility functions are represented by

$$u_i(w, e|r_w, r_e) = m(we) - c_i(e) + \mu_i(m(we) - m(r_w)) + \mu_i(c_i(e) - c_i(r_e)).$$

The first component of utility, $m_i(we) - c_i(e)$, is standard consumption utility obtained from working $e$ tasks and earning $we$. Consumption utility is complemented with a reference-dependent, psychological component of utility, for which the utility from realized earnings $m_i(we)$ is compared to the utility of reference-point earnings $m_i(r_w)$ under a

[10]Throughout, our theoretical analysis will use the Kőszegi and Rabin (2006, 2007) formulation. An earlier literature also provided formulations of reference dependence grounded in rational expectations, but without the equilibrium concepts we use to analyze behavior (Bell, 1985; Loomes and Sugden, 1986).

piece-wise linear gain-loss function $\mu_i$, where

$$\mu_i(z) = \begin{cases} \eta z & z \geq 0 \\ \eta \lambda_i z & z < 0 \end{cases} .$$

Intuitively, if an outcome falls short of the reference point by a difference of $z$, this leads to a reduction of utility by $\eta\lambda_i$ times this difference. An outcome that exceeds the reference point increases utility by $\eta$ times the difference, where $\eta > 0$. Thus, $\lambda_i$ represents individual gain-loss attitude and can either exhibit loss-aversion where losses are felt more severely than commensurate gains, $\lambda_i > 1$, or gain-seeking where gains are felt more severely than commensurate losses, $\lambda_i < 1$. If $\lambda_i = 1$, the individual is considered "loss-neutral". Throughout the analysis, we assume that $m(we) = we$ and constant for all individuals, that $c_i(e)$ is an increasing at least twice-differentiable convex function, and normalize $\eta = 1$ for all individuals.

Kőszegi and Rabin (2006, 2007) propose that agents hold the entire distribution of expected outcomes as their referent. Each potential realization is compared to each potential reference point and weighted by the relevant densities. In order to close the model, Kőszegi and Rabin (2006, 2007) equip it with the rational expectations Choice-Acclimating Personal Equilibrium (CPE) concept. Intuitively, a choice is a CPE if the agent's expected utility from this choice given their expectation of this choice as the referent exceeds the expected utility of any alternative choice given the expectation of that alternative choice as the referent. We consider the CPE identification (and estimation) of gain-loss attitudes in Stage 1 of our experimental design, and the CPE comparative statics in Stage 2 of our experimental design.

### 2.2.1 Stage 1 Estimates of Gain-Loss Attitudes

Consider an uncertain piece rate condition in Stage 1, $(0.5, w_l; 0.5, w_h)$, $w_h > w_l$. The individual chooses effort, $e_i$, knowing that with 50% chance they will earn either $e_i \times w_l$ or

$e_i \times w_h$. The associated CPE utility for such an effort choice, $e_i$, is

$$0.5w_l e_i + 0.5w_h e_i - 0.25(\lambda_i - 1)(w_h e_i - w_l e_i) - c_i(e_i).$$

If the individual faces a fixed piece rate, $w$, then CPE utility reduces to

$$u(we_i|we_i) = we_i - c_i(e_i),$$

In choosing a functional form for the cost of effort, our pre-registered analysis follows Augenblick and Rabin (2018) by assuming $c_i(e_i) = \frac{1}{\alpha_i \gamma_i}(e_i + 10)^{\gamma_i}$, where 10 represents the required minimum number of tasks that all subjects must complete.[11] Note that this formulation permits individual variation in $\gamma_i$ and $\alpha_i$, the parameters of the cost function.

The optimal effort choice, $e^*_{i,U}$, with uncertain piece rates thus satisfies

$$0.5w_l + 0.5w_h - 0.25(\lambda_i - 1)(w_h - w_l) = \frac{1}{\alpha_i}(e^*_{i,U} + 10)^{\gamma_i - 1}. \tag{1}$$

Similarly, in the fixed piece rate setting, the optimal effort choice, $e^*_{i,F}$ satisfies

$$w = \frac{1}{\alpha_i}(e^*_{i,F} + 10)^{\gamma_i - 1}. \tag{2}$$

These two equations provide an intuitive formulation for identifying gain-loss attitudes. If the individual has $\lambda_i = 1$, then any two conditions with $0.5w_l + 0.5w_h = w$ will deliver identical choices. Loss neutral individuals with $\lambda_i = 1$ are invariant to such mean-preserving wage spreads. Loss-averse individuals with $\lambda_i > 1$ will respond to mean-preserving uncertainty by exhibiting $e^*_{i,U} < e^*_{i,F}$, as $\lambda_i > 1$ lowers the left hand side of (1) relative to the left hand side of (2). Conversely, gain-seeking individuals with $\lambda_i > 1$ will

[11] As Augenblick and Rabin (2018) point out: "The parameter $\alpha$ is necessary and represents the exchange rate between effort and the payment amount. If instead $c_i(e_i) = \frac{1}{\gamma_i}(e + 10)^{\gamma_i}$, a requirement such as linear marginal costs (which necessitates $\gamma_i = 2$), would also imply that the marginal cost of $e_i$ tasks is exactly $e_i$ monetary units, regardless of the task type or the payment currency."

respond to mean-preserving uncertainty by exhibiting $e^*_{i,U} > e^*_{i,F}$ as $\lambda_i < 1$ increases the left hand side of (1) relative to the left hand side of (2). Hence, the sensitivity of effort choice to uncertain wage spreads identifies gain-loss attitudes.[12]

This simple intuition on identification motivates a reduced form measure of gain-loss attitudes. Specifically, we consider the individual regression

$$log(e_i + 10) = C + g_i log(\overline{w}) - l_i log(1 + \Delta w) + \epsilon_i. \tag{3}$$

The variable $\overline{w} = 0.5w_l + 0.5w_h$ for uncertain piece rates, and $\overline{w} = w$ for fixed piece rates. The variable $\Delta w = w_h - w_l$ for uncertain piece rates, and $\Delta w = 0$ for fixed piece rates. The regression estimate, $\hat{l}_i$, captures the negative of the elasticity of labor supply to wage uncertainty, and so should closely correspond to the theoretical quantity $\lambda_i$.

In order to provide a structural estimate of the parameter, $\lambda_i$, we conduct the non-linear regression corresponding to the log-transformed marginal conditions (1) and (2) with additive shocks

$$log(e_i + 10) = \frac{1}{\gamma_i - 1} log(\alpha_i \left[\overline{w} - 0.25(\lambda_i - 1)\Delta w\right]) + \epsilon_i, \tag{4}$$

[12] Our formulation assumes that utility of money, $m(\cdot)$, is linear. If individuals had diminishing marginal utility of money, one would expect a potential deviation between $e^*_{i,U}$ and $e^*_{i,F}$ even if $\lambda = 1$. Indeed if $m(\cdot)$ were concave the optimal responses with $\lambda = 1$ would be calculated from marginal conditions

$$m'(we^*_{i,F})w = \frac{1}{\alpha_i}(e^*_{i,F} + 10)^{\gamma_i - 1}$$

and

$$0.5m'(w_l e^*_{i,U})w_l + 0.5m'(w_h e^*_{i,U})w_h = \frac{1}{\alpha_i}(e^*_{i,U} + 10)^{\gamma_i - 1}.$$

These two values will differ to the extent that marginal utility changes over the range $[w_l * e, w_h * e]$. For values of $e$ around 40 tasks and a range of $w_h - w_l \approx 0.1 - 0.2$ this corresponds to a \$4-8 range. Changes in marginal utility over such ranges would have to be dramatic to deliver perceptible effects on behavior and would deliver calibrational implausibilities at larger stakes. Moreover, if one were to attribute differences between $e^*_{i,U}$ and $e^*_{i,F}$ to changes in marginal utility, one would predict null effects (and no heterogeneity) in Stage 2 of our design.

where $\overline{w}$ and $\Delta w$ are defined as above. Because the data are potentially censored at $e_i = 0$ and $e_i = 100$, we use a maximum likelihood tobit method.[13] The product of such an exercise is an estimated triple, $(\hat{\alpha}_i, \hat{\gamma}_i, \hat{\lambda}_i)$, capturing gain-loss attitudes alongside auxiliary parameters of the individual's cost function.

Individual estimates of gain-loss attitudes, $\hat{\lambda}_i$, are likely to be estimated with error. Appendix A.2 develops a standard Bayesian shrinkage exercise leveraging distributional information on all $\hat{\lambda}_i$ and the estimated errors, $\hat{\sigma}_{\lambda_i}$. This exercise, effectively a random-effects meta-analysis on our data, maps from an individual's value of $\hat{\lambda}_i$ and $\hat{\sigma}_{\lambda_i}$ to an expected value $E[\hat{\lambda}_i]$. Intuitively, the method takes into account the population mean of all estimates and the standard error of each individual's estimate: estimates with a high standard error are imprecise and carry little information. The outcome of this exercise is that imprecise estimates of $\hat{\lambda}_i$ are shrunk to the sample average in proportion to their imprecision. Whether we use reduced form or structural measures, shrunk to account for measurement error or not, our results are effectively unchanged.

### 2.2.2 Heterogeneous Effects of Stage 2 Low vs. High Conditions

We now consider how individuals behave when offered an earnings structure $(p, H; q, L; 0.5, w)$ where $L < H$; that is, individuals have a 50% chance of earning a piece-rate, $w$, per unit

[13] In order to arrive at useful starting values, we first estimate equation (4) on only the fixed piece rate data, which eliminates $\lambda_i$ from estimation, with common starting values for $\alpha_i$ and $\gamma_i$. We then use resulting estimates for $\alpha_i$ and $\gamma_i$ as starting values for the combined fixed piece rate and uncertain piece rate data set. Anyone for whom the first estimation step fails to converge, we retain the original starting values. Additionally, our implementation placed box constraints on the parameters $\gamma_i \in (1, 4)$ and $\lambda_i \in (0, 3)$ by estimating the parameters $g_i$ and $l_i$ such that $\gamma_i = 1 + 3(1/(1 + exp(g_i)))$ and $\lambda_i = 3(1/(1 + exp(l_i)))$. Parameters $\gamma_i$ and $\lambda_i$ along with their standard errors were recovered via the delta method. The restriction on $\gamma_i$ requires costs to be convex, but not overly so. The restriction on $\lambda_i$ ensures reasonable bounds; $\lambda_i > 3$ is ruled out under CPE since it has unrealistic implications—including violations of First Order Stochastic Dominance (see Masatlioglu and Raymond 2016 for more details). Operationally, allowing $\lambda_i \geq 3$ generates additional problems as it leads to undefined allocation values (a root of a negative number) unless $\gamma_i = 2$. Our procedure began with the starting values $\alpha_i = 594$, $g_i = -2$, $l_i = 1$ (for the combined data), and a standard deviation of $\epsilon$, $log(\sigma) = 2$. See Appendix A.2 for a more detailed description of the estimation procedure. Appendix A.2 also shows that altering these starting values changes some estimated quantities and the fraction of subjects for whom convergence is achieved, but does not alter any of our general conclusions.

of effort, a $p$% chance of earning \$$H$ regardless of effort, and a $q = (50-p)$% chance of earning \$$L$ regardless of effort. Following the development of Gneezy et al. (2017), we study the effects of an increase in $p$ when $L \leq we_i^* \leq H$.[14] In Appendix A.3 we derive the CPE choice, $e_i^*$, in this case satisfying

$$0.5w\left[1 + (p-q)(\lambda_i - 1)\right] = c_i'(e_i^*), \tag{5}$$

and the effect of increasing the probability of the high outcome, $p$, while keeping $p+q = 0.5$ as

$$\frac{\partial e_i^*}{\partial p}|_{p+q=0.5} = \frac{(\lambda_i - 1)w}{c_i''(e_i^*)}.$$

This effect contrasts with that of alternative models of the reference point, where $\frac{\partial e_i^*}{\partial p}|_{p+q=0.5} = 0$. As the outside possibility unrelated to effort, $(p, H; q, L)$, increases in expectation, KR individuals should change their level of effort. Moreover, the direction of the response is governed by gain-loss attitudes, $\lambda_i$. Following from the equation above,

$$\lambda_i > 1 \implies \frac{\partial e_i^*}{\partial p}|_{p+q=0.5} > 0$$
$$\lambda_i < 1 \implies \frac{\partial e_i^*}{\partial p}|_{p+q=0.5} < 0.$$

In our implementation we set $H = \$20$, $L = \$0$, $w = 0.20$, and vary $p$ from 0.05 in the Low condition to 0.45 in the High condition. This implementation leads to the following theoretical prediction for heterogeneous treatment effects.

**Prediction 1.** Loss-averse individuals ($\lambda_i > 1$) should be more willing to work in the High condition relative to the Low condition. Gain-seeking individuals should be less willing to work in the High condition relative to the Low condition.

[14]For all other rank cases, there is no predicted treatment effect (see Appendix A.3 for details).

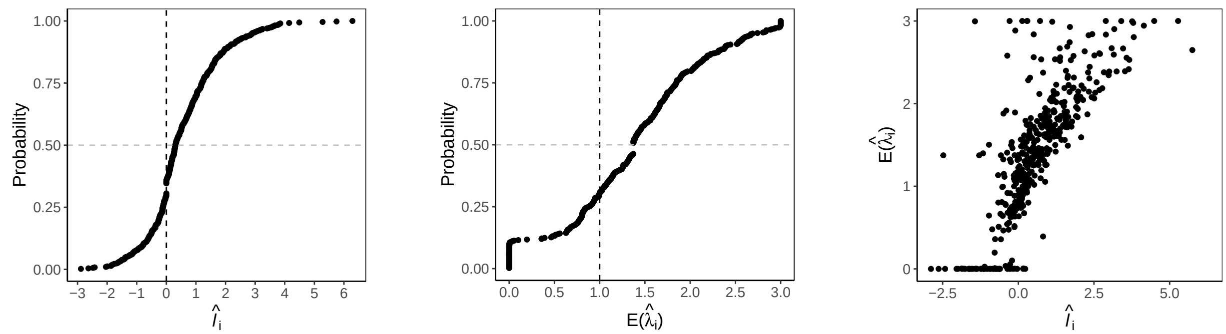
## 2.3 Results From The Labor Supply Experiment

**Stage 1: The distribution of gain-loss attitudes in labor supply.** In Stage 1 our 500 subjects each make 30 effort choices, 10 for fixed piece rates and 20 for uncertain piece rates. In Appendix Table A1, we present the mean, median, and interquartile range for each choice, along with the proportion of observations censored at the extreme allocations of $e_i = 0$ or $e_i = 100$. Overall, subjects exhibit increasing labor supply functions, being willing to complete more tasks for greater fixed piece rates. On average, subjects are willing to complete fewer tasks under uncertain piece rates relative to fixed rates of equal mean. Within the context of our KR analysis, this implies loss aversion on average. Importantly, Appendix Table A1 also documents substantial heterogeneity. At every piece rate, whether fixed or uncertain, the interquartile range covers a wide portion of the choice space. This, in turn, suggests substantial heterogeneity in both costs and gain-loss attitudes.

In order to provide an initial indication on the extent of heterogeneity, we estimate equation (3) for every individual. Figure 1, Panel A plots the distribution of reduced form gain-loss attitudes, $\hat{l}_i$, capturing the elasticity of labor supply with respect to wage uncertainty. Of the 500 subjects, 65.4% exhibit $\hat{l}_i < 0$, indicating reduced-form loss aversion, while 30.6% exhibit $\hat{l}_i > 0$, indicating reduced-form gain seeking.

While our reduced-form approach is estimable on all study subjects, our structural maximum-likelihood approach yields estimates of $\hat{\lambda}_i$ and the standard error $\hat{\sigma}_{\lambda_i}$ for a subset 451 of 500 subjects (90%).[15] Without adjusting for measurement error, the average value of $\hat{\lambda}_i$ is 1.31, the standard deviation is 0.88, and 58.76% have $\hat{\lambda}_i > 1$. Deploying the shrinkage adjustment described in Appendix A.2, we find an average value of $E[\hat{\lambda}_i] = 1.37$, with a standard deviation of 0.77, and 69.2% have $E[\hat{\lambda}_i] > 1$. Figure 1, Panel B plots the distribution of structural gain-loss attitudes accounting for shrinkage, $E[\hat{\lambda}_i]$.

[15] Our estimation method converges and delivers an estimate of $\hat{\lambda}_i$ for an additional 31 subjects but without standard errors indicating the extent of imprecision. Without an indication of imprecision we cannot conduct the shrinkage adjustment proposed. For a total of 334 subjects, we estimate quantities and standard errors for all parameters. In Appendix Table A2 we show that our results are robust to using each of these samples in turn with unadjusted or shrinkage adjusted estimates of gain-loss attitudes.




(a) CDF of reduced form estimates

(b) CDF of structural estimates

(c) Reduced form and structural estimates

Figure 1: Stage 1: Gain-loss attitudes in the labor supply experiment

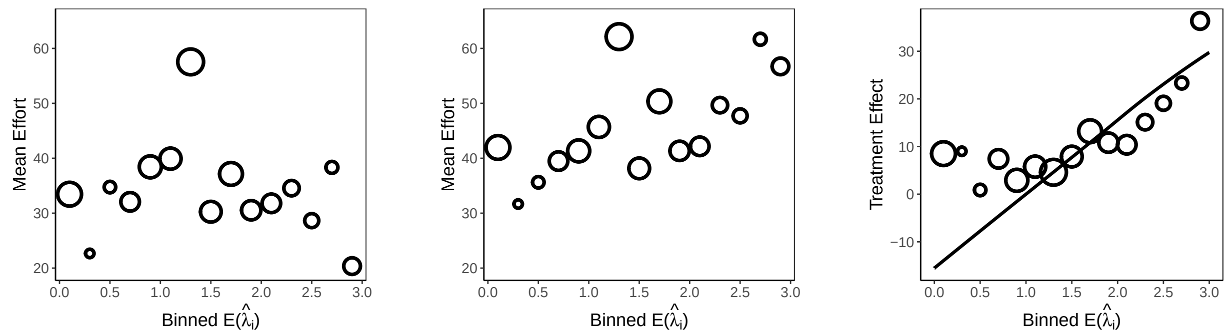
*Notes:* Panel (a) and (b) show CDFs of the reduced form and structural measures of gain-loss attitudes, respectively. Note that gain-seeking behavior corresponds to values of $\hat{l}_i < 0$ in the reduced form but $E[\hat{\lambda}_i] < 1$ in the structural estimates. Based on $N = 451$ observations. Panel (c) displays the relationship between both measures (Pearson's $r = 0.79,\ p < 0.01$)

In both the reduced form and structural measures of gain-loss attitudes, approximately 70% of subjects are loss-averse, while 30% are gain-seeking. Panel C of Figure 1 documents the correlation between the two measures. The intuitive connection between the elasticity of wage uncertainty and gain-loss attitudes is supported in the data as $E[\hat{\lambda}_i]$ and $\hat{l}_i$ have a Pearson's $r = 0.79$ ($p < 0.01$).

**Stage 2: Heterogeneous treatment effects of Low vs. High.** With our estimates of each subject's gain-loss attitude in hand, we analyze effort choice across the Stage 2 conditions, Low vs. High.[16]

Figure 2 illustrates the relationship between Stage 1 measures of gain-loss attitudes, $E[\hat{\lambda}_i]$, and Stage 2 behavior. We construct fifteen equally spaced bins of $E[\hat{\lambda}_i]$ and calculate the average behavior in each bin. Panel A illustrates a slight negative relationship between $E[\hat{\lambda}_i]$ and effort in the Low condition; more loss-averse subjects choose lower levels of effort. In contrast, Panel B illustrates a substantial positive relationship between $E[\hat{\lambda}_i]$ and effort

[16]Our main (pre-registered) analysis exploits the within feature of the experiment, leveraging each subject's answers to both Condition Low and Condition High. Appendix Table A3 shows that the point estimates are similar when we only use data from the first condition that subjects saw.




(a) Effort in Condition Low (b) Effort in Condition High (c) Treatment effect

Figure 2: Stage 2: Heterogeneous treatment effects in the labor supply experiment

*Notes:* Panel (a) shows average effort across 15 bins of $E[\hat{\lambda}_i]$ in Condition Low ($N = 451$). Panel (b) shows average effort across 15 bins of $E[\hat{\lambda}_i]$ in Condition High ($N = 451$). Panel C provides the treatment effect, corresponding to the differences between Panels (a) and (b), as well as the KR CPE prediction for subjects with different $E[\hat{\lambda}_i]$. The bubble size indicates the number of subjects in the corresponding bin.

in the High condition; more loss-averse subjects choose higher levels of effort. Panel C presents the relationship between individual treatment effects, i.e. the difference between conditions, and $E[\hat{\lambda}_i]$. Greater values of $E[\hat{\lambda}_i]$ in Stage 1 correspond with greater treatment effects in Stage 2.

Alongside the empirical relationship between treatment effects and gain-loss attitudes, Figure 2 Panel C presents EBRD predictions for heterogeneous treatment effects. This prediction is generated to indicate the heterogeneity in response associated with gain-loss attitudes. At the median value of the estimated $\alpha_i = 609$ and $\gamma_i = 2.26$, we calculate a predicted treatment effect for each individual at their specific value of $E[\hat{\lambda}_i]$ based on (5) for the High and Low conditions.[17] EBRD predicts substantial heterogeneity in treatment effects. As the outside payment increases in expectation, loss-averse individuals should grow more willing to work, while gain-seeking individuals should grow less so. The data are consistent with these predictions. Recall that alternate models of reference dependence

[17]This formulation associates all predicted heterogeneity with heterogeneous gain-loss attitudes and ignores any potential correlation between the gain-loss attitudes and parameters of the cost-function for the purpose of prediction.

predict zero treatment effect in this experimental design, and, moreover, zero heterogeneity therein.

Table 1: Heterogeneous treatment effects in the labor supply experiment

| | *Dependent Variable: Effort* | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Condition High | 10.412 | 7.903 | 0.612 |
| | (1.040) | (1.095) | (2.340) |
| Reduced form ($\hat{l}_i$) | | -3.269 | |
| | | (1.053) | |
| Condition High * Reduced form ($\hat{l}_i$) | | 4.449 | |
| | | (1.119) | |
| $E[\hat{\lambda}_i]$ | | | -2.306 |
| | | | (1.945) |
| Condition High * $E[\hat{\lambda}_i]$ | | | 6.789 |
| | | | (1.740) |
| Constant (Condition Low) | 35.742 | 37.585 | 40.126 |
| | (1.422) | (1.616) | (3.017) |
| R-Squared | 0.025 | 0.034 | 0.030 |
| # Observations | 1000 | 1000 | 902 |
| $H_0$ : Zero Treatment Effect (H-L) | $F_{1,499} = 100.23$ | $F_{1,499} = 52.08$ | $F_{1,450} = 0.07$ |
| | $(p < 0.01)$ | $(p < 0.01)$ | $(p = 0.79)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Effort in Low | | $F_{1,499} = 9.64$ | $F_{1,450} = 1.41$ |
| | | $(p < 0.01)$ | $(p = 0.24)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,499} = 15.81$ | $F_{1,450} = 15.23$ |
| | | $(p < 0.01)$ | $(p < 0.01)$ |

*Notes*: Ordinary least squares regression explaining each subject's effort choice. Each subject provides two observations: one with their effort in Condition Low, and one with their effort in Condition H. Clustered standard errors at the individual level in parentheses. The "Reduced form ($\hat{l}_i$)" measure captures the negative of the elasticity of labor supply to wage uncertainty, as estimated by $l_i$ in equation 3. Null hypotheses tested for 1) zero treatment effect (Condition High coefficient= 0); 2) no relationship between gain-loss attitudes and behavior in Condition Low behavior ($E[\hat{\lambda}_i]$ or $\hat{l}_i = 0$); 4) constant treatment effect over gain-loss attitudes (Condition High* $E[\hat{\lambda}_i]$ or Condition High* $\hat{l}_i = 0$). $F$-statistics and two-sided $p$-values reported.

Complementing the visual illustration in Figure 2, we report regression results in Table 1. Column 1 shows an aggregate specification, wherein treatment effects are assumed to be homogeneous. In the Low condition, subjects choose around 36 tasks, while in the High condition they choose around 46 tasks. This aggregate treatment effect of approximately

10 tasks is significant at all conventional levels, $F_{1,499} = 100.23$ ($p < 0.01$). This observed aggregate treatment effect is in the same direction as what would be implied by EBRD under universal loss aversion: subjects should grow more willing to work as the reference point increases in expectation. Importantly, the average treatment effect in column 1 potentially aggregates differential treatment effects for loss-averse and gain-seeking subjects. In columns (2) and (3) we interact treatment with reduced form and structural measures of gain-loss attitudes for the relevant samples. Both measures are highly correlated with the effect of treatment and in both specifications we reject the null hypothesis of zero heterogeneity in treatment, $F_{1,499} = 15.81$ ($p < 0.01$) and $F_{1,450} = 15.23$ ($p < 0.01$), respectively. More loss averse subjects have greater increases in willingness to work as they move from the Low to the High condition. These results are notably supportive of the EBRD formulation of the reference point: individuals respond to the change in expectations across Low and High conditions, and differentially so depending on their gain-loss attitudes. Alternative formulations of the reference point predict zero treatment effect and zero heterogeneity therein, and, thus, are rejected by our labor supply results.

Without accounting for heterogeneous gain loss attitudes, the average treatment effect reported in column 1 combines the potentially different-signed effects of loss-averse and gain-seeking subjects. Aggregating such different signed effects leads to two potential analytical issues. First, if theoretical treatment effects are non-linear in $\lambda_i$, then the average treatment effect is not necessarily equal to the treatment effect at the average parameter value. In Appendix A.4, we show that in the setting of labor supply the relationship between $\lambda_i$ and treatment differences is governed by the shape of the marginal cost function, $c_i'(e)$. Intuitively, if marginal costs are linear in equation (5), then $e_i^*$ in each condition and the corresponding treatment effects are linear in $\lambda_i$. And, the shape of marginal costs governs the potential for non-linear aggregation. As our median estimate of $\gamma_i = 2.26$ implies approximately linear marginal costs, this first aggregation issue does not have a great influence in our labor supply setting. As shown in Figure 2, theoretical treatment effects are nearly linear in gain-loss attitudes. We emphasize that this is not a general result, but does help to explain why the average treatment effect is positively signed and consistent with the average preference in our labor supply experiment.

Second, even if the average treatment effect is equal to the treatment effect of the average preference, the standard deviation of treatment effects will be influenced by heterogeneity. If gain-loss attitudes were universal, all individuals would have the same treatment effect in theory. Variation across individuals could derive from noise or heterogeneity in other parameters, and examining the treatment effect would require appropriately powering for the extent of variation expected. If individuals are heterogeneous in gain-loss attitudes, appropriately powering an examination of the average treatment effect grows more challenging: the standard deviation of expected behavior grows and the required sample size grows for well-powered designs. In Appendix A.4, we explore these issues of power and heterogeneity in detail.

**Gain-loss attitudes across domains.** Prior work has documented linkages between gain-loss attitudes measured with and without risk, coupling measures of small-stakes risk aversion with exchange behavior in standard endowment effect experiments (see, e.g., Dean and Ortoleva, 2015; Gächter et al., 2022). This work documents sizeable correlations between different measures, ranging from 0.3 to 0.6.

Appendix Figure A3 provides the distribution of gain-loss attitudes calculated using CPE from subjects' lottery choices. The mean and median $\lambda$ are 1.58 and 1.5, respectively. As in the labor supply setting, we find substantial heterogeneity in gain-loss attitudes across subjects. We classify a sizable minority of 25 percent as gain-seeking. We find that gain-loss attitudes estimated from lottery choices are correlated with the structural estimates of gain-loss attitudes based on labor supply decisions, but not dramatically so (Pearson's $r = 0.090,\ p = 0.057$; Spearman's $\rho = 0.098,\ p = 0.037$). And, we find that our lottery measure of gain-loss attitudes has no predictive power for treatment effects in Stage 2. These findings suggest that though heterogeneity is similar across domains, gain-loss attitudes at the individual level are potentially more domain-specific than previously thought.

# 3 Exchange experiment

## 3.1 Experimental Design

The basic structure of the exchange experiment closely follows that of the labor supply experiment. Stage 1 serves to elicit gain-loss attitudes at the individual level. Stage 2 features a manipulation of expectations adapted to the exchange setting.

**Stage 1: Measuring gain-loss attitudes.** At the beginning of the experiment, subjects saw equally-sized pictures and descriptions of two objects. They were then randomly assigned a private cubicle in which they found one of the two objects. We informed them that the object in front of them was in their possession, and that they were invited to inspect it more closely.[18] After three minutes allotted for inspection of the object, we asked subjects three questions. First, for each object subjects were asked "How much do you like this object?" with a Likert response scale ranging from 0="Not at all" to 8="Very much". Second, for each object they were asked "How much would you want to have this product?" using the same response scale. Third, they were asked "If you had to choose one of the objects, which one would you prefer to keep?". These three unincentivized preference statements are the raw data from which our estimates of gain-loss attitudes are constructed.

After subjects provided their preference statements, the experimenter randomly selected half of all subjects in the session based on a draw from a lotto drum that was clearly visible to all subjects. The experimenter replaced the endowed good with the alternative good for each of the selected subjects. This random replacement of Stage 1 objects was conducted

[18]Crucially, we did not say that they "own" the object, and we asked them to not remove the packaging yet.

to provide subjects with an experience of probabilistic exchange and to generate exogenous variation in the objects obtained in Stage 1.

**Stage 2: Experimental manipulation of expectations.** The procedures in Stage 2 were purposefully similar to those in Stage 1. In a separate room, subjects saw pictures and descriptions of two different objects. Upon returning to their private cubicle they would find one of the two objects, which we again assigned randomly. We study two between-subjects conditions, with randomization at the session level.[19] In both conditions, subjects decide whether they would like to retain their assigned object or exchange it. The two conditions differ in the probability that exchange will be forced regardless of their statement. In the Low condition, subjects face a 0% chance that exchange will be forced. That is, this condition is equivalent to a standard exchange setting common to endowment effect experiments. In the High condition, subjects are forced to exchange their object with 50% chance regardless of choice. The chance of forced exchange was based on a draw from a lotto drum that was visible to all subjects. Within EBRD models, the Low and High conditions induce different expectations of the final object to be obtained and so induce different reference points. This, in turn, leads to different willingness to exchange across the two conditions. In the neoclassical model or models with backwards looking reference points, the probability of forced exchange should have no effect on optimal choice.

**Procedures and pre-registration.** The objects used for the exchange experiment comprise a USB stick, a set of three erasable pens, a picnic mat, and a thermos.[20] We selected these four objects on the basis of a pre-experimental survey evaluation of 12 candidate objects. We put particular emphasis on ruling out complementarities between items across rounds. The former two (USB stick and pens) and the latter two objects (picnic mat and thermos) each constituted a pair. Across the two stages, each subject encountered each

[19] We present our analysis with robust standard errors in the main text and Appendix Tables A10 through A12 reproduce our results with standard errors clustered at the session level. Statistical significance is enhanced with clustering, and so we decided to provide the more conservative values in the main text.

[20] Pictures and information presented to subjects are reproduced in Appendix D.

pair of objects exactly once. The use of each pair as the Stage 1 pair was counterbalanced at the session level.

The total sample for the exchange experiment consists of 1024 subjects recruited from the BonnEconLab at University of Bonn in Germany. In total, 59 percent (603 of 1024 subjects) were randomly assigned to Probabilistic Forced Exchange. An initial sample of 607 subjects participated in June and July 2015, and a pre-registered replication sample of a further 417 subjects participated in July 2018 (Goette et al., 2018, AEARCTR-0003124).[21] Subjects received a participation fee of 6 euros and also two of the four objects used in the experiment according to their endowments, choices, and chance. A full set of screenshots for our experiment, implemented in *ztree* (Fischbacher, 2007), can be found in Appendix D.

## 3.2 Estimating Gain-Loss Attitudes and Heterogeneous Theoretical Predictions

We again derive theoretical predictions using the Kőszegi and Rabin (2006, 2007) EBRD model in the exchange setting. We consider the two-dimensional utility function over object $X$ and object $Y$,

$$u_i(\mathbf{c}|\mathbf{r}) = m_X + m_Y + \mu_i(m_X - r_X) + \mu_i(m_Y - r_Y),$$

where $\mathbf{c} = (m_X, m_Y)$ refers to consumption utility associated with the quantity of each object, and $\mathbf{r} = (r_X, r_Y)$ similarly refers to reference utility. Thus, an individual's utility function consists of two components: consumption utility, $m_X + m_Y$, and gain-loss utility, $\mu_i(m_X - r_X) + \mu_i(m_Y - r_Y)$. We let $m_X, r_X \in \{0, X\}$, and $m_Y, r_Y \in \{0, Y\}$ denote both

[21] While the experiment carried in June and July 2015 was not pre-registered, the one carried in July 2018 was pre-registered. In the main body of the paper we pool the results from both experiments, but Appendix Table A9 shows that the main results replicate in both samples. There were a few very minor differences between the original sessions and those in the replication, which are also described in Appendix B.6.

the outcome and the corresponding utility of zero or one unit of object X, and zero or one unit of object Y, respectively. For our primary analysis we assume utilities, $X$ and $Y$ to be homogeneous in the population, but we also investigate heterogeneity in valuations in Appendix B.2.[22] As before, we assume piecewise linear gain-loss attitudes, with potential heterogeneity in loss-aversion or gain-seeking, $\lambda_i$, and $\eta = 1$ for all individuals. We consider the identification (and estimation) of gain-loss attitudes in Stage 1 of our experimental design, and the CPE comparative statics in Stage 2 of our experimental design.

#### 3.2.1 Stage 1 Estimates of Gain-Loss Attitudes

In Stage 1 of our design subjects are explicitly endowed with an object and then asked to provide preference statements about that object and an alternative. These statements are made without knowledge of any possibility of actual exchange. Hence, theoretically, the reference point is fixed at the endowed object.[23] An individual endowed with $X$ will state a preference in the form of a higher liking value for $X$, higher wanting value for $X$, or hypothetical choice of $X$ if $u_i(X, 0|X, 0) - u_i(0, Y|X, 0) > \delta$, where $\delta$ captures the possibility of equal rating levels.[24] Under our functional form assumptions such a preference statement occurs if

$$(1 + \lambda_i) - 2\frac{Y}{X} - \delta_X > 0,$$

where $\delta_X \equiv \frac{\delta}{X}$. Similarly, an individual endowed with $X$ would state a preference for $Y$ if

$$2\frac{Y}{X} - (1 + \lambda_i) - \delta_X > 0.$$

[22]The exercise elaborated in Appendix B.2 assumes homogeneous gain-loss attitudes and heterogeneous valuations as the source of variation in behavior in Stage 1. This formulation is clearly rejected by the heterogeneous treatment effects observed in our exchange study.

[23]Though implausible given our design, potential alternative formulations might be to assume that subjects believe they can change their reference point from $X$ to $Y$ or to assume subjects consider retaining their endowed object, $X$, *and* gaining the alternative, $Y$ (evaluating utility of $Y$ as $X + (1 + \eta)Y$). Importantly, both of these formulations would imply that Stage 1 statements reveal no information on gain-loss attitudes, $\lambda_i$. Hence, both would yield null predictions for heterogeneous treatment effects in Stage 2. As such, the results we document invalidate these formulations.

[24]Note that $\delta = 0$ for our hypothetical choice data as there was no possibility of stating indifference.

An individual would state equal preferences if neither inequality were satisfied. These two equations provide an intuitive formulation for identifying gain-loss attitudes. Controlling for the relative utility of the two objects, $\frac{Y}{X}$, an individual with a greater value of $\lambda_i$ should be more likely to prefer their endowment and less likely to prefer the alternative.

This simple intuition on identification motivates a reduced-form measure of gain-loss attitudes based on residual preference for endowed objects. First, we conduct a principal components analysis on the three preference statements in Stage 1 and reduce the data to the first principal component. Within our data the first component captures around 70 percent of the variation in relative wanting, relative liking, and hypothetical choice statements. We then regress this component on Stage 1 object assignment. The residuals of this regression summarize a residual preference for the endowed or the alternative object accounting for the average preference. An individual who disproportionately likes their assigned object relative to average preferences is plausibly more loss averse than one who exhibits a residual in the opposite direction. Hence, we consider these residuals as a reduced form measure of gain-loss attitudes, $\hat{l}_i$.

Residual preference for assigned objects could also partially reflect heterogeneity in the intrinsic utilities, $\frac{Y}{X}$. Because subjects are assigned new objects in Stage 2, heterogeneity in $\frac{Y}{X}$ in Stage 1 is orthogonal to any subsequent treatment effects. Hence, the interpretation of Stage 1 measures as being driven by heterogeneity in $\frac{Y}{X}$ is rejected by the heterogeneous treatment effects observed in Stage 2.

In order to provide a structural estimate of the parameter, $\lambda_i$, we make a number of assumptions. First, rather than assuming deterministic choice, we posit that an individual endowed with $X$ will state a relative preference for $X$ with probability

$$\pi_{X|X} = Prob((1 + \lambda_i) - 2\frac{Y}{X} - \delta_X > \epsilon),$$

a relative preference for $Y$ with probability

$$\pi_{Y|X} = Prob(2\frac{Y}{X} - (1+\lambda_i) - \delta_X > \epsilon),$$

and, where appropriate, would provide equal ratings for the two objects with probability $\pi_{E|X} = 1 - \pi_{X|X} - \pi_{Y|X}$. Symmetric formulations are assumed for individuals endowed with object $Y$. Within this structure, $\epsilon$ can be interpreted as capturing idiosyncratic variation in the $\frac{Y}{X}$ parametrically, or noise in response. We assume $Prob(\cdot)$ is the logistic function leading to logit choice. Second, we assume that $\lambda_i$ is drawn from a log-normal distribution with $log(\lambda_i) \sim N(\mu_\lambda, \sigma^2_\lambda)$, leading to a mixed logit formulation. Third, we assume the deterministic portion of relative utility, $\frac{Y}{X}$, is homogeneous in the population, and a parameter to be estimated. And fourth, we assume $\delta_X = 0.55$, a value that our prior research indicated to be an appropriate aggregate value.[25] These assumptions permit us to estimate the parameters of the distribution of gain-loss attitudes $N(\hat{\mu}_\lambda, \hat{\sigma}^2_\lambda)$ based on Stage 1 data.[26]

Moving from the estimated distribution of gain-loss attitudes to an expected value of $E[\hat{\lambda}_i]$ for each individual is a straightforward step. As proposed in Train (2009), from $N(\hat{\mu}_\lambda, \hat{\sigma}^2_\lambda)$ we simulate the distribution of $\hat{\lambda}_i$ and the corresponding distributions of preference statements. We then calculate the expected simulated value, $E[\hat{\lambda}_i]$, for each possible Stage 1 collection of preference statements. This exercise of mapping from preference statements to a conditional expectation of gain-loss attitudes takes into account the possibility

[25]See Appendix B.6 for these prior estimates. We found some sensitivities of the value $\sigma^2_\lambda$ to attempting to estimate $\delta_X$ alongside the other parameters. The challenge is intuitive: a larger value of $\delta_X$ implies individuals should more frequently give the two objects equal ratings. All else equal, a higher variance of gain-loss attitudes is required to justify the relative infrequency of such observations. Appendix Table A6 provides analysis setting $\delta_X$ at several different values and demonstrating corresponding sensitivity for the variance of gain-loss attitudes.

[26]It is also straightforward to alter the assumptions of this formulation to estimate heterogeneity in intrinsic utilities, $\frac{Y}{X}$, rather than gain-loss attitudes. Such an exercise is presented in Appendix B.2, and yields estimates of aggregate loss aversion and substantial variation in object valuations. As noted above, interpreting Stage 1 measures as being driven by heterogeneous utilities rather than heterogeneous gain-loss attitudes leads to the prediction of no heterogeneous treatment effects in Stage 2, and thus is rejected by the data.

of noise as the preference statements are simulated assuming logit errors.[27] Appendix B.1 provides additional details and Appendix Table A7 provides examples of the corresponding mappings from preference statements to $E[\hat{\lambda}_i]$.

### 3.2.2 Heterogeneous Effects of Low vs. High Conditions

Consider the Low condition, in which subjects are asked whether endowed with object $X$ they prefer $X$ or $Y$. In this setting, the two potential CPE selections are $\{(X, 0), (0, Y)\}$ (the first reflecting a choice not to exchange, and the second the choice to exchange). The individual can support not exchanging in a CPE if $u_i(X, 0|X, 0) \geq u_i(0, Y|0, Y)$. Given our assumptions, this condition is equivalent to the threshold $X_{L,i} \geq Y$, i.e., the agent keeps their endowed object if it has weakly greater consumption utility than the alternative object.[28]

Next, consider the environment in the High condition. With probability 0.5, the agent, assumed endowed with $X$, will be forced to exchange $X$ for $Y$ regardless of their choice. If the individual wishes to retain their object, they are subject to a stochastic reference point, as with probability 0.5 their object will be exchanged regardless of their choice. Now, the potential CPE selections for someone endowed with $X$ are $\{0.5(X, 0) + 0.5(0, Y), (0, Y)\}$, with the first element reflecting attempting not to exchange and the second reflecting

[27] The reason we could not perform this joint estimation in the labor supply experiment is that we would have had to map $101^{30}$ preference profiles (since there are 30 decisions and subjects can choose 101 options for each of them), which is not technically possible. In the exchange experiment, this approach is feasible since we only have 3 decisions and 9 possible options.

[28] It has been noted before that the CPE formulation predicts that individuals exchange in standard endowment effect designs only on the basis of consumption utility, and so fails to predict an endowment effect. The Kőszegi and Rabin (2006, 2007) EBRD model is also equipped with several alternative equilibrium concepts and refinements, Personal Equilibrium (PE) and Preferred Personal Equilibrium (PPE), the former of which can rationalize and endowment effect. Importantly, PE, PPE, and CPE all share common comparative statics for the change from Low to High conditions: loss-averse individuals should grow more willing to exchange in High relative to Low, while gain-seeking individuals should grow less willing to exchange in High relative to Low. Appendix B.1 presents all three forms of the Kőszegi and Rabin (2006, 2007) model's application to this design for completeness.

exchange, as before. They can support attempting not to exchange as a CPE if

$$u_i(0.5(X,0) + 0.5(0,Y)|0.5(X,0) + 0.5(0,Y)) \geq u_i(0,Y|0,Y),$$

which, under our functional form assumptions, reduces to the threshold

$$X_{H,i} \geq \frac{1 + 0.5(\lambda_i - 1)}{1 + 0.5(1 - \lambda_i)} Y.$$

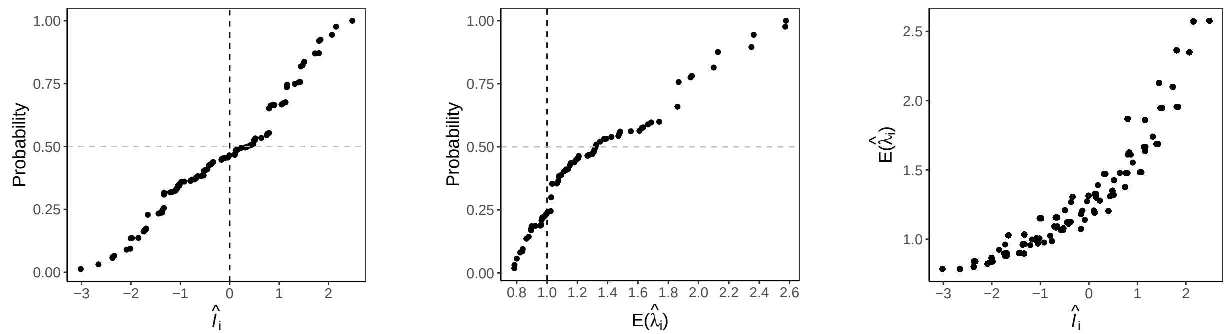
The manipulation of probabilistic forced exchange changes the CPE threshold for not exchanging from $X_{L,i} = Y$ in the Low condition to $X_{H,i} = \frac{1+0.5(\lambda_i-1)}{1+0.5(1-\lambda_i)} Y$ in the High condition.

Note that gain-loss attitudes govern the difference in response between the Low and High conditions. If $\lambda_i = 1$, then $X_{L,i} = X_{H,i}$ and individuals should exhibit identical behavior in the two conditions. If individuals are loss-averse, $\lambda_i > 1$, then $X_{L,i} < X_{H,i}$. If higher values for object $X$ are required to support not exchanging in the High condition, this implies that loss-averse individuals are more willing to exchange in High than in Low. In contrast, if individuals are gain-seeking $\lambda_i < 1$, then $X_{L,i} > X_{H,i}$, and gain-seeking individuals are less willing to exchange in High than in Low. These observations lead to the following prediction for heterogeneous treatment effects.

**Prediction 2.** Loss-averse individuals ($\lambda_i > 1$) should be more willing to exchange in the High condition relative to the Low condition. Gain-seeking individuals ($\lambda_i < 1$) should be less willing to exchange in the High condition relative to the Low condition.

### 3.3 Results From The Exchange Experiment

**Stage 1: The distribution of gain-loss attitudes in exchange.** Fifty-seven percent of subjects state that they would hypothetically choose their endowed object, 45 percent provide a higher liking rating for their endowed object compared to 33 percent for the alternative, and 45 percent provide a higher wanting rating for their endowed object compared




(a) CDF of reduced form estimates

(b) CDF of structural estimates

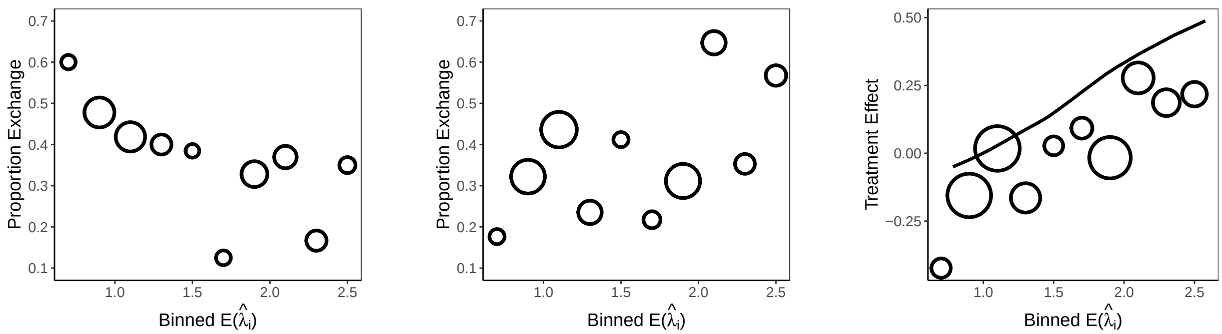
(c) Reduced form and structural estimates

Figure 3: Stage 1: Gain-loss attitudes in the exchange experiment

*Notes:* Panel (a) and (b) show CDFs of the reduced form and structural measures of gain-loss attitudes, respectively. Note that gain-seeking behavior corresponds to values of $\hat{l}_i < 0$ in the reduced form but $E[\hat{\lambda}_i] < 1$ in the structural estimates. Based on $N = 1024$ observations. Panel (c) displays the relationship between both measures (Pearson's $r = 0.95,\ p < 0.01$).

to 32 percent for the alternative. The different preference statements are remarkably correlated within individual. The pairwise Pearson correlations between hypothetical choice, relative liking, and relative wanting statements all exceed 0.7. Given random assignment of endowed objects and the counterbalanced design, the distributions of preference statements should be identical between endowed and alternative objects. Instead, all three distributions show a clear preference for the subject's endowed object relative to the alternative. For each measure we reject the null hypothesis that stated preferences are equal over the endowed and alternative objects.[29] These collected preference statements show a clear endowment effect, and so are indicative of loss aversion on average. However, we also document substantial heterogeneity. Thirty-eight percent of subjects (385 of 1024) state that they would hypothetically choose, strictly like, and strictly want their endowed object. And, twenty-six percent of subjects (262 of 1024) exhibit the opposite pattern of hypothetically choosing, strictly liking, and strictly wanting the alternative object. This heterogeneity in statements is suggestive of heterogeneous gain-loss attitudes.

[29]Two sided t-tests comparing "Endowed>Alternative" to "Alternative>Endowed" are significant for all statements (Liking: $t = 5.48$, Wanting: $t = 5.86$, Hypothetical Choice: $t = 6.06$, $p < 0.01$ for all comparisons).




(a) Exchange Condition Low (b) Exchange Condition High (c) Treatment effect

Figure 4: Stage 2: Heterogeneous treatment effects in the exchange experiment

*Notes:* Panel (a) shows the proportion of subjects deciding to exchange their endowed object across bins of $E[\hat{\lambda}_i]$ in Condition Low ($N = 417$). Panel (b) shows the proportion of subjects deciding to exchange across bins of $E[\hat{\lambda}_i]$ in Condition High ($N = 603$). Panel C provides the treatment effect, corresponding to the differences between Panels (a) and (b), as well as the KR CPE prediction for subjects with different $E[\hat{\lambda}_i]$. The bubble size indicates the number of subjects in the corresponding bin.

Figure 3 shows the distributions of our reduced form and structural measures of gain-loss attitudes along with the relationship between the two. As in the labor supply study, we document substantial variation in gain-loss attitudes, irrespective of which measure we rely on. Appendix Table A5 provides the structural estimates for the distribution of gain-loss attitudes, $N(\mu_\lambda, \sigma^2_\lambda)$, alongside the auxiliary parameters for relative utilities, $\frac{X}{Y}$, for each pair of objects; and Appendix Table A7 provides the mapping from preference statements to individual estimates of $E[\hat{\lambda}_i]$ under these estimates. Within our sample, $E[\hat{\lambda}_i]$ has mean 1.49 and median 1.32. We calculate that 76% of subjects are loss-averse, $E[\hat{\lambda}_i] > 1$, while 24% are gain-seeking $E[\hat{\lambda}_i] < 1$. This is closely in line with our labor supply findings. In addition, also as in our labor supply experiment, we observe a strong correlation between the reduced form and structural measures of gain-loss attitudes (Pearson's $r = 0.95$, $p < 0.01$).

**Stage 2: Heterogeneous treatment effects of Low vs. High.** Given our estimates of gain-loss attitudes, we analyze exchange behavior in the Stage 2 conditions, Low vs. High. Figure 4 provides a visual illustration of the connections between Stage 1 gain-loss attitudes and Stage 2 behavior. In each panel, we construct 10 equally spaced bins of $E[\hat{\lambda}_i]$ and calculate the average willingness to exchange in each bin. Panel A indicates a negative relationship between $E[\hat{\lambda}_i]$ and the probability of exchange in the Low condition: more loss-averse subjects are less likely to exchange their endowment for the alternative.[30] This relationship reverses in Panel B, which documents a positive relationship between $E[\hat{\lambda}_i]$ and willingness to exchange in the High condition. Panel C documents the relationship between $E[\hat{\lambda}_i]$ and estimated treatment effects: subjects with greater values of $E[\hat{\lambda}_i]$ exhibit larger treatment effects.

Also graphed in Figure 4, Panel C is the predicted relationship between gain-loss attitudes and CPE treatment effects. This relationship is generated by simulating choice probabilities from a logistic choice model based upon the CPE thresholds for $X_{L,i}$ and $X_{H,i}$ at the estimated values of $\frac{Y}{X}$ and $E[\hat{\lambda}_i]$. Appendix B.2 provides details of this simulation. This predicted relationship shows substantial heterogeneity in treatment effects. As the probability of forced exchange increases between the Low and the High conditions, loss-averse individuals should grow more willing to exchange and gain-seeking individuals should grow less so. The data are consistent with these predictions. Recall that alternate models of reference dependence and models that attribute Stage 1 behavior to other forces like heterogeneous valuations for the goods would not make such heterogeneous predictions.

Figure 4 is supported by regression analyses reported in Table 2. Column 1 examines the average treatment effect without accounting for heterogeneity in gain-loss attitudes. In Condition Low, 38 percent of subjects choose to exchange. Comparing this value to the neoclassical benchmark of 50 percent indicates a significant endowment effect in Condition Low, $F_{1,1022} = 25.66$, $(p < 0.01)$. The substantial endowment effect observed in Condition Low is unaffected by probabilistic forced exchange. In contrast to the prediction of EBRD models with universal loss aversion (which would predict a positive treatment effect), we

[30]Within the Kőszegi and Rabin (2006, 2007) model's CPE construct, this correlation is not predicted as exchange in the Low condition should be independent of gain-loss attitudes. However, in the alternative PE construct, this correlation is predicted by the theory. See Appendix B.1.3 for details.

Table 2: Heterogeneous treatment effects in the exchange experiment

| | *Dependent Variable: Exchange (=1)* | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Condition High | -0.004 | -0.004 | -0.340 |
| | (0.031) | (0.031) | (0.087) |
| Reduced form ($\hat{l}_i$) | | -0.050 | |
| | | (0.015) | |
| Condition High* Reduced form ($\hat{l}_i$) | | 0.077 | |
| | | (0.020) | |
| $E[\hat{\lambda}_i]$ | | | -0.136 |
| | | | (0.041) |
| Condition High * $E[\hat{\lambda}_i]$ | | | 0.225 |
| | | | (0.054) |
| Constant (Condition Low) | 0.380 | 0.380 | 0.584 |
| | (0.024) | (0.023) | (0.067) |
| R-Squared | 0.000 | 0.014 | 0.017 |
| # Observations | 1024 | 1024 | 1024 |
| $H_0$ : Zero Treatment Effect (H-L) | $F_{1,1022} = 0.01$ | $F_{1,1020} = 0.02$ | $F_{1,1020} = 15.12$ |
| | $(p = 0.91)$ | $(p = 0.90)$ | $(p < 0.01)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Exchange in L | | $F_{1,1020} = 10.69$ | $F_{1,1020} = 11.23$ |
| | | $(p < 0.01)$ | $(p < 0.01)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,1020} = 14.65$ | $F_{1,1020} = 17.25$ |
| | | $(p < 0.01)$ | $(p < 0.01)$ |

*Notes*: Ordinary least square regressions. The "Reduced form ($\hat{l}_i$)" variable captures the residuals of a regression that explains object choice with the principal component of the three preference measures. Robust standard errors in parentheses. Null hypotheses tested for: 1) zero treatment effect (Condition High coefficient= 0); 2) no relationship between gain-loss attitudes and behavior in Condition Low behavior ($E[\hat{\lambda}_i]$ or $\hat{l}_i = 0$); 3) constant treatment effect over gain-loss attitudes (Condition High * $E[\hat{\lambda}_i]$ or Condition High * $\hat{l}_i = 0$). $F$-statistics and two-sided $p$-values reported.

find that Condition High decreases the probability of exchange by -0.4 percentage points, and we fail to reject that this treatment effect is zero, $F_{1,1022} = 0.01$ $(p = 0.91)$.

The null effect in column 1 masks substantial heterogeneity in treatment effects over gain-loss attitudes. In columns 2 and 3, we interact treatment with reduced form and structural measures of gain-loss attitudes. Both measures are highly correlated with the effect of treatment and in both specifications we reject the null hypothesis of zero heterogeneity in treatment, $F_{1,1020} = 14.65$ ($p < 0.01$) and $F_{1,1020} = 17.25$ ($p < 0.01$), respectively. More loss averse subjects have greater increases in their willingness to exchange as they move from the Low to the High condition. Consistent with EBRD, individuals respond to the change in expectations across Low and High conditions and differentially so depending on their gain-loss attitudes. Alternative formulations of the reference point predict zero treatment effect and zero heterogeneity therein, and, thus, are rejected by our exchange study results.

Without accounting for heterogeneous gain loss attitudes, the average treatment effect reported in column 1 of -0.4 percentage points aggregates different-signed effects of loss-averse and gain-seeking subjects. As noted in Section 2.3, aggregation presents two potential issues related to non-linearity of treatment effects and power. In Appendix B.4, we show that in the exchange setting the relationship between $\lambda_i$ and treatment differences for exchange probability can be concave, with the negative effects for gain-seeking individuals being of greater absolute magnitude than the positive effects for loss-averse individuals. This leads to substantial aggregation issues in our setting as the average treatment effect may be substantially understated relative to the treatment effect of the average preference. This may help to explain why the average treatment effect is indeed null. Indeed, our simulations in Figure 4, Panel C indicate a predicted average treatment effect of only 0.08. Appendix B.4 shows that appropriate power for experiments exploring aggregate effects recognizing aggregation and power issues induced by heterogeneity would require around double the sample size of a homogeneous treatment effect at the average preference.

In sum, the results on the heterogeneity of gain-loss attitudes and its predictive power for the behavioral effect of a shift in the expectations-based reference point closely mirror those of the labor supply experiment. This is despite the fact that the two sets of findings rely on entirely distinct experimental paradigms and leverage different approaches for identifying gain-loss attitudes.

# 4 Conclusion

Prior work testing reference-dependent preferences assumes universal loss aversion. This paper studies the role of heterogeneity in gain-loss attitudes, and explores its implications for identifying models of the reference point. Failing to acknowledge heterogeneity in gain-loss attitudes is critical both because comparative statics used to test different formulations of the reference point can change sign depending on the level of gain-loss attitudes and because such heterogeneity is an empirical reality. In two laboratory experiments, we show that once one accounts for heterogeneity in gain-loss attitudes, experimental tests are strikingly supportive of Expectations-Based Reference Dependence (EBRD) formulations of reference points.

Our large-sample experiments show that the existing body of evidence on heterogeneity in gain-loss attitudes is not a mere artifact of measurement error or behavioral noise. Instead, by showcasing its out-of-sample predictive power, we document that gain-seeking behavior has a substantive interpretation that can be productively used in theory testing. The striking consistency of our findings across our two experimental settings attests to the robustness and importance of recognizing heterogeneity.

Conceptually, the importance of recognizing parameter heterogeneity in identifying behavioral predictions hinges on two issues: non-linearity in aggregation and statistical power. First, treatment effects need not aggregate linearly over the dimension of heterogeneity, so ignoring heterogeneity can confound inference. The severity of this concern differs by model and context, and we, ourselves, show a potentially more pronounced aggregation problem in our study of exchange behavior than in our study of labor supply. Similar concerns have been highlighted in other decision domains such as intertemporal choice (Weitzman 2001; Jackson and Yariv 2014). Second, even under linear aggregation, heterogeneity influences power considerations. An empirical study that is theoretically well-powered under the assumption of preference homogeneity may be under-powered if there is actual heterogeneity, which may lead to false conclusions from null findings. Both issues are of first-order importance for interpreting empirical tests of theories that likely feature parameters with real-world heterogeneity.

There is no universally accepted measurement of gain-loss attitudes, and each candidate has unique advantages and potential drawbacks. In the two designs presented in this manuscript, we elicit gain-loss attitudes in markedly different ways. In our labor supply study, we estimate gain-loss attitudes from a large number of incentivized decisions and treat each decision as isolated for the purposes of estimating gain-loss attitudes. Such approaches facilitate estimation, but fail to account for the possibility that the reference point (EBRD or otherwise) depends upon the entire body of choice problems. In our exchange behavior study, by contrast, we estimate gain-loss attitudes from hypothetical non-choice data, circumventing this challenge but creating the concern that the measures are not incentivized. Importantly, regardless of these differences in domain and measurement technique, we find quite similar distributions of gain-loss attitudes in our two studies. Whether measured using incentivized labor supply or hypothetical exchange choices, around three quarters of subjects are measured to be loss averse and one quarter gain seeking.

Additionally, we also elicit gain-loss attitudes for the 500 subjects in our labor supply study using a more traditional approach of lottery choices. There, as well, we identify a sizable minority of gain-seeking subjects. This suggests a potential component of portability for measures of gain-loss attitudes across domains, albeit with the recognition that this portability is not perfect: gain-loss attitudes measured in lottery choice are not significantly predictive of treatment effects in labor supply. More work evaluating the extent of heterogeneity in gain-loss attitudes across domains, linking heterogeneity across measurement techniques, and evaluating measure portability is greatly needed.

Though we provide results on the role of EBRD in the two main paradigms used to test models of reference-dependent preferences, the considerations that motivate this paper equally apply to the role of gain-loss attitudes in other classes of theories and applications. Heterogeneity matters not only for tests of non-expectations-based forms of reference dependence, such as current or backward-looking elements (e.g., Bowman et al. 1999), but also for other field settings in which loss aversion has been shown to play a role, such as job search (DellaVigna et al. 2017), insurance choice (Barseghyan et al. 2013) or tax compliance (Engström et al. 2015).

Beyond the context of gain-loss attitudes, our work contributes to a growing literature in behavioral economics that acknowledges the importance of (structurally) recognizing heterogeneity in behavioral parameters (see DellaVigna 2018 for a recent review). Our paper shows that taking the theoretical implications of heterogeneity seriously—instead of treating it as a nuisance—can deliver more comprehensive tests of behavioral theories and potentially reconcile conflicting evidence.

# References


Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman (2011). “Reference points and effort provision”. In: *The American Economic Review*, pp. 470–492.

Augenblick, Ned and Matthew Rabin (Mar. 2018). “An Experiment on Time Preference and Misprediction in Unpleasant Tasks”.

Barseghyan, Levon, Francesca Molinari, Ted O’Donoghue, and Joshua C Teitelbaum (2013). “The nature of risk preferences: Evidence from insurance choices”. In: *American economic review* 103.6, pp. 2499–2529.

Bell, David E. (1985). “Disappointment in Decision Making under Uncertainty”. In: *Operations Research* 33.1, pp. 1–27.

Bowman, David, Deborah Minehart, and Matthew Rabin (1999). “Loss aversion in a consumption–savings model”. In: *Journal of Economic Behavior & Organization* 38.2, pp. 155–178.

Brown, Alexander L, Taisuke Imai, Ferdinand Vieider, and Colin Camerer (2021). “Meta-analysis of empirical estimates of loss-aversion”. In: *Available at SSRN 3772089*.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler (1997). “Labor supply of New York City cabdrivers: One day at a time”. In: *The Quarterly Journal of Economics*, pp. 407–441.

Campos-Mercade, Pol, Lorenz Goette, Alexandre Kellogg, and Charles Sprenger (2021). “Reference-Dependent Effort Provision under Heterogeneous Loss Aversion: Pre-Analysis Plan”. In: *AEA RCT Registry. March 02.*

Cantillo, Vıctor, Johanna Amaya, and Juan de Dios Ortúzar (2010). “Thresholds and Indifference in Stated Choice Surveys”. In: *Transportation Research Part B* 44, pp. 753–763.

Cerulli-Harms, Annette, Lorenz Goette, and Charles Sprenger (Feb. 2019). "Randomizing Endowments: An Experimental Study of Rational Expectations and Reference-Dependent Preferences". In: *American Economic Journal - Microeconomics* 11.1, pp. 185–207.

Chapman, Jonathan, Mark Dean, Pietro Ortoleva, Erik Snowberg, and Colin Camerer (2017). *Willingness to Pay and Willingness to Accept are Probably Less Correlated Than You Think*. Tech. rep. National Bureau of Economic Research.

Chapman, Jonathan, Erik Snowberg, Stephanie Wang, and Colin Camerer (2018). *Loss Attitudes in the US Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE)*. Tech. rep. National Bureau of Economic Research.

Chen, Daniel L., Martin Schonger, and Chris Wickens (Mar. 2016). "oTree - An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9, pp. 88–97.

Dean, Mark and Pietro Ortoleva (2015). "Is it all connected? A testing ground for unified theories of behavioral economics phenomena". In: *A Testing Ground for Unified Theories of Behavioral Economics Phenomena (August 13, 2015)*.

DellaVigna, Stefano (2018). "Structural behavioral economics". In: *Handbook of Behavioral Economics: Applications and Foundations 1*. Vol. 1. Elsevier, pp. 613–723.

DellaVigna, Stefano, Attila Lindner, Balázs Reizer, and Johannes F Schmieder (2017). "Reference-dependent job search: evidence from Hungary". In: *The Quarterly Journal of Economics*.

Engström, Per, Katarina Nordblom, Henry Ohlsson, and Annika Persson (2015). "Tax compliance and loss aversion". In: *American Economic Journal: Economic Policy* 7.4, pp. 132–64.

Erev, Ido, Eyal Ert, and Eldad Yechiam (2008). “Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions”. In: *Journal of Behavioral Decision Making* 21.5, pp. 575–597.

Ericson, Keith M. Marzilli and Andreas Fuster (2011). “Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments”. In: *The Quarterly Journal of Economics* 126.4, pp. 1879–1907.

Fehr, Ernst and Lorenz Goette (2007). “Do workers work more if wages are high? Evidence from a randomized field experiment”. In: *American Economic Review* 97.1, pp. 298–317.

Fischbacher, Urs (2007). “z-Tree: Zurich toolbox for ready-made economic experiments”. In: *Experimental economics* 10.2, pp. 171–178.

Gächter, Simon, Eric J Johnson, and Andreas Herrmann (2022). “Individual-level loss aversion in riskless and risky choices”. In: *Theory and Decision* 92.3, pp. 599–624.

Gneezy, Uri, Lorenz Goette, Charles Sprenger, and Florian Zimmermann (2017). “The Limits of Expectations-Based Reference Dependence”. In: *Journal of the European Economic Association*.

Goette, Lorenz, Thomas Graeber, Alexandre Kellogg, and Charles Sprenger (2018). “Heterogeneity of Loss Aversion and Expectations-Based Reference Points: Replication”. In: *AEA RCT Registry. July 05.*

Harinck, Fieke, Eric Van Dijk, Ilja Van Beest, and Paul Mersmann (2007). “When gains loom larger than losses reversed loss aversion for small amounts of money”. In: *Psychological Science* 18.12, pp. 1099–1105.

Heffetz, Ori and John A. List (2014). “Is the Endowment Effect an Expectations Effect?” In: *Journal of the European Economic Association* 12.5, pp. 1396–1422.

Jackson, Matthew O and Leeat Yariv (2014). “Present bias and collective dynamic choice in the lab”. In: *American Economic Review* 104.12, pp. 4184–4204.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler (1990). “Experimental Tests of the Endowment Effect and the Coase Theorem”. In: *Journal of Political Economy* 98.6, pp. 1325–1348.

Knetsch, Jack and Wei-Kang Wong (2009). “The endowment effect and the reference state: Evidence and manipulations”. In: *Journal of Economic Behavior & Organization* 71.2, pp. 407–413.

Kőszegi, Botond and Matthew Rabin (2006). “A Model of Reference-Dependent Preferences”. In: *The Quarterly Journal of Economics* 121.4, pp. 1133–1165.

— (2007). “Reference-Dependent Risk Attitudes”. In: *The American Economic Review* 97.4, pp. 1047–1073.

— (2009). “Reference-Dependent Consumption Plans”. In: *The American Economic Review* 99.3, pp. 909–936.

Loomes, Graham and Robert Sugden (1986). “Disappointment and Dynamic Consistency in Choice under Uncertainty”. In: *Review of Economic Studies* 53.2, pp. 271–82.

Masatlioglu, Yusufcan and Collin Raymond (2016). “A Behavioral Analysis of Stochastic Reference Dependence”. In: *American Economic Review* 106.9, pp. 2760–82.

McFadden, Daniel (1974). “Conditional Logit Analysis of Qualitative Choice Behavior”. In: *Frontiers in Econometrics*. Ed. by Paul Zarembka. New York: Academic Press. Chap. 4.

Nicolau, Juan L (2012). “Asymmetric tourist response to price: loss aversion segmentation”. In: *Journal of Travel Research* 51.5, pp. 568–676.

Odean, Terrance (1998). “Are Investors Reluctant to Realize Their Losses?” In: *The Journal of Finance* 53.5, pp. 177–1798.

Rabin, Matthew (2000). “Risk Aversion and Expected Utility Theory: A Calibration Theorem”. In: *Econometrica* 68.5, pp. 1281–1292.

Raven, John and Jean Raven (2003). “Raven Progressive Matrices”. In: *Handbook of Nonverbal Assessment*. Ed. by R. Steve McCallum. Boston, MA: Springer US, pp. 223–237.

Slovic, Paul, Baruch Fischhoff, Sarah Lichtenstein, Bernard Corrigan, and Barbara Combs (1977). “Preference for insuring against probable small losses: Insurance implications”. In: *Journal of Risk and insurance*, pp. 237–258.

Smith, Alex (2019). “Lagged Beliefs and Reference-Dependent Utility”. In: *Journal of Economic Behavior & Organization* 167, pp. 331–340.

Sokol-Hessner, Peter, Ming Hsu, Nina G Curley, Mauricio R Delgado, Colin F Camerer, and Elizabeth A Phelps (2009). “Thinking like a trader selectively reduces individuals’ loss aversion”. In: *Proceedings of the National Academy of Sciences* 106.13, pp. 5035–5040.

Sprenger, Charles (2015). “An endowment effect for risk: Experimental tests of stochastic reference points”. In: *Journal of Political Economy* 123.6, pp. 1456–1499.

Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Weitzman, Martin L (2001). “Gamma discounting”. In: *American Economic Review* 91.1, pp. 260–271.

# Appendix For Online Publication

## Appendix Table of Contents

# Appendix A Additional Development, Analysis, and Results for Labor Supply Study

## A.1 Theoretical Considerations for Labor Supply

We present the theoretical framework of the Kőszegi and Rabin (2006, 2007) EBRD formulation as applied to an individual's labor supply decision. An agent's utility consists of two components—consumption utility derived from earned wages and the (negative) cost of exerting effort, and psychological utility derived from comparing the realized wage and effort level to the agent's expectations. Formally, this is represented by

$$u_i(w, e|r_w, r_e) = m(we) - c_i(e) + \mu_i(m(we) - m(r_w)) + \mu_i(c_i(e) - c_i(r_e)).$$

The first component of utility, $m(we) - c_i(e)$, is standard consumption utility obtained from working $e$ tasks and earning $we$. Consumption utility is complemented with a reference-dependent, psychological component of utility, for which the utility from realized earnings $m(we)$ is compared to the utility of reference-point earnings $m(r_w)$ under a piece-wise linear gain-loss function $\mu_i$, where

$$\mu_i(z) = \begin{cases} \eta z & z \geq 0 \\ \eta\lambda_i z & z < 0. \end{cases}$$

Intuitively, if an outcome falls short of the referent by a difference of $z$, this leads to a reduction of utility by $\eta\lambda_i$ times this difference. An outcome that exceeds the reference point increases utility by $\eta$ times the difference, where $\eta > 0$. Thus, $\lambda_i$ represents individual gain-loss attitude and can either exhibit loss-aversion where losses are felt more severely than commensurate gains, $\lambda_i > 1$, or gain-seeking where gains are felt more severely than commensurate losses, $\lambda_i < 1$. If $\lambda_i = 1$, there is the individual is considered 'loss-neutral'. Throughout the analysis, we assume that $m(we) = we$ and constant for all individuals, that $c_i(e)$ is an increasing at least twice-differentiable convex function, and normalize $\eta = 1$ for all individuals.

Kőszegi and Rabin (2006, 2007) propose that agents hold the entire distribution of the outcome space as their expectation. Each potential realization is compared to every other potential realization and weighted by the relevant densities. In the labor supply context, decision-makers face a potentially stochastic schedule of wages and must commit to an effort level prior to the realization of wages. Thus, when considering the utility of an effort level $e'$, the agent computes the expected consumption utility given the known wage distribution as well as the expected gain-loss utility. Mathematically, this is represented as a double integral over the stochastic reference points ($\mathbf{r} = (r_w, e')$) and the stochastic consumption realizations ($\mathbf{c} = (w, e')$):

$$U_i(F|G) = \int \int u_i(\mathbf{c}|\mathbf{r}) dG(\mathbf{r}) dF(\mathbf{c}),$$

where $F, G$ represent the lotteries over the wage-outcome space at a fixed level of effort.

### A.1.1 Choice Acclimating Personal Equilibrium (CPE)

In order to close the model, Kőszegi and Rabin (2006, 2007) equip it with the rational equilibrium concept known as CPE:

*Choice Acclimating Personal Equilibrium (CPE):* A choice $F \in \mathcal{D}$, where $\mathcal{D}$ is the possible outcome space, is a choice-acclimating personal equilibrium if

$$U_i(F|F) \geq U_i(F'|F') \ \forall \ F' \in \mathcal{D}.$$

In our context, the effort level $e_i^*$ is a CPE if its associated ex-ante utility—given the distribution of wages it induces—is the largest of all the possible effort choices given the ex-ante distributions they respectively induce. In deriving comparative static predictions throughout the following sections, we will assume that agents seek to maximize their CPE utility.

### A.1.2 CPE Comparative Statics

We consider how CPE individuals behave when offered a wage $(p, H; q, L; 0.5, w)$ where $L < H$; that is, individuals have a 50% chance of earning a piece-rate, $w$ per unit of effort, a $p$% chance of earning \$$H$, and a $q = (0.5 - p)$% chance of earning \$$L$ regardless of effort. The CPE utility induced by a prospective effort level, $e_i$ is given by

$$U((p, H; q, L; 0.5, we_i)|(p, H; q, L; 0.5, we_i)) =$$

$$\begin{cases} pH + qL + 0.5we_i + (1 - \lambda_i)\left[pq(H - L) + 0.5p(H - we_i) + 0.5q(L - we_i)\right] - c_i(e_i) & we_i < L < H \\ pH + qL + 0.5we_i + (1 - \lambda_i)\left[pq(H - L) + 0.5p(H - we_i) + 0.5q(we_i - L)\right] - c_i(e_i) & L < we_i < H \\ pH + qL + 0.5we_i + (1 - \lambda_i)\left[pq(H - L) + 0.5p(we_i - H) + 0.5q(we_i - L)\right] - c_i(e_i) & L < H < we_i. \end{cases}$$

Following the appendix of Gneezy et al. (2017), we study the effects of an increase in $p$ by signing the derivative $\frac{\partial e_i^*}{\partial p}|_{p+q=0.5}$ when $L \leq we_i \leq H$. When the considered level of effort yields earnings between $L$ and $H$, the optimal level of effort can be found by studying the first order condition of

$$0.5w\left[1 + (p - q)(\lambda_i - 1)\right] = c_i'(e_i^*).$$

Defining $\bar{P} = p + q = 0.5$ and $p - q = 2p - \bar{P} = 2p - 0.5$, we can sign the partial derivative as

$$\frac{\partial e_i^*}{\partial p}|_{p+q=0.5} = (c_i'^{-1})'(0.5w[1 + (2p - 0.5)(\lambda - 1)]) * (\lambda - 1)w.$$

By the inverse function theorem, $(c'^{-1})'(0.5w[1+(2p-0.5)(\lambda_i-1)]) * (\lambda_i-1)w = \frac{1}{c_i''(e_i^*)}$ where $0.5w[1+(2p-0.5)(\lambda_i-1)] = c'(e_i^*)$. Thus,

$$\frac{\partial e_i^*}{\partial p}|_{p+q=0.5} = \frac{(\lambda_i-1)w}{c''(e_i^*)}$$

and by the assumed convexity of $c_i(\cdot)$, we know $c_i''(e^*) > 0$ so that

$$\lambda_i > 1 \implies \frac{\partial e_i^*}{\partial p}|_{p+q=0.5} > 0$$
$$\lambda_i < 1 \implies \frac{\partial e_i^*}{\partial p}|_{p+q=0.5} < 0.$$

Thus, under CPE, loss-averse individuals are predicted to increase their effort whereas gain-seeking individuals are predicted to decrease their effort in response to an increasing in $p$, holding fixed $p+q=0.5$.

For completeness, we also discuss the other two cases: $we_i < L < H$ and $L < H < we_i$. First, consider $we_i < L < H$. The first order condition yielding optimal effort is

$$0.5w\left[1+(p+q)\eta(\lambda-1)\right] = c'(e),$$

and because $c_i'(e_i)$ is continuous and differentiable, $c_i'^{-1}(e_i)$ exists and the optimal $e_i^*$ is

$$e_i^* = c_i'^{-1}\left(0.5w\left[1+(p+q)(\lambda_i-1)\right]\right).$$

Turning back to $\frac{\partial e_i^*}{\partial p}|_{1-p-q=0.5}$, let $p+q=\bar{P}=0.5$—since changes in $p$ must leave $p+q$ constant, we have that $\frac{\partial e_i^*}{\partial p}|_{1-p-q=0.5} = 0$ in this case. Next, consider $L < H < we_i$. Again, we examine the first order condition given by

$$0.5w\left[1-(p+q)(\lambda_i-1)\right] = c'(e_i),$$

and

$$e_i^* = c_i'^{-1}(0.5w[1-(p+q)(\lambda_i-1)]),$$

again yielding $\frac{\partial e_i^*}{\partial p}|_{1-p-q=0.5} = 0$.

## A.2 Estimation and Calculation of Gain Loss Attitudes in Labor Supply

We now discuss how to identify gain-loss preferences in the context of real effort. Consider how the introduction of mean-preserving spreads over wages affect individuals across the gain-loss types: loss-averse individuals would suddenly be exposed to gains and losses at each potential effort level, and because losses loom larger than gains for these types, they would prefer to work fewer tasks under this wage structure. Gain-lovers instead weight the losses relatively less than the gains, so that additional effort can generate even more positive surprises, leading to increases in effort provision.

Theoretically, a CPE agent facing a deterministic wage maximizes the following utility function:

$$u(we_i|we_i) = we_i - c_i(e_i),$$

so that the optimal effort choice, $e_i^*(w)$, satisfies the first order condition $w = c_i'(e_i)$. Variation in $w$ traces out a cost of effort curve, which as pre-registered we assume takes the functional form $c_i(e_i) = \frac{1}{\alpha\gamma_i}(e_i + 10)^{\gamma_i}$ as in Augenblick and Rabin (2018), where 10 represents the required minimum number of tasks that all subjects must complete. The marginal consideration is thus

$$\frac{1}{\alpha}(e_i + 10)^{(\gamma_i - 1)} = w.$$

By introducing a mean-preserving spread of these wages, we are able to identify the gain-loss parameter $\lambda_i$. To see this, consider the piece-rate $(0.5, w_l; w_h)$, $(w_h > w_l)$, which represents a contract under which the agent exerts effort $e_i$ knowing that with 50% chance, they will earn either $e_i \times w_l$ or $e_i \times w_h$. The associated CPE utils for such an effort choice, $e_i$, is then

$$0.5w_le_i + 0.5w_he_i - 0.25(\lambda_i - 1)(w_he_i - w_le_i) - c_i(e_i),$$

where $c_i(e_i)$ is as described above. The optimal effort choice under this wage structure, $e_i^*$, must then satisfy the first order condition

$$0.5w_l + 0.5w_h - 0.25(\lambda_i - 1)(w_h - w_l) = \frac{1}{\alpha}(e_i + 10)^{\gamma_i - 1}.$$

In order to provide a structural estimate of the parameter, $\lambda_i$, we conduct the non-linear regression corresponding to the log-transformed marginal conditions above with additive shocks

$$log(e_i + 10) = \frac{1}{\gamma_i - 1} log(\alpha_i \left[\overline{w} - 0.25(\lambda_i - 1)\Delta w\right]) + \epsilon_i, \tag{6}$$

where we define the variable $\overline{w} = 0.5w_l + 0.5w_h$ for uncertain piece rates, and $\overline{w} = w$ for fixed piece rates. We also define the variable $\Delta w = w_h - w_l$ for uncertain piece rates, and $\Delta w = 0$ for fixed piece rates. Because the data are potentially censored at $e_i = 0$ and $e_i = 100$, we use a maximum likelihood tobit method.

To set meaningful starting values, we first estimate equation (6) on only the fixed piece rate data, which eliminates $\lambda_i$ from estimation, with starting values $\alpha_i = 594$ and $\gamma_i = 3.64$. We then use resulting estimates for $\alpha_i$ and $\gamma_i$ as starting values for the combined fixed piece rate and uncertain piece rate data set along with a starting value $\lambda_i = 0.81$. Anyone for whom the first estimation step fails to converge, we retain the original starting values.[31]

Following the theoretical insights of our model, our implementation places box constraints on the parameters $\gamma_i \in (1, 4)$ and $\lambda_i \in (0, 3)$ by estimating the parameters $g_i$ and $l_i$ such that $\gamma_i = 1 + 3(1/(1 + exp(g_i)))$ and $\lambda_i = 3(1/(1 + exp(l_i)))$. The restriction on $\gamma$ requires costs to be convex, but not overly so. The restriction on $\lambda_i$ within reasonable bounds; $\lambda_i > 3$ is ruled out under CPE since it has unrealistic implications—including violations of First Order Stochastic Dominance (see Masatlioglu and Raymond 2016 for

[31]The starting values were picked such that the estimates converged for as many subjects as possible. The estimates remain fairly consistent when choosing alternative starting values. For example, choosing $\lambda_i = 1$ yields convergence for one fewer subject, but we find similarly that 31% of the subjects are gain-seeking and equivalent results for the heterogeneity analysis. We also find qualitatively equivalent results when choosing $\lambda_i = 2$.

more details). Operationally, allowing $\lambda_i \geq 3$ generates additional problems as it leads to undefined allocation values (a root of a negative number) unless $\gamma_i = 2$. Parameters $\gamma_i$ and $\lambda_i$ along with their standard errors are recovered via the delta method.

The product of such an exercise is an estimated triple, $(\hat{\alpha}_i, \hat{\gamma}_i, \hat{\lambda}_i)$, capturing gain-loss attitudes alongside auxiliary parameters of the individual's cost function. Our estimation method converges and delivers an estimate of $\hat{\lambda}_i$ and $\hat{\sigma}_{\lambda_i}$ for a subset 451 of 500 subjects (the average value of $\hat{\lambda}_i$ is 1.31 with a standard deviation of 0.88). We further find convergence of $\hat{\gamma}_i$ for 359 subjects (average: 2.29; standard deviation: 0.53) and of $\hat{\alpha}_i$ for 346 subjects (average: 686; standard deviation: 151).

### A.2.1 Classifying Individual Gain-Loss Attitudes Accounting for Errors

Note that individual estimates of the parameters are likely to be estimated with error. This appendix section develops a standard Bayesian shrinkage exercise leveraging distributional information on all $\hat{\lambda}_i$ and the estimated errors, $\hat{\sigma}_{\lambda_i}$. This exercise, effectively a random-effects meta-analysis on our data, maps from an individual's value of $\hat{\lambda}_i$ and $\hat{\sigma}_{\lambda_i}$ to an expected value $E[\hat{\lambda}_i]$. The outcome of this exercise is that imprecise estimates of $\hat{\lambda}_i$ are shrunk to the sample average in proportion to their imprecision. While we focus on the $E[\hat{\lambda}_i]$, the values of which are used in the main body of the paper, we additionally perform the same analysis for $E[\hat{\gamma}_i]$ for robustness checks in appendix table A2.

For each of the 451 subjects for whom we find convergence, we have values for $\hat{\lambda}_i$ and $\hat{\sigma}_{\lambda_i}$. In our estimation process, we assume that the true value of $\lambda_i$ follows a normal distribution $N(\hat{\lambda}_i, \hat{\sigma}^2_{\lambda_i})$. We further assume that the population distribution of $\lambda_i$ follows a $N(\mu, \sigma^2)$, which we can estimate. Our goal is to estimate $E[\hat{\lambda}_i]$, which we define as the expected $\lambda_i$ given that $\lambda_i$ is drawn from the population distribution.

Note that if $\lambda_i$ follows $N(\mu, \sigma)$ and our observations are $N(\hat{\lambda}_i, \hat{\sigma}^2_{\lambda_i})$ then they are joint normal with mean vector $(\mu, \mu)$ and covariance matrix

$$\begin{pmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \hat{\sigma}^2_{\lambda_i} + \sigma^2 \end{pmatrix}.$$

Then, the conditional distribution of the true $\lambda_i$ given an observation $\hat{\lambda}_i$ is normal with expectation

$$E[\hat{\lambda}_i] = \mu + \frac{\sigma^2}{\hat{\sigma}^2_{\lambda_i} + \sigma^2}(\hat{\lambda}_i - \mu),$$

or, alternatively,

$$E[\hat{\lambda}_i] = \frac{\mu \hat{\sigma}^2_{\lambda_i} + \hat{\lambda}_i \sigma^2}{\hat{\sigma}^2_{\lambda_i} + \sigma^2}.$$

We use this result to assign a value $E[\hat{\lambda}_i]$ to each of the 451 for whom we find convergence of both $\hat{\lambda}_i$ and standard error $\hat{\sigma}^2_{\lambda_i}$. This exercise yields an average value of $E[\hat{\lambda}_i] = 1.37$, with a standard deviation of 0.77, and a proportion 69.2% of the subjects having $E[\hat{\lambda}_i] > 1$. Without adjusting for measurement error, the average value of $\hat{\lambda}_i$ is 1.31, the standard deviation is 0.88, and 58.76% have $\hat{\lambda}_i > 1$. Figure A1 plots the relationship between both estimates, showing that both estimates are very highly correlated.

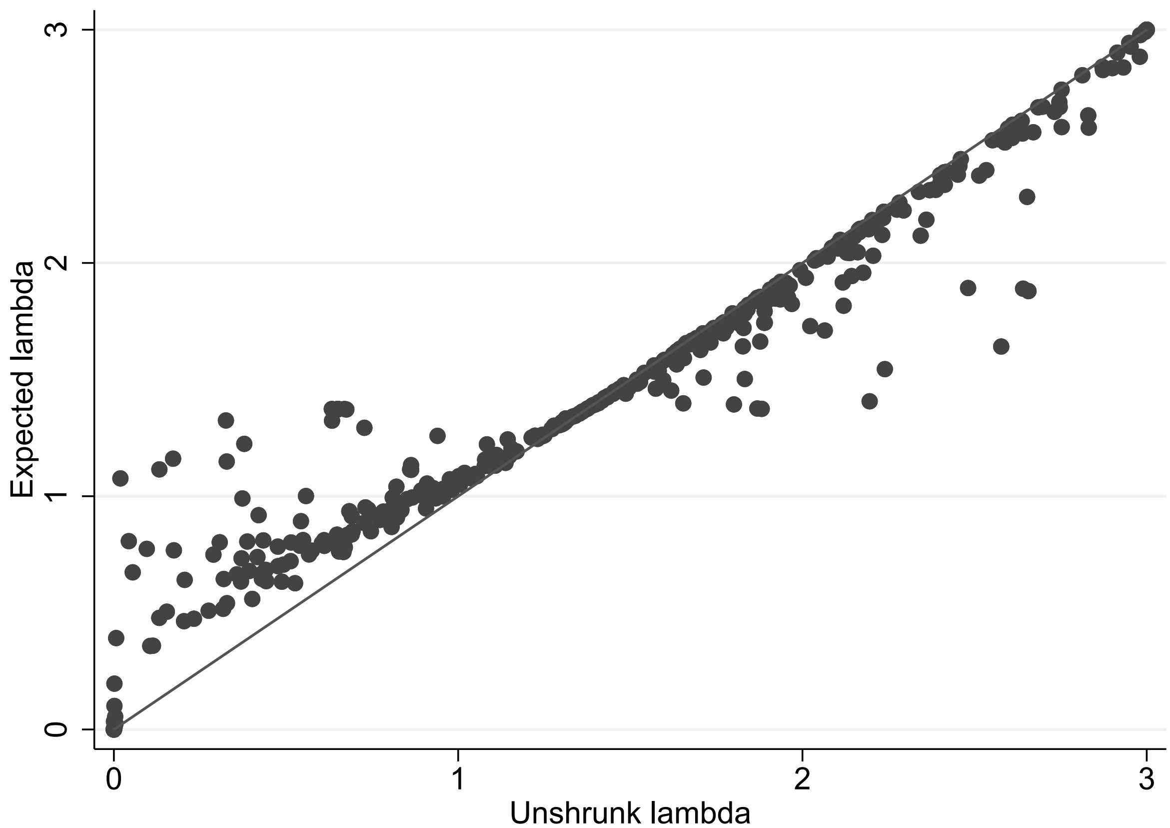



Figure A1: Relationship between $\hat{\lambda}_i$ and $E[\hat{\lambda}_i]$

## A.3 Predicting Heterogeneous Treatment Effects in Labor Supply

Appendix Section A.1 established the first order condition for the critical case of $L \leq we_i \leq H$ as

$$0.5w\left[1+(p-q)(\lambda_i-1)\right] = c_i'(e_i^*).$$

In our implementation we set $H = \$20$, $L = \$0$, $w = 0.20$, hold $p + q = 0.5$ fixed, and vary $p$ from 0.05 in the Low condition to 0.45 in the High condition. Under the assumed functional form $c_i(e_i) = \frac{1}{\alpha_i \gamma_i}(e_i + 10)^{\gamma_i}$, where 10 represents the required tasks, this yields solutions

$$\begin{aligned} e^*_{i,L} &= \left(\alpha_i 0.10\left[1-0.4(\lambda_i-1)\right]\right)^{\frac{1}{\gamma_i-1}} - 10, \\ e^*_{i,H} &= \left(\alpha_i 0.10\left[1+0.4(\lambda_i-1)\right]\right)^{\frac{1}{\gamma_i-1}} - 10, \end{aligned}$$

and a treatment effect as a function of the key parameters of interest,

$$\begin{aligned} TE(\lambda_i, \gamma_i, \alpha_i) &= e^*_{i,H} - e^*_{i,L} \\ &= \left(\alpha_i 0.10\left[1+0.4(\lambda_i-1)\right]\right)^{\frac{1}{\gamma_i-1}} - \left(\alpha_i 0.10\left[1-0.4(\lambda_i-1)\right]\right)^{\frac{1}{\gamma_i-1}}. \end{aligned}$$

This treatment effect will depend on the structure of the cost function (i.e., $\alpha_i$ and $\gamma_i$) along with gain-loss attitudes. In order to focus on the heterogeneity of treatment effects associated with gain-loss attitudes, the analysis of Figure 2 in the main text presents predictions for $TE(\lambda_i, Median(\gamma_i), Median(\alpha_i))$, at the median estimated values of $\gamma_i = 2.26$ and $\alpha_i = 609$ and each individual's specific value of $E[\hat{\lambda}_i]$. Reconducting this exercise at each individual's estimated value of $\gamma_i$ and $\alpha_i$ (when all estimates are available) does not meaningfully change the analysis.

## A.4 Non-Linear Aggregation of Labor Supply Treatment Effects and Statistical Power

Having established an individual's theoretical treatment effect,

$$\begin{aligned} TE(\lambda_i, \gamma_i, \alpha_i) &= e^*_{i,H} - e^*_{i,L} \\ &= \left(\alpha_i 0.10\left[1 + 0.4(\lambda_i - 1)\right]\right)^{\frac{1}{\gamma_i - 1}} - \left(\alpha_i 0.10\left[1 - 0.4(\lambda_i - 1)\right]\right)^{\frac{1}{\gamma_i - 1}}, \end{aligned}$$

we can consider aggregation of treatment effects into an average treatment effect,

$$\overline{TE(\lambda_i, \gamma_i, \alpha_i)} = \frac{1}{N} \sum TE(\lambda_i, \gamma_i, \alpha_i).$$

When will the average treatment effect deviate from the treatment effect of the average gain-loss attitude, $\overline{\lambda}_i$? Note that for quadratic costs, $\gamma_i = 2$, the marginal cost function is linear, and so treatment effects are a linear function of $\lambda_i$ and $\alpha_i$.

$$TE(\lambda_i, 2, \alpha_i) = \left(\alpha_i 0.10\left[0.8(\lambda_i - 1)\right]\right).$$

If $\lambda_i$ and $\alpha_i$ are independent, then averaging over these two linear dimensions of heterogeneity will not lead to deviations between the average treatment effect and the treatment effect of the average gain-loss attitude.

Outside of the linear marginal cost case of $\gamma = 2$, aggregation will not necessarily be linear. Holding $\alpha$ and $\gamma$ fixed across individuals, Figure A2 plots $TE(\lambda_i, \gamma, \alpha)$ for various values of $\gamma$. For $\gamma = 1.25$ and $\gamma = 2.75$, marginal costs are non-linear and $TE(\lambda_i, \gamma, \alpha)$ is similarly a non-linear function of $\lambda_i$. In such cases, the average treatment effect does not necessarily correspond to the treatment effect of the average preference. Moreover, the relationship between non-linearity in marginal costs and treatment effects illustrated in Figure A2 may lead average treatment effects to overstate the case for loss aversion. Of course, much depends on the distribution of gain loss attitudes and the shape of costs, but

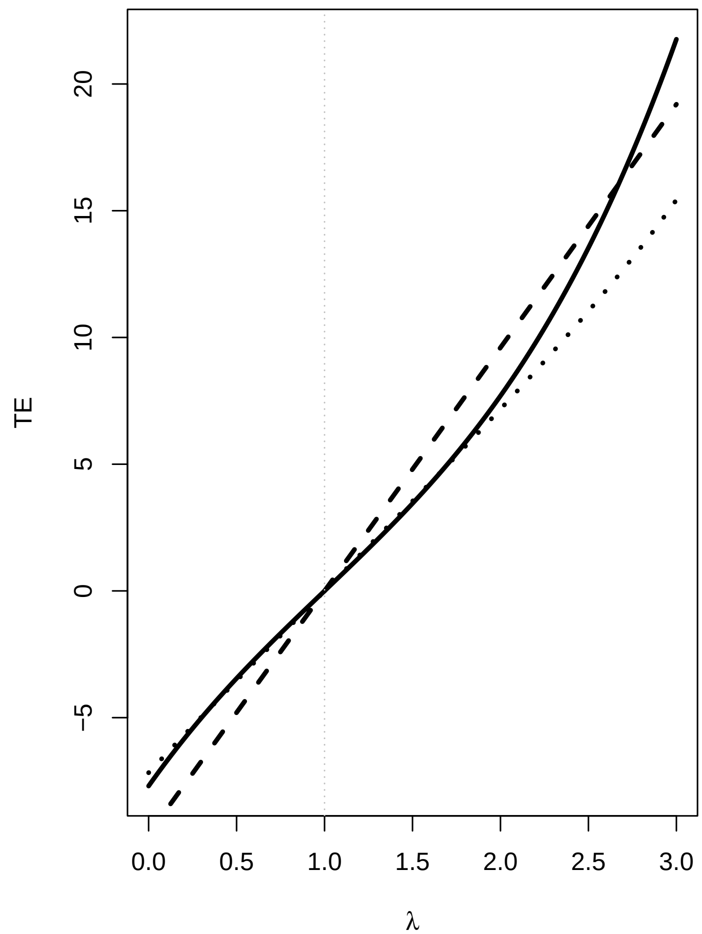



Figure A2: Predicted treatment effects by gain-loss attitudes

*Notes:* This figure represents predicted treatment effects for different values of $\lambda$ (x-axis) and $\gamma$. The full, dashed, and dotted lines represent $\gamma = 1.25$, $\gamma = 2$, and $\gamma = 2.75$, respectively.

if the average preference is loss averse, then due to the convexity apparent in Figure A2 for some values of $\gamma$ one could obtain substantially upwards-biased average treatment effects.

In our estimation exercise, we obtain a median value of $\gamma_i = 2.26$. Given the proximity of this central tendency to $\gamma = 2$, our estimates do not suggest much non-linearity in treatment effects over $\lambda_i$. Nonetheless, even if the average treatment effect is the treatment effect of the average preference, aggregating over different gain-loss types can affect the power of any conducted experimental test. Ignoring heterogeneity in $\alpha$, and assuming linear aggregation, $\gamma = 2$, then

$$\begin{aligned}\overline{TE(\lambda_i, 2, \alpha)} &= TE(\overline{\lambda_i}, 2, \alpha) \\ &= \kappa\overline{\lambda_i} - \kappa,\end{aligned}$$

where $\kappa = \alpha \cdot 0.10 \cdot 0.8$. The theoretical standard deviation of the treatment effect is thus

$$sd(TE(\lambda_i, 2, \alpha)) = \kappa \cdot sd(\lambda_i).$$

Under our median estimate of $\alpha_i = 609$, $\kappa \approx 48$. We estimate an average (measurement error adjusted) $\lambda_i = 1.37$ with a standard deviation of 0.77. Absent any other source of variation, we would thus expect an average treatment effect of 17.76 with a standard deviation of 36.96 under our estimated distribution of gain-loss attitudes.

A study that is theoretically powered assuming homogeneous gain-loss attitudes and straightforward sampling variation will have different power considerations when accounting for this additional source of variation introduced by heterogeneity. Consider an EBRD labor supply experiment conducted with approximately 100 subjects. Absent heterogeneity in gain-loss attitudes, a treatment effect of 17.8 would be powered at 80% with 100 subjects if the standard deviation of treatment effects due to sampling variation alone were approximately 63. If heterogeneity and sampling variation were independent (and thus additive in variance to yield standard deviation $\sqrt{63^2 + 37^2)} \approx 73$), this same treatment effect would require approximately 135 observations to appropriately power accounting for the above heterogeneity. Hence, accounting for heterogeneity in gain-loss attitudes can substantially alter the power considerations associated with testing average treatment effects in labor supply designs, even under linear aggregation.

### A.5 Additional Tables and Figures for Labor Supply Study

Table A1: Descriptive statistics in the labor supply study

| Decision | Fixed w. | Low w. | High w. | Mean | SD | Q25 | Q50 | Q75 | Fr. 100 | Fr. 0 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.05 | | 0 | 19.43 | 27.49 | 0 | 8.5 | 26.25 | 0.06 | 0.37 |
| 2 | 0.1 | | 0 | 25.1 | 29.47 | 0 | 14 | 40 | 0.07 | 0.28 |
| 3 | 0.125 | | 0 | 31.55 | 30.94 | 5 | 21 | 50 | 0.08 | 0.2 |
| 4 | 0.15 | | 0 | 38.68 | 33.45 | 10 | 30 | 61 | 0.12 | 0.15 |
| 5 | 0.175 | | 0 | 48.83 | 34.22 | 20 | 45.5 | 80 | 0.2 | 0.06 |
| 6 | 0.2 | | 0 | 39.05 | 35.27 | 9 | 28 | 65 | 0.16 | 0.15 |
| 7 | 0.225 | | 0 | 42.85 | 35.15 | 11 | 36.5 | 71 | 0.16 | 0.14 |
| 8 | 0.25 | | 0 | 46.6 | 35.37 | 14 | 42 | 80 | 0.19 | 0.12 |
| 9 | 0.275 | | 0 | 51.59 | 34.81 | 20 | 50 | 85 | 0.21 | 0.09 |
| 10 | 0.3 | | 0 | 61.7 | 32.99 | 34 | 60 | 100 | 0.31 | 0.01 |
| 11 | | 0 | 0.1 | 18.59 | 27.45 | 0 | 6 | 25.25 | 0.06 | 0.38 |
| 12 | | 0 | 0.2 | 23.89 | 29.44 | 0 | 11.5 | 36 | 0.07 | 0.3 |
| 13 | | 0.025 | 0.225 | 31.74 | 30.4 | 8 | 21.5 | 50 | 0.09 | 0.17 |
| 14 | | 0.05 | 0.25 | 38.52 | 31.72 | 12 | 31 | 56 | 0.12 | 0.12 |
| 15 | | 0.075 | 0.275 | 48.82 | 32.82 | 20 | 43 | 73.5 | 0.19 | 0.04 |
| 16 | | 0.1 | 0.3 | 35.71 | 32.12 | 10 | 28 | 50 | 0.13 | 0.14 |
| 17 | | 0.125 | 0.325 | 41.01 | 32.52 | 12.75 | 35 | 60 | 0.14 | 0.11 |
| 18 | | 0.15 | 0.35 | 45.72 | 33.23 | 17 | 42 | 71 | 0.16 | 0.1 |
| 19 | | 0.175 | 0.375 | 51.73 | 33.94 | 21 | 50 | 83.25 | 0.2 | 0.08 |
| 20 | | 0.2 | 0.4 | 60.79 | 32.78 | 30 | 60 | 100 | 0.3 | 0.01 |
| 21 | | 0.025 | 0.275 | 31.27 | 30.51 | 7 | 20 | 50 | 0.09 | 0.16 |
| 22 | | 0 | 0.3 | 27.39 | 30.12 | 2 | 16.5 | 45 | 0.08 | 0.22 |
| 23 | | 0.025 | 0.325 | 36.88 | 31.15 | 10 | 30 | 55.25 | 0.1 | 0.11 |
| 24 | | 0.05 | 0.35 | 43.95 | 32.71 | 16 | 40 | 64 | 0.16 | 0.08 |
| 25 | | 0 | 0.4 | 36.84 | 32.87 | 10 | 25 | 56 | 0.13 | 0.12 |
| 26 | | 0.075 | 0.375 | 38.96 | 32.87 | 10 | 30 | 60 | 0.15 | 0.1 |
| 27 | | 0.05 | 0.45 | 40.51 | 32.24 | 13 | 32.5 | 60 | 0.13 | 0.09 |
| 28 | | 0 | 0.5 | 32.29 | 31.71 | 6 | 20 | 50 | 0.1 | 0.17 |
| 29 | | 0.1 | 0.5 | 49.54 | 33.33 | 20 | 44 | 79.25 | 0.2 | 0.04 |
| 30 | | 0 | 0.6 | 39.74 | 33.47 | 11 | 30 | 61.25 | 0.15 | 0.1 |

*Notes*: This table shows the decisions that subjects faced in the labor supply study, with their respective descriptive statistics. The first ten decisions consisted of a fixed wage (Fixed w.), and decisions 11 to 30 consisted of a stochastic wage (with a low wage, Low w., and a high wage, High w.). For each decision, this table shows the average number of tasks that subjects decide to solve (Mean), the standard deviation (SD), the 25-percentile effort choice (Q25), the median effort choice (Q50), the 75-percentile effort choice (Q75), the fraction of decisions to solve 100 tasks (Fr. 100), and the fraction of decisions to solve 0 tasks (Fr. 0).

Table A2: Heterogeneous treatment effects: Robustness checks

| | Full sample | | | Full var-cov matrix | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Condition High | 7.903 | 14.893 | 18.022 | 5.528 | 15.551 | 59.417 |
| | (1.095) | (5.602) | (10.363) | (1.213) | (7.631) | (18.424) |
| Reduced form ($\hat{l}_i$) | -3.269 | | | -5.603 | | |
| | (1.053) | | | (1.324) | | |
| Condition High * Reduced form ($\hat{l}_i$) | 4.449 | | | 6.713 | | |
| | (1.119) | | | (1.437) | | |
| $\hat{\lambda}_i$ | | -4.625 | | | -7.896 | |
| | | (1.618) | | | (1.898) | |
| Condition High * $\hat{\lambda}_i$ | | 6.000 | | | 8.715 | |
| | | (1.379) | | | (1.901) | |
| $E[\hat{\lambda}_i]$ | | | -4.373 | | | -9.229 |
| | | | (2.355) | | | (3.114) |
| Condition High * $E[\hat{\lambda}_i]$ | | | 9.905 | | | 8.856 |
| | | | (2.135) | | | (2.215) |
| $\hat{\gamma}_i$ | | -25.102 | | | -47.104 | |
| | | (3.322) | | | (5.534) | |
| Condition High * $\hat{\gamma}_i$ | | -5.298 | | | -7.464 | |
| | | (2.241) | | | (3.185) | |
| $E[\hat{\gamma}_i]$ | | | -16.366 | | | -68.224 |
| | | | (6.488) | | | (30.233) |
| Condition High * $E[\hat{\gamma}_i]$ | | | -9.304 | | | -26.979 |
| | | | (4.216) | | | (7.681) |
| Constant (Condition Low) | 37.585 | 100.732 | 84.398 | 43.241 | 154.775 | 208.053 |
| | (1.616) | (8.520) | (15.943) | (2.031) | (13.303) | (71.838) |
| $N$ | 1000 | 964 | 698 | 668 | 668 | 668 |
| $R^2$ | 0.033 | 0.248 | 0.099 | 0.045 | 0.454 | 0.232 |

*Notes*: Ordinary least squares regression explaining each subject's treatment effect (dependent variable) with their Stage 1 decisions. Models (1) and (2) use the raw $\hat{\lambda}_i$ and $\gamma$ without shrinkage. Models (3) and (4) restrict the sample to subjects for whom we can estimate a full-rank variance-covariance matrix. Robust standard errors in parentheses.

Table A3: Heterogeneous treatment effects in the labor supply experiment - Between subject

| | *Dependent Variable: Effort* | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Condition High | 13.392 | 9.529 | 5.757 |
| | (2.924) | (3.299) | (6.237) |
| Reduced form ($\hat{l}_i$) | | -4.176 | |
| | | (1.322) | |
| Condition High * Reduced form ($\hat{l}_i$) | | 6.554 | |
| | | (2.357) | |
| $E[\hat{\lambda}_i]$ | | | -2.464 |
| | | | (2.655) |
| Condition High * $E[\hat{\lambda}_i]$ | | | 4.782 |
| | | | (4.097) |
| Constant (Condition Low) | 33.444 | 36.219 | 38.697 |
| | (1.973) | (2.276) | (4.187) |
| # Observations | 500 | 500 | 451 |
| $H_0$ : Zero Treatment Effect (H-L) | $F_{1,499} = 20.97$ | $F_{1,499} = 8.34$ | $F_{1,450} = 0.85$ |
| | ($p < 0.01$) | ($p < 0.01$) | ($p = 0.36$) |
| $H_0$ : Gain-Loss Attitudes $\perp$ Effort in Low | | $F_{1,499} = 9.97$ | $F_{1,450} = 0.86$ |
| | | ($p < 0.01$) | ($p = 0.35$) |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,499} = 7.73$ | $F_{1,450} = 1.36$ |
| | | ($p < 0.01$) | ($p = 0.24$) |

*Notes*: This table replicates Table 1 using a between-subjects analysis. This means that we only use the first decision (either Condition Low or Condition High) that each subject answered. While we have less statistical power with this specification, note that all coefficients are very similar to those reported in Table 1.

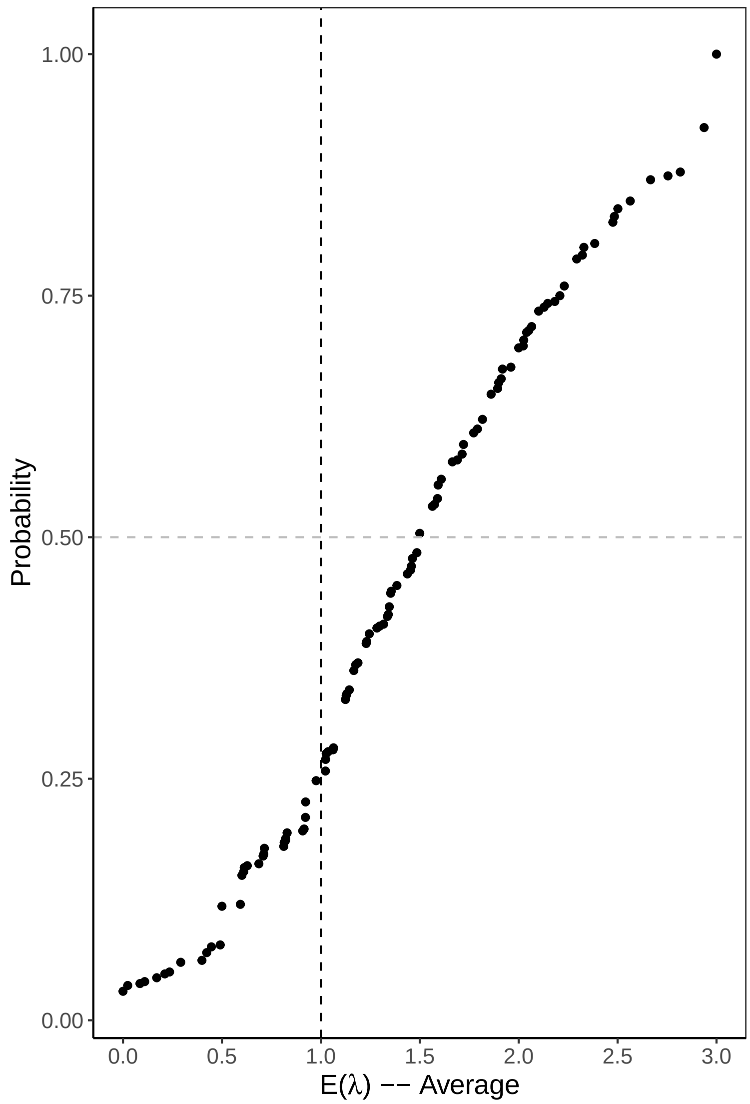



Figure A3: Distribution of $E[\hat{\lambda}_i]$ estimated from subjects' lottery choices

## A.6 Reconciliation with Pre-Analysis Plan of the Labor Study

In this section we report the methodology and corresponding analyses that we pre-registered for the labor supply study (Campos-Mercade et al. 2021, AEARCTR-0007277). By and large, the main text of the paper closely follows the pre-analysis plan. There are however two key points to discuss in order to reconcile the analyses in the main text with the pre-analysis plan: the number of subjects and the estimation of gain-loss attitudes (Stage 1). We discuss each of these points below, and we then show that using the pre-registered strategy for estimating gain-loss attitudes does not meaningfully change the estimation of the heterogeneous treatment effects (Stage 2).

### A.6.1 Sample size

Our power analyses showed that a sample between 500 and 600 subjects would give us enough power to detect the hypothesized heterogeneous treatment effects of the gain-loss attitudes. We hence pre-registered that we would gather between 500 and 800 subjects. The reason why we pre-registered a range is that we were unaware of how many subjects we would be able to recruit using the UC San Diego Economics Laboratory, online, and in the middle of the COVID-19 pandemic. Once we started collecting data we found out that recruiting subjects was harder than anticipated, with very few subjects signing up for our last sessions. We hence decided to stop as soon as we hit the pre-registered lower bound of 500 subjects.

### A.6.2 Stage 1: Identifying Gain-Loss Attitudes

As in the main text of the paper (Section 2.2), we pre-registered that we would use a maximum likelihood tobit method using $c_i(e_i) = \frac{1}{\gamma_i}(e + 10)^{\gamma_i}$ to estimate gain-loss attitudes. In the main text of the present paper, however, we go one step further and note that the estimated $\hat{\lambda}_i$ using this method are likely to be estimated with error. We then develop a standard Bayesian shrinkage exercise in Appendix A.2 and use the resulting $E[\hat{\lambda}_i]$ throughout the paper as our main gain-loss attitude measure.

The results are qualitatively similar whether we use $\hat{\lambda}_i$ or $E[\hat{\lambda}_i]$. Without adjusting for measurement error, the average value of $\hat{\lambda}_i$ is 1.31 and 41.24% have gain-seeking attitudes. Deploying the shrinkage adjustment described in Appendix A.2, we find an average value of $E[\hat{\lambda}_i] = 1.37$ and 30.8% have gain-seeking attitudes. Hence, regardless of the method that we use, we find similar average values of $\lambda$ and a sizable share of subjects with gain-seeking attitudes.

### A.6.3 Stage 2: Identifying Gain-Loss Attitudes

Table A4 performs the same analysis as in Table 1 but using unshrunk measures of $\lambda$. We further restrict the sample to those subjects for whom our structural model retrieves a value for $\hat{\lambda}_i$ (this is, 482 out of the 500 subjects).

Using this approach, we find qualitatively the same results as in the main text. Column 1 shows an aggregate specification, wherein treatment effects are assumed to be homogeneous. We find that the aggregate treatment effect of approximately 10 tasks is significant at all conventional levels, $F_{1,481} = 99.76$ ($p < 0.01$). In columns (2) and (3) we interact treatment with reduced form and structural measures of gain-loss attitudes for the relevant samples. Both measures are highly correlated with the effect of treatment and in both specifications we reject the null hypothesis of zero heterogeneity in treatment, $F_{1,481} = 15.08$ ($p < 0.01$) and $F_{1,450} = 17.95$ ($p < 0.01$), respectively. More loss averse subjects have greater increases in willingness to work as they move from the Low to the High condition.

Hence, while in order to account for noise we deviate from our exact pre-registered plan on how to calculate gain-loss attitudes, our results and conclusions remain unchanged.

Table A4: Heterogeneous treatment effects in the labor supply experiment

| | *Dependent Variable: Effort* | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Condition High | 10.523 | 7.933 | 2.630 |
| | (1.054) | (1.133) | (1.907) |
| Reduced form ($\hat{l}_i$) | | -3.369 | |
| | | (1.067) | |
| Condition High * Reduced form ($\hat{l}_i$) | | 4.477 | |
| | | (1.153) | |
| $\hat{\lambda}_i$ | | | -5.062 |
| | | | (1.651) |
| Condition High * $\hat{\lambda}_i$ | | | 5.908 |
| | | | (1.394) |
| Constant (Condition Low) | 35.871 | 37.821 | 42.634 |
| | (1.447) | (1.672) | (2.737) |
| $N$ | 964 | 964 | 964 |
| $H_0$ : Zero Treatment Effect (H-L) | $F_{1,481} = 99.76$ | $F_{1,481} = 49.06$ | $F_{1,481} = 1.90$ |
| | ($p < 0.01$) | ($p < 0.01$) | ($p = 0.17$) |
| $H_0$ : Gain-Loss Attitudes $\perp$ Effort in Low | | $F_{1,481} = 9.97$ | $F_{1,481} = 9.40$ |
| | | ($p < 0.01$) | ($p < 0.01$) |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,481} = 15.08$ | $F_{1,481} = 17.95$ |
| | | ($p < 0.01$) | ($p < 0.01$) |

*Notes*: Ordinary least squares regression explaining each subject's effort choice. Each subject provides two observations: one with their effort in Condition Low, and one with their effort in Condition H. The table is restricted to subjects for whom we our structural model retrieves their $\hat{\lambda}_i$ values (this is, 482 out of the 500 subjects). Clustered standard errors at the individual level in parentheses. The "Reduced form ($\hat{l}_i$)" measure captures the negative of the elasticity of labor supply to wage uncertainty, as estimated by $l_i$ in equation 3. Null hypotheses tested for 1) zero treatment effect (Condition High coefficient= 0); 2) no relationship between gain-loss attitudes and behavior in Condition Low behavior ($\hat{\lambda}_i$ or $\hat{l}_i = 0$); 4) constant treatment effect over gain-loss attitudes (Condition High* $\hat{\lambda}_i$ or Condition High* $\hat{l}_i = 0$). $F$-statistics and two-sided $p$-values reported.

# Appendix B Additional Development, Analysis, and Results for Exchange Study

## B.1 Theoretical Considerations for Exchange Study

We examine the predictions of the Kőszegi and Rabin (2006, 2007) EBRD formulation in simple exchange settings with two objects, recognizing heterogeneity of gain-loss attitudes. Consider a two-dimensional utility function over two objects of interest, object $X$ and object $Y$. Let $\mathbf{c} = (m_X, m_Y)$ and $\mathbf{r} = (r_X, r_Y)$ represent vectors of intrinsic utility and reference utility, respectively. As described in Section 3.2.1, overall utility is described by

$$u_i(\mathbf{c}|\mathbf{r}) = m_X + \mu_i(m_X - r_X) + m_Y + \mu_i(m_Y - r_Y),$$

where

$$\mu(z) = \begin{cases} \eta z & \text{if } z \geq 0 \\ \eta\lambda_i z & \text{if } z < 0. \end{cases}$$

In this piece-wise linear gain-loss function, the parameter $\eta$ captures the magnitude of changes relative to the reference point, and $\lambda_i$ captures individual gain-loss attitudes. If $\lambda_i > 1$, the individual is loss-averse, experiencing losses more than commensurately-sized gains. If $\lambda_i < 1$, the individual is gain-seeking, experiencing gains more than commensurately-sized losses. We normalize $\eta = 1$ for all individuals and restrict consumption utilities, $X$ and $Y$ to be homogeneous. We explore heterogeneity of consumption utilities in our estimation exercises of Appendix B.2.

### B.1.1 Choice Acclimating Personal Equilibrium (CPE)

Unless exogenously determined, the vector $\mathbf{r}$ is established as part of a consistent forward-looking plan for behavior. Kőszegi and Rabin (2006, 2007) posit a reference-dependent expected utility function $U_i(F|G)$, taking as input a distribution $F$ over consumption outcomes, $\mathbf{c}$, which are valued relative to a distribution $G$ of reference points, $\mathbf{r}$. That is

$$U_i(F|G) = \int \int u_i(\mathbf{c}|\mathbf{r}) dF(\mathbf{c}) dG(\mathbf{r}).$$

A *Personal Equilibrium* is a situation where, given that the decision-maker expects as a referent some distribution, $F$, they indeed prefer $F$ as a consumption distribution over all alternative consumption distributions, $F'$. Ex-ante optimal behavior has to accord with expectations of that behavior. Formally, given a choice set, $\mathcal{D}$, of lotteries, $F$, over consumption outcomes $\mathbf{c} = (m_X, m_Y)$, KR's *Personal Equilibrium* states the following:
*Personal Equilibrium (PE):* A choice $F \in \mathcal{D}$, is a personal equilibrium if

$$U_i(F|F) \geq U_i(F'|F) \ \forall \ F' \in \mathcal{D}.$$

Regardless of endowment, if object X is to be chosen in a PE, then $\mathbf{r} = (X, 0)$, and if object Y is to be chosen in a PE then $\mathbf{r} = (0, Y)$.

Given the potential for the multiplicity of PE selections, the KR model is constructed with a notion of equilibrium refinement, *Preferred Personal Equilibrium* (PPE), and an alternate non-PE criterion, *Choice-Acclimating Personal Equilibrium* (CPE). In both of these constructs, ex-ante utility is used as a basis for selection and, hence, for making more narrow predictions. For ease of explication, we focus our analysis on the CPE criterion. We also provide theoretical analyses under the PE and PPE approaches. Importantly, all three formulations share common comparative statics, and therefore make qualitatively similar predictions, for our KR test.

Given a choice set, $\mathcal{D}$, of lotteries, $F$, over consumption outcomes $\mathbf{c} = (m_X, m_Y)$, *Choice-Acclimating Personal Equilibrium* states the following:
*Choice-Acclimating Personal Equilibrium (CPE):* A choice $F \in \mathcal{D}$, is a choice-acclimating personal equilibrium if

$$U_i(F|F) \geq U_i(F'|F') \ \forall \ F' \in \mathcal{D}.$$

Under CPE, an individual selects between options like $[\mathbf{c},\mathbf{r}] = [(X,0),(X,0)]$ and $[\mathbf{c},\mathbf{r}] = [(0,Y),(0,Y)]$.[32]

### B.1.2 CPE Comparative Statics

The CPE concept noted above requires consistency between the distributions of $\mathbf{c}$ and $\mathbf{r}$. We consider a baseline simple exchange condition, Condition Low, for an individual endowed with object $X$. We focus on the choice set consisting of pure strategy choices $\mathcal{D} = \{(X,0),(0,Y)\}$, with the first element reflecting choosing not to exchange and the second choosing to exchange.

In this setting, there are two potential CPE selections, $[\mathbf{c},\mathbf{r}] = [(X,0),(X,0)]$ and $[\mathbf{c},\mathbf{r}] = [(0,Y),(0,Y)]$. The individual can support not exchanging, $[\mathbf{c},\mathbf{r}] = [(X,0),(X,0)]$, in a CPE if

$$U_i(X,0|X,0) \geq U_i(0,Y|0,Y),$$

which, under our functional form assumptions, becomes

$$X_{L,CPE,i} \geq Y. \tag{7}$$

Figure A4 graphs the Condition Low CPE cutoff, $\underline{X}_{L,CPE} = Y$, the smallest value of $X$ at which the individual can support not exchanging, which is constant for all values of the gain-loss parameter, $\lambda$. The value $\underline{X}_{L,CPE} = Y$ implies that choice in Condition Low is governed only by intrinsic utility. This represents the inability of CPE to rationalize the standard endowment effect. This prediction is not shared by the PE formulation, wherein the value of gain-loss attitudes tunes the set of permissible PE choices and can lead to an endowment effect (see below). Nonetheless, the critical comparative static shared by both formulations is delivered by comparing exchange behavior in this baseline Condition Low with Condition High's probabilistic forced exchange.

[32] Note that a selection need not be PE in order to be CPE. The alternate concept, PPE requires $F$ and $F'$ to be PE, rather than simply elements of $\mathcal{D}$.

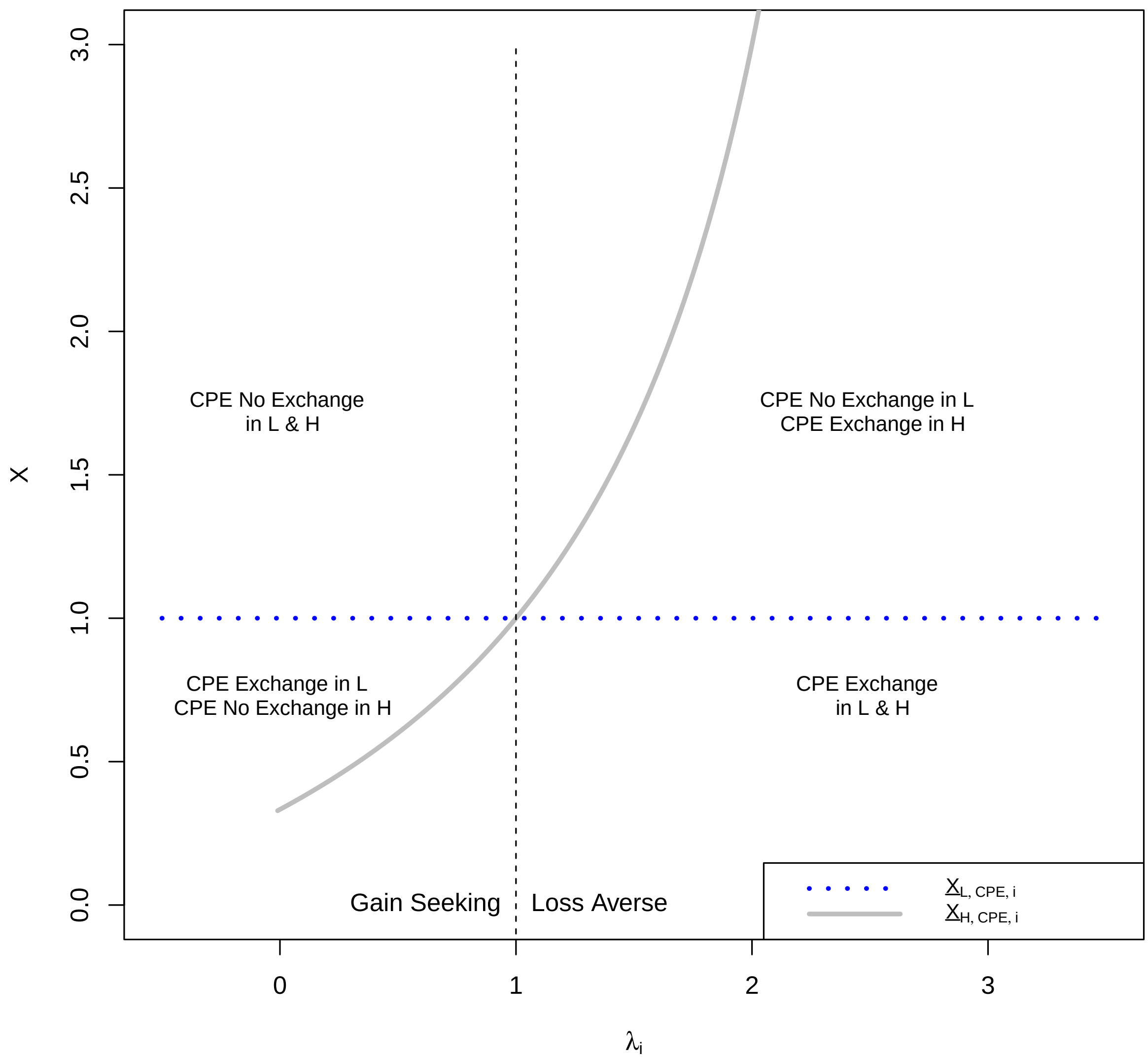



Figure A4: Gain-Loss Attitudes and Theoretical CPE Strategy Thresholds

*Notes:* Threshold values for CPE for agent endowed with $X$, assuming $Y = 1$ and $\eta = 1$.

Now, consider an environment of probabilistic forced exchange, Condition High. As shown in Section 3.2.1, agents can support attempting not to exchange as a CPE if

$$U_i(0.5(X,0) + 0.5(0,Y)|0.5(X,0) + 0.5(0,Y)) \geq U_i(0,Y|0,Y),$$

which, under our functional form assumptions, becomes

$$\begin{aligned} 0.5X + 0.5Y + 0.25(1-\lambda_i)(X+Y) &\geq Y \\ X_{H,CPI,i} &\geq \frac{1+0.5(\lambda_i - 1)}{1+0.5(1-\lambda_i)}Y. \end{aligned}$$

The manipulation of probabilistic forced exchange changes the CPE threshold from $\underline{X}_{L,CPE,i} = Y$ to $\underline{X}_{H,CPE,i} = \frac{1+0.5(\lambda_i-1)}{1+0.5\eta(1-\lambda)}Y$. Figure A4 illustrates the changing CPE cutoff values associated with not exchanging. In Condition High, the individual can support attempting to retain $X$ in CPE on the basis of both intrinsic utility and gain-loss attitudes.

The gain-loss parameter, $\lambda_i$, tunes precisely how behavior should change between Conditions Low and High. Figure A4 is partitioned into four regions. Two critical regions of changing CPE choice are identified. For $X > Y$ and $\lambda_i > 1$, it is CPE to not exchange in Condition Low, and CPE to exchange in Condition High. This region has been the basis of prior experimental tests under the assumption of universal loss aversion; such individuals become more willing to exchange when probabilistically forced. Ignored to date is the region where $X < Y$ and $\lambda_i < 1$. In this region, it is CPE to exchange in Condition Low, and CPE to not exchange in Condition High. In contrast to the loss-averse prediction, such gain-seeking individuals become less willing to exchange when probabilistically forced. The KR comparative static for the difference between Condition Low and Condition High changes sign at $\lambda_i = 1$.

### B.1.3 Additional Theoretical Analysis: PE and PPE

We now provide additional theoretical development for heterogeneity in response to probabilistic forced exchange under Personal Equilibrium (PE) and the PE refinement, Preferred Personal Equilibrium, PPE. Throughout, our maintained assumptions will be $X, Y, \lambda_i, \eta > 0$. We begin with the restrictions on behavior implied by PE. To begin, we focus on Condition Low and a choice set consisting of pure strategy choices $\mathcal{D} = \{(X, 0), (0, Y)\}$. In this setting, there are two potential PE selections, $[\mathbf{c}, \mathbf{r}] = [(X, 0), (X, 0)]$ and $[\mathbf{c}, \mathbf{r}] = [(0, Y), (0, Y)]$. The individual can support not exchanging, $[\mathbf{c}, \mathbf{r}] = [(X, 0), (X, 0)]$, in a PE if

$$U_i(X, 0|X, 0) \geq U_i(0, Y|X, 0),$$

or

$$X_{L,PE,i} \geq \frac{2}{1 + \lambda_i} Y. \tag{8}$$

Note that the smallest value of $X$ at which the individual can support not exchanging, $\underline{X}_{L,PE,i} = \frac{2}{1+\lambda_i} Y$, is inferior to $Y$ if $\lambda_i > 1$. As such, loss-averse individuals with $\lambda_i > 1$ may be able support not exchanging $X$ for $Y$ even if $Y$ would be preferred on the basis of intrinsic utility alone. This describes the mechanism by which the KR model generates an endowment effect in PE. Similarly, the individual can support exchanging, $[\mathbf{c}, \mathbf{r}] = [(0, Y), (0, Y)]$, if

$$U_i(0, Y|0, Y) \geq U_i(X, 0|0, Y),$$

or

$$X_{L,PE,i} \leq \frac{1 + \lambda_i}{2} Y.$$

The highest value of $X$ at which the agent can support exchanging, $\overline{X}_{L,PE,i} = \frac{1+\lambda_i}{2} Y$, increases linearly with $\lambda$. For $\underline{X}_{L,PE,i} \leq X \leq \overline{X}_{L,PE,i}$, there will be multiple equilibria, with the agent able to support both exchanging and not exchanging as a PE.

Note that for gain-seeking individuals with $\lambda_i < 1$ it is also possible for $\overline{X}_{L,PE,i} < X < \underline{X}_{L,PE,i}$, such that no pure strategy PE selection from the assumed $\mathcal{D}$ exists. In this region, if $\mathcal{D}$ were to include all mixtures of exchanging and not exchanging, there would be a mixed strategy PE of not exchanging with a given probability, $p$. Below, we provide this analysis. Figure A5 provides the pure strategy PE cutoffs associated with exchanging not exchanging in Condition Low.

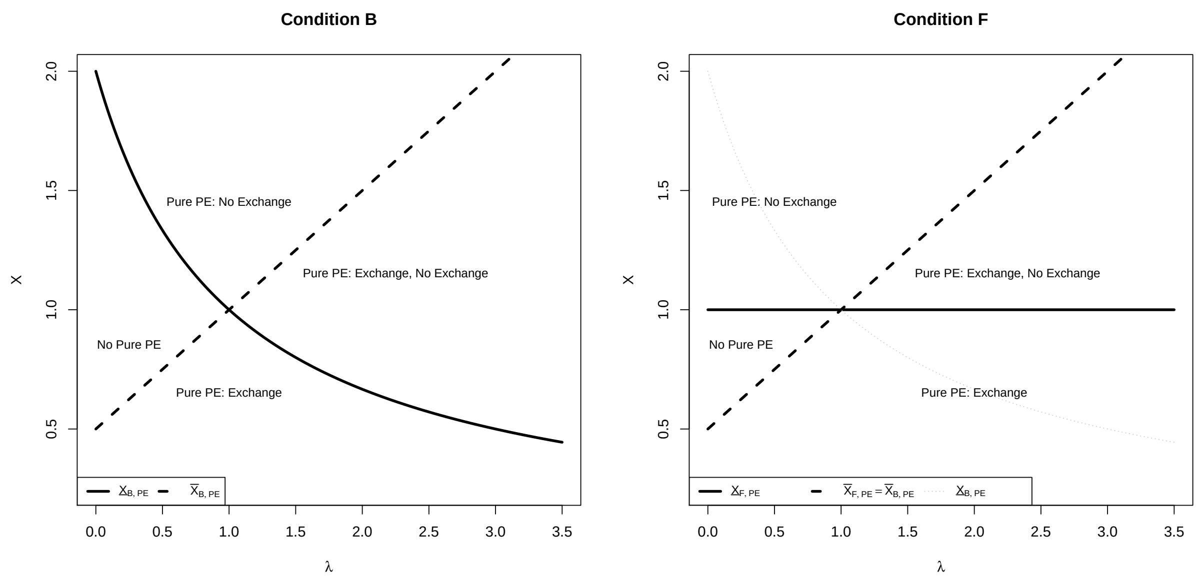



Figure A5: Gain-Loss Attitudes and Theoretical Pure PE Strategy Thresholds

*Notes:* Threshold values for pure strategy PE for agent endowed with $X$, assuming $Y = 1$ and $\eta = 1$.

Now, consider Condition High. The potential selections for someone endowed with $X$ are $\mathcal{D} = \{0.5(X,0) + 0.5(0,Y), (0,Y)\}$, with the first element reflecting attempting not to exchange and the second reflecting exchange, as before. The individual can support attempting not to exchange in a PE if

$$U_i(0.5(X,0) + 0.5(0,Y)|0.5(X,0) + 0.5(0,Y)) \geq U_i(0,Y|0.5(X,0) + 0.5(0,Y)),$$

or

$$X_{H,PE,i} \geq Y. \tag{9}$$

Under forced exchange, the individual can support attempting to retain $X$ in PE only on the basis of intrinsic utility values, regardless of the level of $\lambda$.

Though probabilistic forced exchange alters the PE considerations associated with not exchanging, it leaves unchanged the PE considerations associated with exchanging. The individual can support exchanging in PE if

$$U_i(0, Y|0, Y) \geq U_i(0.5(X, 0) + 0.5(0, Y)|0, Y),$$

which as before is

$$X_{H,PE,i} \leq \frac{1 + \lambda_i}{2} Y.$$

Hence, $\overline{X}_{H,PE,i} = \overline{X}_{L,PE,i}$.

The manipulation of probabilistic forced exchange changes the PE cutoff for not exchanging from $\underline{X}_{L,PE,i} = \frac{2}{1+\lambda_i} Y$ to $\underline{X}_{H,PE,i} = Y$. There is no longer any possibility in PE for a loss-averse individual to support keeping their object if $Y > X$. A loss-averse individual with $\lambda_i > 1$ and valuation $\underline{X}_{L,PE,i} < X < \underline{X}_{H,PE,i}$ moves from a position of multiple PE in Condition Low, to having a unique PE to exchange in Condition High. Such an individual plausibly grows more willing to exchange when moving from Condition Low to Condition High. Similarly, a gain-seeking individual with $\lambda_i < 1$ and valuation $\underline{X}_{H,PE,i} < X < \underline{X}_{B,PE,i}$ moves from a position of no pure strategy PE in Condition Low to having a unique PE of exchange in Condition High. Such an individual plausibly grows less willing to exchange when moving from Condition Low to Condition High. Figure A5, illustrates these changing pure strategy PE considerations from Condition High to Condition Low. The direction of these comparative statics is identical to that of our CPE analysis in the main text.

### B.1.4 PE Mixed Strategy Analysis

To provide more complete analysis, particularly when there is no pure strategy PE, we now elaborate PE and PPE formulations when the choice set $\mathcal{D}$ includes all available mixtures of exchanging and not exchanging. For Condition Low , we assume $\mathcal{D}_B = \{p \in [0, 1] : p(X, 0) + (1 - p)(0, Y)\}$, allowing all mixtures of exchange and no exchange to be chosen. A given mixture, $p$, will be PE if

$$U_i(p(X,0)+(1-p)(0,Y)|p(X,0)+(1-p)(0,Y)) \geq$$
$$U_i(q(X,0)+(1-q)(0,Y)|p(X,0)+(1-p)(0,Y)) \;\forall\; q \;\in [0,1],$$

or

$$pX+(1-p)Y+p(1-p)(1-\lambda_i)(X+Y) \geq$$
$$qX+(1-q)Y+(1-q)p(Y-\lambda_X)+q(1-p)(X-\lambda_i Y) \;\forall\; q \in [0,1].$$

For a given $p$, let $\mathbf{q}^*(p) \equiv \{argmax_q U_i(q,p)\} \equiv \{argmax_q U_i(q(X,0)+(1-q)(0,Y)|p(X,0)+(1-p)(0,Y))\}$. The brackets indicate that $\mathbf{q}^*(p)$ may be a set. A mixture, $p \in [0,1]$, is PE if $p \in \mathbf{q}^*(p)$.

Note that

$$\begin{aligned}\frac{\partial U_i(q,p)}{\partial q} &= X-Y-p(Y-\lambda_i X)+(1-p)(X-\lambda_i Y)\\ &= 2X-(1+\lambda_i)Y-p(1-\lambda_i)(Y+X)\end{aligned}$$

is constant for a given $p$, as $U(q,p)$ is linear in $q$. If $\frac{\partial U_i(q,p)}{\partial q} > (<)\ 0$, then it will attain a unique maximum $\mathbf{q}^*(p) = \{1\}(\{0\})$. As such, any strict mixtures, $p \in (0,1)$, for which $\frac{\partial U_i(q,p)}{\partial q} \neq 0$ cannot be PE. Note that this development implies that not exchanging with certainty, $p=1$, will be PE if $\frac{\partial U_i(q,1)}{\partial q} \geq 0$, or

$$\begin{aligned}2X-(1+\lambda_i)Y-(1-\lambda_i)(Y+X) &\geq 0,\\ X &\geq \frac{2}{(1+\lambda_i)}Y,\end{aligned}$$

which corresponds to the pure strategy threshold noted above, $\underline{X}_{L,PE,i}$. Similarly, exchanging with certainty, $p = 0$, will be PE if $\frac{\partial U_i(q,0)}{\partial q} \leq 0$, or

$$
\begin{aligned}
2X - (1+\lambda_i)Y &\leq 0 \\
X &\leq \frac{(1+\lambda_i)}{2}Y,
\end{aligned}
$$

which corresponds to the pure strategy threshold, $\overline{X}_{L,PE,i}$. For values of $X$ such that

$$\frac{2}{(1+\lambda_i)}Y \leq X \leq \frac{(1+\lambda_i)}{2}Y,$$

$p = 1$ and $p = 0$ will be PE.

Strict mixtures, $p \in (0,1)$, for which $\frac{\partial U_i(q,p)}{\partial q} = 0$, $p \in \mathbf{q}^*(p)$, as all values of $q$, including $q = p$, attain the maximum. For each parameter constellation, $X$, $Y$, $\lambda_i$, if there exists a candidate mixture

$$p \in (0,1) \ s.t \ p = \frac{2X - (1+\lambda_i)Y}{(1-\lambda_i)(Y+X)}$$

such a $p$ is PE. Note that there will be at most one strict mixture PE. This strict mixture will be a proper probability provided $\frac{2X-(1+\lambda_i)Y}{(1-\lambda_i)(Y+X)} \in (0,1)$. For such a proper mixture probability to exist for $\lambda_i > 1$, it must be that

$$\frac{2}{(1+\lambda_i)}Y < X < \frac{(1+\lambda_i)}{2}Y.$$

That is, if $\lambda_i > 1$, both pure strategies, $p = 0$ and $p = 1$, are PE, and the required preferences are strict, there will also be a strict mixture PE. In contrast, for such a proper probability mixture to exist for $\lambda_i < 1$, it must be that

$$\frac{(1+\lambda_i)}{2}Y < X < \frac{2}{(1+\lambda_i)}Y.$$

That is, if $\lambda_i < 1$, and neither pure strategy, $p = 0$ or $p = 1$, are PE, there will be a strict mixture PE.

Figure A6 summarizes the PE considerations in Condition Low recognizing the possibility of mixed strategy equilibria with the corresponding value of the mixture probability noted. In contrast to the pure strategy analysis of Figure A5, for $\lambda_i < 1$ within the

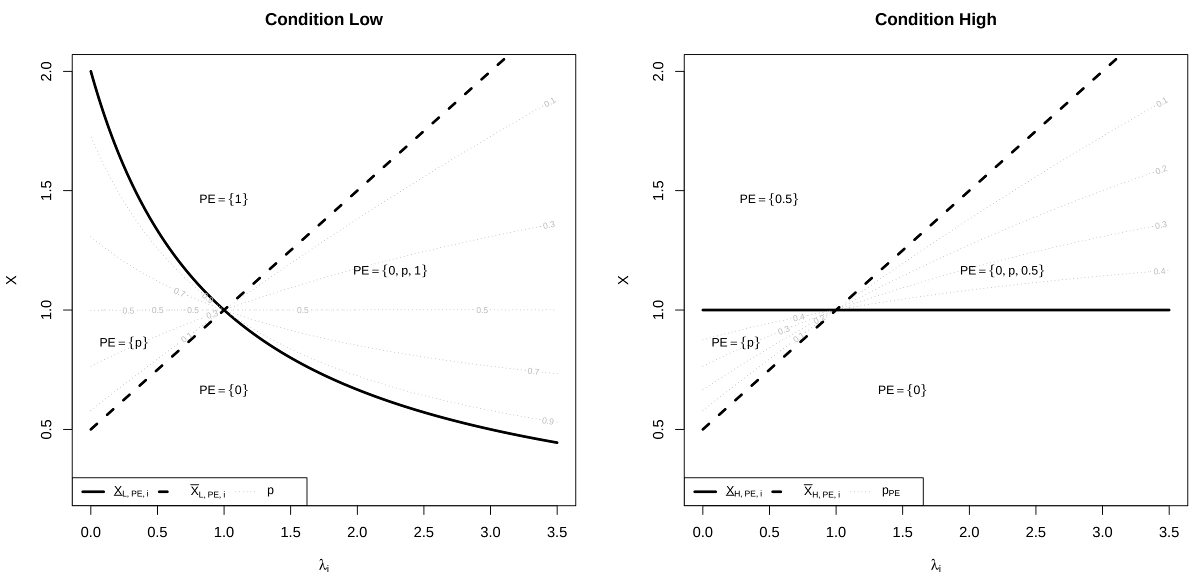



Figure A6: Gain-Loss Attitudes and Theoretical PE Strategy Thresholds

*Notes:* Threshold values for mixed strategy PE for agent endowed with $X$, assuming $Y = 1$ and $\eta = 1$.

bounds $\frac{(1+\lambda_i)}{2)}Y < X < \frac{2}{(1+\lambda_i)}Y$, there is now a mixed strategy PE. Further, for $\lambda_i > 1$ and $\frac{2}{(1+\lambda_i)}Y < X < \frac{(1+\lambda_i)}{2}Y$ there are three equilibria when accounting for potential mixtures.

Having elaborated the PE restrictions for Condition Low, we proceed to Condition High. Condition High alters the choice set from $\mathcal{D}_L = \{p \in [0,1] : p(X,0) + (1-p)(0,Y)\}$ to $\mathcal{D}_H = \{p \in [0,0.5] : p(X,0) + (1-p)(0,Y)\}$. This alteration induces two potential changes to the PE calculus. First, potential PE choices from Condition Low may not be available in Condition High. Second, lotteries, $q$, that prevent a specific $p$ from being PE may potentially be eliminated.

In Condition High, a given mixture $p \in [0, 0.5]$ will be PE if

$$\begin{aligned} U_i(p(X,0) + (1-p)(0,Y)|p(X,0) + (1-p)(0,Y)) \geq \\ U_i(q(X,0) + (1-q)(0,Y)|p(X,0) + (1-p)(0,Y)) \ \forall \ q \ \in [0, 0.5]. \end{aligned}$$

As before $U(q,p)$ is linear in $q$, and so a boundary strategy of attempting to keep one's object, $(p = 0.5)$ will be PE if

$$\begin{aligned} \frac{\partial U_i(q, 0.5)}{\partial q} = 2X - (1+\lambda_i)Y - 0.5(1-\lambda_i)(Y+X) \geq 0 \\ (1 + 0.5(1+\lambda_i))X \geq (1 + 0.5(1+\lambda_i))Y \\ X \geq Y, \end{aligned}$$

which corresponds to the pure strategy threshold, $\underline{X}_{H,PE,i}$. Similarly, exchanging with certainty, $p = 0$, will be be PE if

$$\begin{aligned} \frac{\partial U(q, 0)}{\partial q} = 2X - (1+\lambda_i)Y \leq 0 \\ X \leq \frac{(1+\lambda_i)}{2} Y, \end{aligned}$$

which corresponds to the pure strategy threshold, $\overline{X}_{H,PE,i} = \overline{X}_{L,PE,i}$.

Again strict mixtures, $p \in (0, 0.5)$, for which $\frac{\partial U_i(q,p)}{\partial q} = 0$, $p \in \mathbf{q}^*(p)$, as all values of $q$, including $q = p$, attain the maximum. For each parameter constellation, $X$, $Y$, $\eta$, $\lambda$, if there exists a candidate mixture

$$p \in (0, 0.5) \ s.t \ p = \frac{2X - (1+\lambda_i)Y}{(1-\lambda_i)(Y+X)}$$

such a $p$ is PE. Note that there will be at most one strict mixture PE. This strict mixture will be a proper probability and within the choice set provided $\frac{2X-(1+\lambda_i)Y}{(1-\lambda_i)(Y+X)} \in (0, 0.5)$. For such a proper mixture probability to exist for $\lambda_i > 1$, it must be that

$$Y < X < \frac{(1+\lambda_i)}{2}Y$$

That is, if $\lambda_i > 1$, both pure strategies, $p = 0$ and $p = 0.5$, are PE, and the required preferences are strict, there will also be a strict mixture PE. In contrast, for such a proper probability mixture to exist for $\lambda_i < 1$, it must be that

$$\frac{(1+\lambda_i)}{2}Y < X < Y.$$

That is, if $\lambda_i < 1$, and neither pure strategy, $p = 0$ or $p = 0.5$, are PE, there will be a strict mixture PE.

Figure A6 summarizes the PE considerations in Condition High recognizing the possibility of mixed strategy equilibria with the corresponding value of the mixture probability noted. Moving from Condition Low to Condition High all mixed strategy PE with $p \in (0.5, 1)$ are eliminated from the choice set. Individuals with $\lambda_i > 1$ and multiple equilibria, $PE = \{0, p > 0.5, 1\}$ in Condition Low have a unique $PE = \{p = 0\}$ in Condition High. Such individuals may exchange less than 100 percent of the time in Condition Low and do so 100 percent of the time in Condition High, growing more willing to exchange. In contrast, individuals with $\lambda_i < 1$ and a unique $PE = \{p > 0.5\}$ in Condition Low, have a unique $PE = \{p = 0.5\}$ in Condition High. Such individuals would attempt to retain their object less than 100 percent of the time in Condition Low and would do so 100 percent of the time in Condition High, growing less willing to exchange. This analysis highlights exactly the intuition laid out with our prior pure strategy analysis and that for the CPE concept. We next turn to PPE analysis to select among multiple PE selections.

### B.1.5 Preferred Personal Equilibrium Analysis

Where there exist multiple PE selections, the KR model is equipped with an equilibrium selection mechanism, *Preferred Personal Equilibrium* (PPE). PPE selects among PE values on the basis of ex-ante utility. Having elaborated the PE values in the Figure A6, it is straightforward to identify the selection, $p$, with the highest value of $U_i(p(X,0)+(1-p)(0,Y)|p(X,0)+(1-p)(0,Y)) = pX+(1-p)Y+p(1-p)\eta(1-\lambda_i)(X+Y)$. In the case of Condition Low, there is a region of multiplicity for $\lambda_i > 1$ where the set of $PE = \{0, p \in (0,1), 1\}$. In this region it is clear that not exchanging, $p = 1$, will yield higher ex-ante utility than exchanging, $p = 0$, if

$$X > Y.$$

If $X > Y$, $p = 1$ will also yield higher ex-ante utility than any PE mixture $p \in (0,1)$ as all mixtures will both lower intrinsic utility (as $X > Y \rightarrow X > pX+(1-p)Y \forall\, p \in (0,1)$) and expose the individual to the overall negative sensations of gain loss embodied in the term $p(1-p)(1-\lambda_i)(X+Y) < 0$ for $\lambda_i > 1$. Following this logic, in Condition Low, multiplicity is resolved via PPE by selecting either $p = 1$ if $X > Y$ or $p = 0$ if $X < Y$.

Similarly, in Condition High, there is a region of multiplicity for $\lambda_i > 1, Y < X < \frac{(1+\lambda_i)}{2}Y$ where the set of $PE = \{0, p \in (0, 0.5), 0.5\}$. Note that for $\lambda_i > 1$, if $X < \frac{(1+\lambda_i)}{2}Y$, then $X < \frac{(1+0.5(\lambda_i-1))}{(1+0.5(1-\lambda_i))}Y = \frac{(1+\lambda_i-0.5(\lambda_i+1))}{(2-0.5(\lambda_i+1))}Y$. That is, in this region of multiplicity, $X$ is below the $X_{H,CPE,i}$ cutoff noted in the main text. Hence, we know that exchanging, $p = 0$, yields higher ex-ante utility than attempting not to exchange, $p = 0.5$, in this region. It suffices to check which of the remaining PE selections $\{0, p = \frac{2X-(1+\lambda_i)Y}{(1-\lambda_i)(Y+X)} \in (0, 0.5)\}$ provide higher utility. For this key mixture,

$$p = \frac{2X-(1+\lambda_i)Y}{(1-\lambda_i)(Y+X)}$$

$$(1-p) = \frac{(1-\lambda_i)(Y+X)}{(1-\lambda_i)(Y+X)} - \frac{2X-(1+\lambda_i)Y}{(1-\lambda_i)(Y+X)}$$

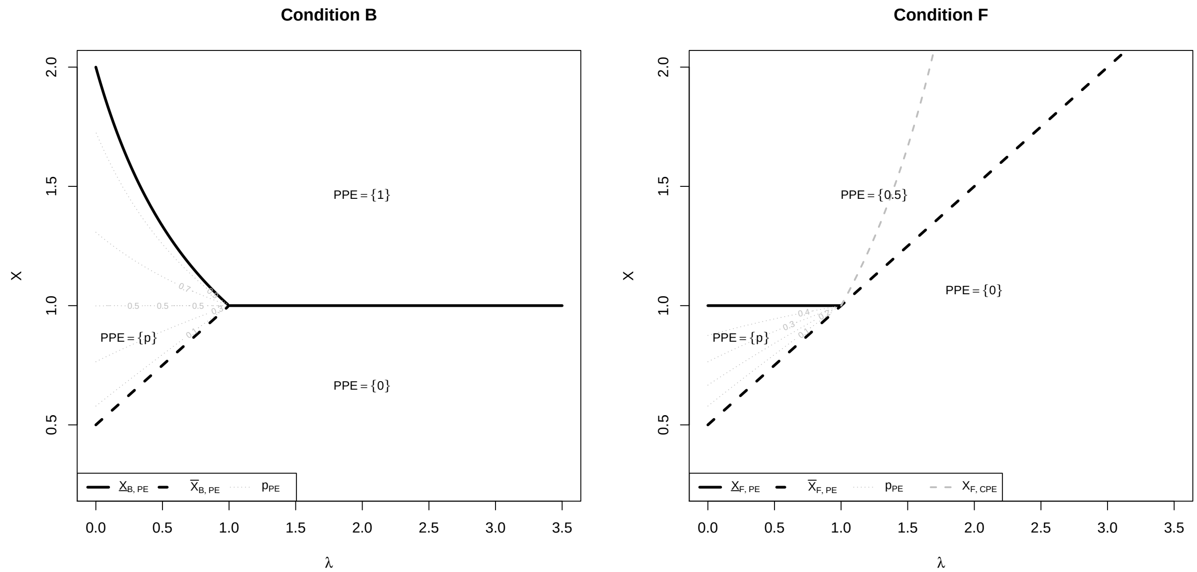



Figure A7: Gain-Loss Attitudes and Theoretical PPE Strategy Thresholds

*Notes:* Threshold values for PPE for agent endowed with $X$, assuming $Y = 1$ and $\eta = 1$.

The PPE selection will be $p = 0$ provided

$$
\begin{aligned}
Y &> pX + (1-p)Y + p(1-p)(1-\lambda_i)(X+Y) \\
Y &> X + (1-p)(1-\lambda_i)(X+Y) \\
Y &> X + \left[\frac{(1-\lambda_i)(Y+X)}{(1-\lambda_i)(Y+X)} - \frac{2X-(1+\lambda_i)Y}{(1-\lambda_i)(Y+X)}\right](1-\lambda_i)(X+Y) \\
Y &> X + [(1-\lambda_i)(Y+X) - 2X + (1+\lambda_i)Y] \\
Y - (1+\lambda_i)Y - (1-\lambda_i)Y &> X + (1-\lambda_i)(X) - 2X \\
-Y &> -\lambda_i X \\
X &> \frac{1}{\lambda_i}Y,
\end{aligned}
$$

Which is satisfied as $X > Y$ and $\lambda_i > 1$ in this region.

Figure A7 summarizes the PPE considerations in Conditions Low and High recognizing the possibility of a mixed strategy PPE with the corresponding value of the mixture probability noted. Also graphed in Figure A7 is the relevant CPE cutoff for $\lambda_i > 1$ in Condition

High to reinforce both that in the region of multiplicity exchanging, $p = 0$, yields higher ex-ante utility than attempting not to exchange, $p = 0.5$, and that the restrictions on behavior differ meaningfully between CPE and PPE. Nonetheless, both solution concepts share the same directional comparative statics that individuals with $\lambda_i > 1$ should grow more willing to exchange moving from Condition Low to Condition High, while individuals with $\lambda_i < 1$ should grow less-so.

## B.2 Estimation and Calculation of Gain Loss Attitudes In Exchange Study

In this appendix, we provide the likelihood formulation for our mixed-logit methodology to estimate heterogeneity in gain-loss attitudes and utilities. There are three relative preference statements that subjects provide in Stage 1: relative wanting statements, relative liking statements, and hypothetical choice. Let $i = 1, ..., N$ represent the index for subjects, and let $\{w, l, h\}$ represent the index of the three preference statements, referring to (w)anting, (l)iking, and (h)ypothetical choice, respectively. Let $w, l \in \{-1, 0, 1\}$ correspond to providing a higher rating for the alternative object, providing equal ratings for both objects, and providing a higher rating for the endowed object, respectively. Let $h \in \{-1, 1\}$ correspond to hypothetically choosing the alternative object or the endowed object, respectively.

We begin by presenting a standard logit formulation and then extend to the mixed logit case. Let $G(\cdot)$ represent the CDF of the logistic distribution. For each individual there are three potential probabilities associated with the three potential wanting ratings for those endowed with $X$, $Prob_{w_i,X}$,

$$\begin{array}{llr} Prob_{w_i,X} = & G((1+\lambda_i) - 2\frac{Y}{X} - \delta_X) & if\ w_i = 1 \\ Prob_{w_i,X} = & G(2\frac{Y}{X} - (1+\lambda_i) - \delta_X) & if\ w_i = -1 \\ Prob_{w_i,X} = & 1 - G((1+\lambda_i) - 2\frac{Y}{X} - \delta_X) - G(2\frac{Y}{X} - (1+\lambda_i) - \delta_X) & if\ w_i = 0, \end{array}$$

and three for those endowed with $Y$, $Prob_{w_i,Y}$,

$$\begin{aligned}
Prob_{w_i,Y} = & \quad G(2-(1+\lambda_i)\tfrac{Y}{X}-\delta_X) & if\ w_i=-1 \\
Prob_{w_i,Y} = & \quad G((1+\lambda_i)\tfrac{Y}{X}-2-\delta_X) & if\ w_i=1 \\
Prob_{w_i,Y} = & \quad 1-G(2-(1+\lambda_i)\tfrac{Y}{X}-\delta_X)-G((1+\lambda_i)\tfrac{Y}{X}-2-\delta_X) & if\ w_i=0.
\end{aligned}$$

Similarly, there are three potential probabilities associated with the three potential liking ratings for those endowed with $X$, $Prob_{l_i,X}$,

$$\begin{aligned}
Prob_{l_i,X} = & \quad G((1+\lambda_i)-2\tfrac{Y}{X}-\delta_X) & if\ l_i=1 \\
Prob_{l_i,X} = & \quad G(2\tfrac{Y}{X}-(1+\lambda_i)-\delta_X) & if\ l_i=-1 \\
Prob_{l_i,X} = & \quad 1-G((1+\lambda_i)-2\tfrac{Y}{X}-\delta_X)-G(2\tfrac{Y}{X}-(1+\lambda_i)-\delta_X) & if\ l_i=0,
\end{aligned}$$

and three for those endowed with $Y$, $Prob_{l_i,Y}$,

$$\begin{aligned}
Prob_{l_i,Y} = & \quad G(2-(1+\lambda_i)\tfrac{Y}{X}-\delta_X) & if\ l_i=-1 \\
Prob_{l_i,Y} = & \quad G((1+\lambda_i)\tfrac{Y}{X}-2-\delta_X) & if\ l_i=1 \\
Prob_{l_i,Y} = & \quad 1-G(2-(1+\lambda_i)\tfrac{Y}{X}-\delta_X)-G((1+\lambda_i)\tfrac{Y}{X}-2-\delta_X) & if\ l_i=0.
\end{aligned}$$

Lastly, there are two potential probabilities associated with the two hypothetical choice statements for those endowed with $X$, $Prob_{h_i,X}$,

$$\begin{aligned}
Prob_{h_i,X} = & \quad G((1+\lambda_i)-2\tfrac{Y}{X}) & if\ w_i=1 \\
Prob_{h_i,X} = & \quad G(2\tfrac{Y}{X}-(1+\lambda_i)) & if\ w_i=-1,
\end{aligned}$$

and two for those endowed with $Y$, $Prob_{h_i,Y}$,

$$\begin{aligned}
Prob_{h_i,Y} = & \quad G(2-(1+\lambda_i)\tfrac{Y}{X}) & if\ w_i=-1 \\
Prob_{h_i,Y} = & \quad G((1+\lambda_i)\tfrac{Y}{X}-2) & if\ w_i=1.
\end{aligned}$$

Let $\mathbf{1}_X$ indicate an individual endowed with object $X$. A single individual's choice probability would thus be

$$L_i = (Prob_{w_i,X} \cdot Prob_{l_i,X} \cdot Prob_{h_i,X})^{\mathbf{1}_X} \cdot (Prob_{w_i,Y} \cdot Prob_{l_i,Y} \cdot Prob_{h_i,Y})^{(1-\mathbf{1}_X)},$$

and the grand log likelihood would be

$$\mathcal{L} = \sum_{i=1}^{N} log(L_i)$$

Moving from this logit formulation to our mixed logit formulation is straightforward and follows Train (2009). For estimating the heterogeneity of gain-loss attitudes, we assume that the value $\lambda_i$ is drawn from a log-normal distribution with $log(\lambda_i) \sim N(\mu_{\lambda_i}, \sigma^2_{\lambda_i})$. Let $\theta \equiv (\mu_{\lambda_i}, \sigma^2_{\lambda_i})$, represent the parameters of this distribution, and let $f(\lambda_i|\theta)$ be the distribution of $\lambda_i$ given these parameters. A single individual's choice probabilities are thus

$$L_i = \int L_i(\lambda_i) f(\lambda_i|\theta) d\lambda_i$$

where $L_i(\lambda_i)$ is the individual choice probability evaluated at a given draw of $f(\lambda_i|\theta)$. We construct these choice probabilities through simulation. Let $r = 1, ..., R$ represent simulations of $\lambda_i$ from $f(\lambda_i|\theta)$ at a given set of parameters, $\theta$. Let $\lambda_i^r$ be the $r^{th}$ simulant. We simulate $L_i$ as

$$\check{L}_i = \frac{1}{R} \sum_{r=1}^{R} L_i(\lambda_i^r),$$

And these simulated probabilities replace the standard choice probabilities in the grand log likelihood to create a simulated log likelihood,

$$\mathcal{SL} = \sum_{i=1}^{N} log(\check{L}_i).$$

This simulated log likelihood is maximized to deliver estimates of $\mu_{\lambda_i}$ and $\sigma^2_{\lambda_i}$ alongside the homogeneous utility ratio $\frac{X}{Y}$.

When considering the possibility of heterogeneous utility rather than heterogeneous gain-loss attitudes, the exercise is analogous. We assume that the value $\frac{X}{Y}$ is drawn from a log-normal distribution with $log(\frac{X}{Y}) \sim N(\frac{X}{Y}, \sigma^2_{\frac{X}{Y}})$. Let $\theta' \equiv (\mu_{\frac{X}{Y}}, \sigma^2_{\frac{X}{Y}})$, represent the parameters of this distribution, and let $f(\frac{X}{Y}|\theta')$ be the distribution of $\frac{X}{Y}$ given these parameters. A single individual's choice probabilities are thus

$$L_i = \int L_i(\frac{X}{Y})f(\frac{X}{Y}|\theta')d\frac{X}{Y}$$

where $L_i(\frac{X}{Y})$ is the individual choice probability evaluated at a given draw of $f(\frac{X}{Y}|\theta')$. We construct these choice probabilities through simulation. Let $r = 1, ..., R$ represent simulations of $\frac{X}{Y}$ from $f(\frac{X}{Y}|\theta')$ at a given set of parameters, $\theta'$. Let $\frac{X}{Y}^r$ be the $r^{th}$ simulant. We simulate $L_i$ as

$$\check{L}_i = \frac{1}{R}\sum_{r=1}^{R} L_i(\frac{X}{Y}^r),$$

And these simulated probabilities replace the standard choice probabilities in the grand log likelihood to create a simulated log likelihood,

$$\mathcal{SL} = \sum_{i=1}^{N} log(\check{L}_i).$$

This simulated log likelihood is maximized to deliver estimates of $\mu_{\frac{X}{Y}}$ and $\sigma^2_{\frac{X}{Y}}$ alongside the homogeneous gain-loss parameter, $\lambda_i$. Operationally for implementing both of our simulated likelihood techniques we use 1000 Halton draws for each heterogeneous parameter and implement the code in Stata.

### B.2.1 Classifying Individual Gain-Loss Attitudes Accounting for Errors

Moving from the distribution of gain-loss attitudes to an expected value of $\lambda$ for each individual is a straightforward step after estimation. As proposed by Train (2009), we simulate the distribution of $\lambda$, and calculate the $E[\hat{\lambda}_i]$ for each possible Stage 1 statement profile. For example, under the estimated log-normal density, $g(\lambda)$, one simulates $Prob_{X|X}(\lambda)$, and

Table A5: Method of Simulated Likelihood Estimates

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Estimate | (Std. Error) | Estimate | (Std. Error) |
| | *Heterogeneous* $\lambda$ | | *Heterogeneous* $\frac{Y}{X}$ | |
| *Gain-Loss Attitudes:* | | | | |
| $\hat{\lambda}$ | 1.37 | (0.08) | 1.31 | (0.05) |
| $\hat{\mu_{\lambda}}$ | 0.17 | (0.07) | - | - |
| $\hat{\sigma^2_{\lambda}}$ | 0.29 | (0.21) | - | - |
| *Pair 1 Utilities* (USB Stick (X) - Pen Set (Y)) *:* | | | | |
| $\hat{\frac{Y}{X}}$ (Initial) | 0.62 | (0.04) | 0.62 | (0.03) |
| $\hat{\frac{Y}{X}}$ (Replication) | 0.61 | (0.04) | - | - |
| $\hat{\mu_{\frac{Y}{X}}}$ | - | - | -0.55 | (0.09) |
| $\hat{\sigma^2_{\frac{Y}{X}}}$ | - | - | 0.16 | (0.13) |
| *Pair 2 Utilities* (Picnic Mat (X) - Thermos (Y))*:* | | | | |
| $\hat{\frac{Y}{X}}$ (Initial) | 1.11 | (0.03) | 1.03 | (0.03) |
| $\hat{\frac{Y}{X}}$ (Replication) | 0.88 | (0.04) | - | - |
| $\hat{\mu_{\frac{Y}{X}}}$ | - | - | -0.03 | (0.04) |
| $\hat{\sigma^2_{\frac{Y}{X}}}$ | - | - | 0.12 | (0.08) |
| *Discernibility:* | | | | |
| $\delta_X$ | 0.55 | - | 0.55 | - |
| # Observations | 3,072 | | 3,072 | |
| Log-Likelihood | -2743.13 | | -2751.72 | |
| Akaike's Information Criterion (AIC) | 5498.26 | | 5513.44 | |

*Notes*: Method of simulated likelihood estimates. Standard errors in parentheses.

Table A6: Method of Simulated Likelihood Estimates: Sensitivity Analysis

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Estimate | (Std. Error) | Estimate | (Std. Error) | Estimate | (Std. Error) |
| | *Heterogeneous* $\lambda$ | | *Heterogeneous* $\lambda$ | | *Heterogeneous* $\lambda$ | |
| *Gain-Loss Attitudes:* | | | | | | |
| $\hat{\lambda}$ | 1.29 | (0.04) | 1.37 | (0.08) | 1.64 | (0.21) |
| $\hat{\mu_\lambda}$ | 0.26 | (0.03) | 0.17 | (0.07) | 0.04 | (0.08) |
| $\hat{\sigma^2_\lambda}$ | 0.00 | (0.00) | 0.29 | (0.21) | 0.91 | (0.39) |
| *Pair 1 Utilities* (USB Stick (X) - Pen Set (Y)) *:* | | | | | | |
| $\hat{\frac{Y}{X}}$ (Initial) | 0.64 | (0.03) | 0.62 | (0.04) | 0.57 | (0.04) |
| $\hat{\frac{Y}{X}}$ (Replication) | 0.64 | (0.04) | 0.61 | (0.04) | 0.57 | (0.05) |
| *Pair 2 Utilities* (Picnic Mat (X) - Thermos (Y))*:* | | | | | | |
| $\hat{\frac{Y}{X}}$ (Initial) | 1.10 | (0.03) | 1.11 | (0.03) | 1.13 | (0.04) |
| $\hat{\frac{Y}{X}}$ (Replication) | 0.90 | (0.04) | 0.88 | (0.04) | 0.87 | (0.05) |
| *Discernibility:* | | | | | | |
| $\delta_X$ | 0.50 | - | 0.55 | - | 0.60 | - |
| # Observations | 3,072 | | 3,072 | | 3,072 | |

*Notes*: Maximum likelihood estimates. Standard errors in parentheses.

the expected value of $\lambda_i$ given a preference for $X$ when endowed with $X$ as

$$E[\hat{\lambda}_{i,X|X}] = \int \lambda \frac{Prob_{X|X}(\lambda)g(\lambda)}{\int Prob_{X|X}(\lambda)g(\lambda)d\lambda} d\lambda.$$

For each endowment, subjects could provide one of two hypothetical choice statements, one of three relative liking statements, and one of three relative wanting statements, yielding 18 potential statement profiles. With four endowments, there are 72 potential profiles, each with an implication for the expected value of $\lambda$.[33] We extend the above example to construct the probability of each such profile assuming independence between the simulated probabilities for hypothetical choice, liking, and wanting statements. We simulate statement profiles at 1 million draws from the estimated distribution of gain-loss attitudes assuming logit errors on choice probabilities. This exercise of mapping from preference statements to a conditional expectation of gain-loss attitudes takes into account the possibility of noise as the preference statements are simulated assuming logit errors.

In Table A7, we provide the expected value $E[\hat{\lambda}_i]$, averaged over the initial and replication study, for eight common statement profiles (accounting for 647 of 1024 (63.1 percent)

[33]Note that because we allow for different utilities in our initial study and replication, there are 72 such values for each.

of observations). Consider an endowment of the USB stick: if a subject stated a preference for the USB stick in all three statements they would have $E[\hat{\lambda}_i] = 1.87$, while if they stated a preference for the pen set in all three they would have $E[\hat{\lambda}_i] = 0.78$. Providing the same profiles when endowed with the pen set leads to $E[\hat{\lambda}_i]$ of 1.03 and 2.57, respectively. The values exhibited in Table A7 are intuitive: stating a preference for one's endowed object indicates loss aversion, while stating a preference for the alternative indicates gain lovingness. The magnitudes of $E[\hat{\lambda}_i]$ are tuned by the intrinsic values of the two objects reported in Table A5.

Table A7: Preference Statements and Individual Gain-Loss Classifications

| | | | |
|---|---|---|---|
| Endowed USB Stick | | | |
| HC(USB Stick) > HC(Pen Set)<br>L(USB Stick) > L(Pen Set)<br>W(USB Stick) > W(Pen Set) | $E[\hat{\lambda}_i] = 1.87$, (N=161) | HC(USB Stick) < HC(Pen Set)<br>L(USB Stick) < L(Pen Set)<br>W(USB Stick) < W(Pen Set) | $E[\hat{\lambda}_i] = 0.78$, (N = 32) |
| Endowed Pen Set | | | |
| HC(USB Stick) > HC(Pen Set)<br>L(USB Stick) > L(Pen Set)<br>W(USB Stick) > W(Pen Set) | $E[\hat{\lambda}_i] = 1.03$, (N=111) | HC(USB Stick) < HC(Pen Set)<br>L(USB Stick) < L(Pen Set)<br>W(USB Stick) < W(Pen Set) | $E[\hat{\lambda}_i] = 2.57$, (N= 57) |
| Endowed Picnic Mat | | | |
| HC(Mat) > HC(Thermos)<br>L(Mat) > L(Thermos)<br>W(Mat) > W(Thermos) | $E[\hat{\lambda}_i] = 2.26$, (N=84) | HC(Mat) < HC(Thermos)<br>L(Mat) < L(Thermos)<br>W(Mat) < W(Thermos) | $E[\hat{\lambda}_i] = 0.85$, (N= 67) |
| Endowed Thermos | | | |
| HC(Mat) > HC(Thermos)<br>L(Mat) > L(Thermos)<br>W(Mat) > W(Thermos) | $E[\hat{\lambda}_i] = 0.85$, (N=52) | HC(Mat) < HC(Thermos)<br>L(Mat) < L(Thermos)<br>W(Mat) < W(Thermos) | $E[\hat{\lambda}_i] = 2.18$, (N= 83) |

*Notes:* Implications for $E[\hat{\lambda}_i]$ for 8 key statement profiles, depending on endowment. HC: Hypothetic Choice; L: Liking Rating Score; W: Wanting Rating Score. $E[\hat{\lambda}_i]$ averaged over relevant observation number, $N$, between initial and replication study.

Figure A8 provides the distribution of $E[\hat{\lambda}_i]$ implied by Stage 1 preference statements as the solid black line. This distribution has mean 1.49, median 1.32, with 23 percent of subjects exhibiting $E[\hat{\lambda}_i] < 1$. The distribution of $E[\hat{\lambda}_i]$ is similar in shape and key statistics to the underlying log-normal estimates. However, the distribution of $E[\hat{\lambda}_i]$ does exhibit fewer extreme gain-seeking and loss-averse observations than its underlying distribution. Individual heterogeneity in $E[\hat{\lambda}_i]$ in hand, we are equipped to analyze heterogeneous treatment effects.

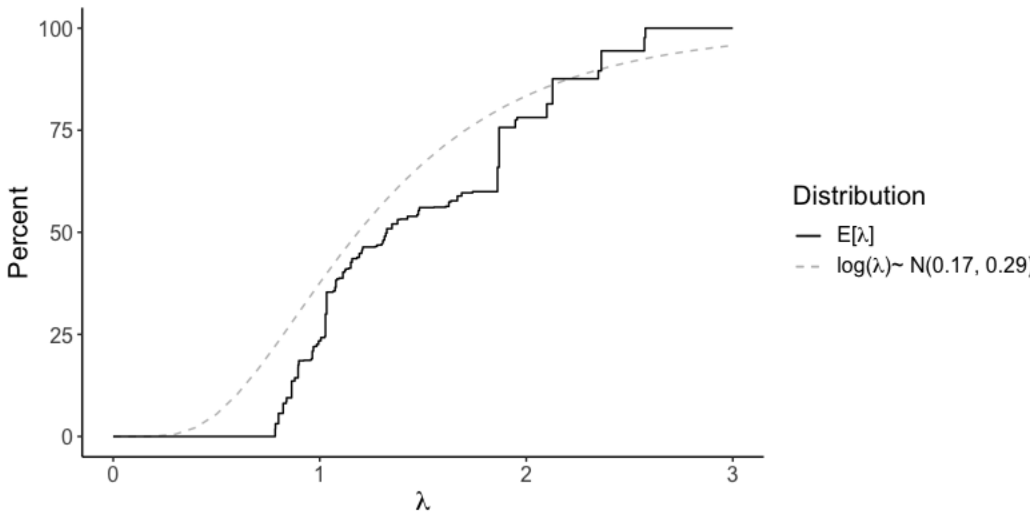



Figure A8: **Estimated and Calculated Distributions of Gain-Loss Attitudes**
*Notes:* The dashed line represents estimated distribution $log(\lambda) \sim N(0.17, 0.29)$. Solid line represents the expected value of $\lambda$ conditional on the Stage 1 statements, $E[\hat{\lambda}_i]$, as described above.

## B.3 Predicting Heterogeneous Treatment Effects in Exchange

Section 3.2.1 establishes the two critical CPE thresholds,

$$X_{L,i} = Y,$$

$$X_{H,i} = \frac{1 + 0.5(\lambda - 1)}{1 + 0.5(1 - \lambda_i)} Y.$$

Under deterministic choice, these CPE thresholds would map to choice probabilities 0 and 1 depending on the relative values of $X$ and $Y$ and the value of $\lambda_i$. In such an environment individual treatment effects on choice probabilities are either 1, 0, or -1 depending on the values of the these same parameters.

We do not assume deterministic choice, but rather stochastic choice. Hence, an individual will choose the alternative $Y$ over their endowed object $X$ in Condition Low, when

$$X + \varepsilon \leq Y$$

or

$$\epsilon \leq \frac{Y}{X} - 1$$

where $\epsilon = \frac{\varepsilon}{X}$ is a draw from mean zero distribution $F(\cdot)$. Hence, the probability of exchange in Condition Low is

$$Prob(Exchange_L) = Prob(\epsilon \leq \frac{Y}{X} - 1) = F(\frac{Y}{X} - 1).$$

Similarly, an individual will exchange in Condition High, when

$$X + \varepsilon \leq \frac{1 + 0.5(\lambda_i - 1)}{1 + 0.5(1 - \lambda_i)} Y,$$

and

$$Prob(Exchange_H) = F(\frac{1 + 0.5(\lambda_i - 1)}{1 + 0.5(1 - \lambda_i)} \frac{Y}{X} - 1).$$

This yields an individual treatment effect as a function of the parameters of interest,

$$\begin{aligned} TE(\lambda_i, X, Y) &= Prob(Exchange_H) - Prob(Exchange_L) \\ &= F(\frac{1 + 0.5(\lambda_i - 1)}{1 + 0.5(1 - \lambda_i)} \frac{Y}{X} - 1) - F(\frac{Y}{X} - 1). \end{aligned}$$

The analysis of Figure 4 in the main text presents predictions of $TE(\lambda_i, Y, X)$ for each individual at their value of $E[\hat{\lambda}_i]$ and at the estimated value of $\frac{Y}{X}$ for their assigned condition with $F(\cdot)$ assumed to be logistic.

## B.4 Non-Linear Aggregation of Exchange Treatment Effects and Statistical Power

Having established the theoretical treatment effect,

$$\begin{aligned} TE(\lambda_i, X, Y) &= Prob(Exchange_H) - Prob(Exchange_L) \\ &= F(\frac{1+0.5(\lambda_i-1)}{1+0.5(1-\lambda_i)}\frac{Y}{X}-1) - F(\frac{Y}{X}-1). \end{aligned}$$

we can consider aggregation of treatment effects in an average treatment effect,

$$\overline{TE(\lambda_i, X, Y)} = \frac{1}{N}\sum_{i=1}^{N} TE(\lambda_i, X, Y).$$

When will the average treatment effect deviate from the treatment effect of the average gain-loss attitude, $\overline{\lambda_i}$? Note that their are two dimensions of non-linearity in $\lambda_i$ that influence aggregation. First, the CPE threshold determining behavior in Condition High,

$$X_{H,i} = \frac{1+0.5(\lambda-1)}{1+0.5(1-\lambda_i)}Y,$$

is non-linear in $\lambda_i$. Second, given standard functional forms for $F(\cdot)$ like logistic or normal, the probability of exchange is plausibly non-linear in its arguments. Both of these forces will lead to deviations between the average treatment effect and the treatment effect of the average preference. Figure A9, plots $TE(\lambda_i, X, Y)$ with $F(\cdot)$ assumed to be logistic as above with various values for the relative utility $\frac{Y}{X}$.

The non-linear relationships illustrated in Figure A9 may lead average treatment effects to deviate dramatically from the treatment effect of the average preference. Overall the nature of the aggregation problem depends on the relative utility value, $\frac{Y}{X}$. When the alternative good is better than the endowment, $\frac{Y}{X} > 1$, gain-seeking individuals have more extreme negative treatment effects than loss-averse individuals. When $\frac{Y}{X} < 1$ the opposite is true. For $\frac{Y}{X} = 1$, both concave and convex regions of $TE(\lambda_i, X, Y)$ exist and the extent of aggregation problems depends importantly on the underlying distribution of $\lambda_i$. Even with loss aversion on average, the average treatment effect is plausibly muted relative to the treatment effect of the average preference. Given our distributional estimates for $\lambda$ noted in Table A5, and assuming $\frac{X}{Y} = 1$, the average treatment effect would be approximately 0.08 and the treatment effect of the average preference, $\lambda = 1.37$, would be approximately 0.11.

In addition to muted average treatment effects, heterogeneity in gain loss attitudes can influence the power of any conducted experimental test. Given our distributional estimates for $\lambda$ noted in Table A5, and assuming $\frac{X}{Y} = 1$, the average treatment effect would be 0.08 and the standard deviation of treatment effects would be 0.12. As noted above, the treatment effect of the average preference noted in Table A5 is 0.11. A study that is theoretically powered assuming homogeneous gain-loss attitudes and straightforward sampling variation will have different power considerations when accounting for this additional source of variation.

Absent heterogeneity in gain-loss attitudes, a treatment effect of 0.08 or 0.11 on exchange probability (assuming all subjects participate in *both* Low and High conditions, and Low condition exchange probability of 0.5) would be powered at 80% with approximately 600 or 320 subjects, respectively (one mean, standard deviation calculated as sum of independent binomial variances $\sqrt{p(1-p) + (p+TE)(1-(p+TE)))}$). This shows a first challenge to power associated with non-linear aggregation of treatment effects: an average treatment effect that is below the treatment effect of the average preference requires a larger sample to appropriately power. Absent heterogeneity, the standard deviation of a 0.08 treatment effect on exchange probability is $\sqrt{0.5(1-0.5) + 0.58(1-0.58))} \approx 0.7$. If heterogeneity were to be recognized, the expected standard deviation of treatment effects would grow. Assuming an independent effect of the heterogeneity described above would increase the standard deviation slightly to $\sqrt{0.7^2 + 0.12^2} \approx 0.71$, and the required sample size for 80% power would increase to approximately 620 subjects in a within-subject design. Hence, the combined effects of heterogeneity through non-linear aggregation and increased

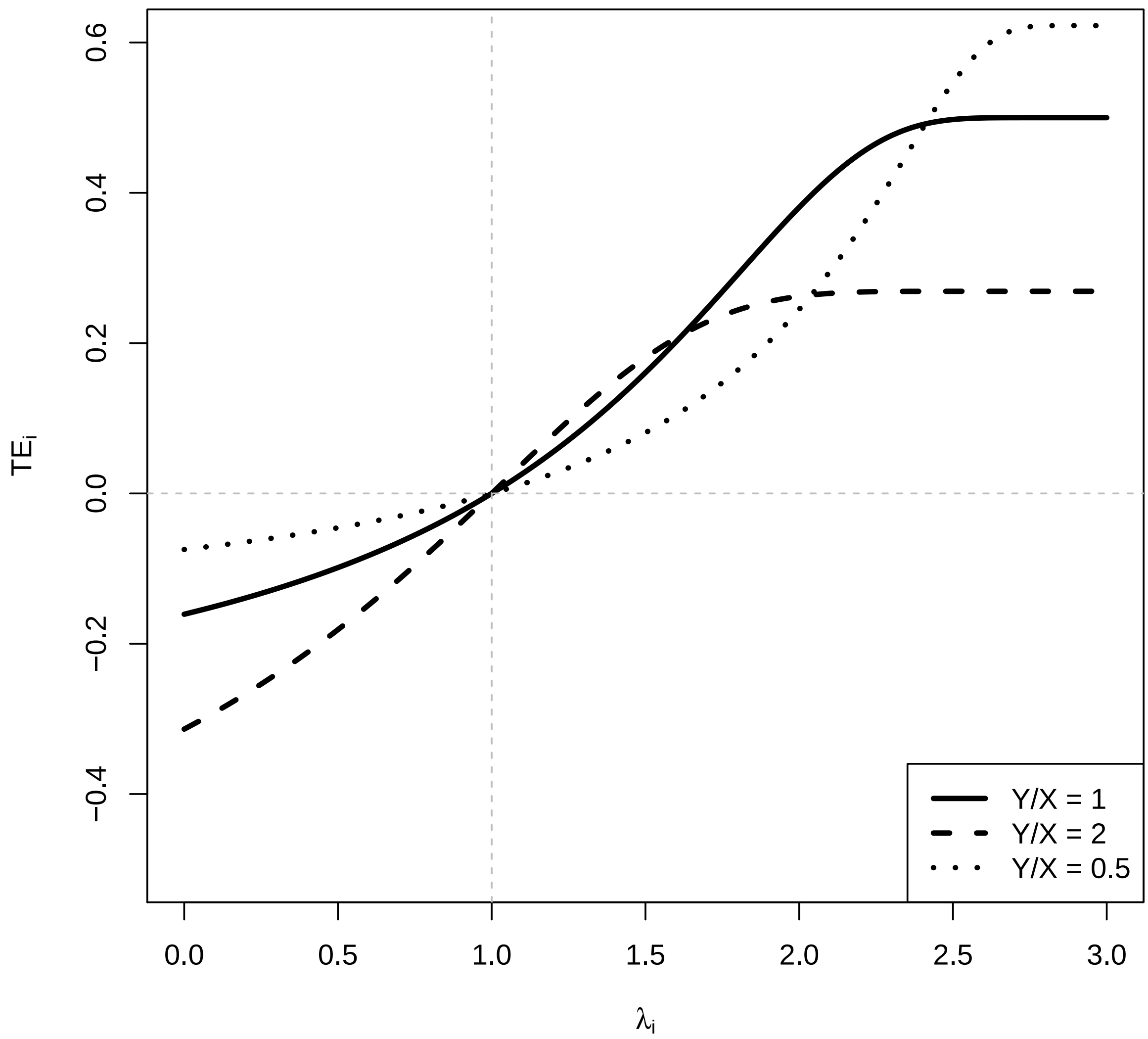



Figure A9: Gain-Loss Attitudes and Predicted Treatment Effects

*Notes:* Figure plots $TE(\lambda_i, X, Y)$ against $\lambda_i$ for various values of $\frac{X}{Y}$ and a logistic distribution of errors, independent and identically distributed in Condition Low and Condition High.

variability of treatment effects can lead to substantially different power calculations than those conducted assuming homogeneous preferences.

## B.5 Additional Results for Exchange Study

**Complementarities Between Stages.** Our results indicate that gain-loss attitudes measured with one pair of objects in Stage 1 are predictive of exchange behavior for a distinct counterbalanced pair of objects in Stage 2. Though we attempted to choose Stage 1 and Stage 2 objects that would have no plausible complementarities, if some un-modeled, unintentional complementarity did exist it might spuriously appear as predictive power across stages. For example, a subject might state a preference for or against both of their endowed objects in order to consume both endowed objects or both alternatives together. Note that this mechanism cannot explain the Stage 2 treatment effect, but could perhaps provide a rationale for the correlations documented between Stage 1 gain-loss attitudes and exchange in Stage 2, Condition Low.

Importantly, our Stage 1 design was constructed with one piece of random variation that serves to break complementarities between objects across stages. After providing their preference statements, half of subjects have their endowed object replaced with the alternative. If our results are reproduced both for subjects who have their endowed object replaced and those who do not, then explanations based upon accidental complementarities cannot be relevant for our results. To explore this possibility, Table A8 reproduces the structural results of Table 2 separately by individuals who do and do not have their Stage 1 endowed object replaced. For both groups, our results are maintained. Appendix Table A11 provides the same analysis with standard errors clustered at the session level and reaches the same statistical conclusions.

**Replication consistency.** Our results to here have combined the data from our initial and replication studies. Table A9 reproduces the structural results of Table 2 separately for the two samples, clustering standard errors at the session level. The null aggregate treat-

Table A8: Stage 2 Behavior and Stage 1 Experience

| | *Dependent Variable: Exchange (=1)* | | | | | |
|---|---|---|---|---|---|---|
| | Stage 1 Object Not Replaced | | | Stage 1 Object Replaced | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Condition High | 0.013 | 0.010 | -0.255 | -0.019 | -0.013 | -0.418 |
| | (0.044) | (0.044) | (0.126) | (0.043) | (0.043) | (0.122) |
| Reduced form ($\hat{l}_i$) | | -0.041 | | | -0.060 | |
| | | (0.022) | | | (0.022) | |
| Condition High * Reduced form ($\hat{l}_i$) | | 0.050 | | | 0.102 | |
| | | (0.029) | | | (0.028) | |
| $E[\hat{\lambda}_i]$ | | | -0.121 | | | -0.153 |
| | | | (0.057) | | | (0.058) |
| Condition High * $E[\hat{\lambda}_i]$ | | | 0.176 | | | 0.272 |
| | | | (0.077) | | | (0.077) |
| Constant (Condition Low) | 0.386 | 0.388 | 0.569 | 0.374 | 0.372 | 0.600 |
| | (0.033) | (0.033) | (0.094) | (0.034) | (0.022) | (0.095) |
| R-Squared | 0.000 | 0.008 | 0.011 | 0.000 | 0.025 | 0.024 |
| # Observations | 511 | 511 | 511 | 513 | 513 | 513 |
| $H_0$ : Zero Treatment Effect (H-L) | $F_{1,509} = .08$ | $F_{1,507} = .05$ | $F_{1,507} = 4.08$ | $F_{1,511} = 0.19$ | $F_{1,509} = 0.09$ | $F_{1,509} = 11.76$ |
| | $(p = 0.77)$ | $(p = 0.82)$ | $(p = 0.04)$ | $(p = .67)$ | $(p = 0.77)$ | $(p < 0.01)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Exchange in L | | $F_{1,507} = 3.68$ | $F_{1,507} = 4.52$ | | $F_{1,509} = 7.49$ | $F_{1,509} = 6.96$ |
| | | $(p = 0.06)$ | $(p = 0.03)$ | | $(p < 0.01)$ | $(p < 0.01)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,507} = 2.92$ | $F_{1,507} = 5.25$ | | $F_{1,509} = 13.13$ | $F_{1,509} = 12.53$ |
| | | $(p = 0.09)$ | $(p = 0.02)$ | | $(P < 0.01)$ | $(p < 0.01)$ |

*Notes*: Ordinary least square regression. Robust standard errors in parentheses. Null hypotheses tested for 1) zero baseline endowment effect regression (Constant coefficient = 0.5); 2) zero treatment effect (Condition High coefficient= 0); 3) no relationship between gain-loss attitudes and behavior in Condition Low behavior ($E[\hat{\lambda}_i]$ coefficient = 0); 4) constant treatment effect over gain-loss attitudes (Condition High * $E[\hat{\lambda}_i] = 0$). $F$-statistics and two-sided $p$-values reported.

ment effect and heterogeneous treatment effects over gain-loss attitudes are produced in both our initial and replication studies. Quantitatively the observed relationships between gain-loss attitudes and exchange behavior are broadly consistent, though the replication has less precise estimates due to the smaller sample size.

Table A9: Replication Consistency, Clustered SE

| | *Dependent Variable: Exchange (=1)* | | | | | |
|---|---|---|---|---|---|---|
| | | Initial Study | | | Replication Study | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Condition High | 0.004 | -0.001 | -0.409 | -0.010 | -0.007 | -0.239 |
| | (0.034) | (0.034) | (0.111) | (0.044) | (0.044) | (0.102) |
| Reduced form ($\hat{l}_i$) | | -0.064 | | | -0.034 | |
| | | (0.022) | | | (0.016) | |
| Condition High * Reduced form ($\hat{l}_i$) | | 0.10 | | | 0.046 | |
| | | (0.027) | | | (0.021) | |
| $E[\hat{\lambda}_i]$ | | | -0.159 | | | -0.103 |
| | | | (0.053) | | | (0.053) |
| Condition High * $E[\hat{\lambda}_i]$ | | | 0.266 | | | 0.161 |
| | | | (0.065) | | | (0.064) |
| Constant (Condition Low) | 0.365 | 0.373 | 0.616 | 0.399 | 0.394 | 0.542 |
| | (0.028) | (0.027) | (0.093) | (0.030) | (0.029) | (0.081) |
| R-Squared | 0.000 | 0.022 | 0.023 | 0.000 | 0.006 | 0.008 |
| # Observations | 607 | 607 | 607 | 417 | 417 | 417 |
| $H_0$ : Zero Treatment Effect (H-L) | $F_{1,30} = .01$ | $F_{1,30} = .0006$ | $F_{1,30} = 13.44$ | $F_{1,21} = 0.05$ | $F_{1,21} = 0.03$ | $F_{1,21} = 5.51$ |
| | $(p = 0.90)$ | $(p = 0.99)$ | $(p < 0.01)$ | $(p = .82)$ | $(p = 0.87)$ | $(p = 0.03)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Exchange in L | | $F_{1,30} = 8.15$ | $F_{1,30} = 9.09$ | | $F_{1,21} = 4.72$ | $F_{1,21} = 3.79$ |
| | | $(p < 0.01)$ | $(p < 0.01)$ | | $(p = 0.03)$ | $(p = 0.07)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,30} = 13.73$ | $F_{1,30} = 16.61$ | | $F_{1,21} = 4.87$ | $F_{1,21} = 6.32$ |
| | | $(p < 0.01)$ | $(p < 0.01)$ | | $(p = 0.03)$ | $(p < 0.01)$ |

*Notes*: Ordinary least square regression. Standard errors clustered at the session level in parentheses. Null hypotheses tested for 1) zero treatment effect (Condition High coefficient= 0); 2) no relationship between gain-loss attitudes and behavior in Condition Low behavior ($E[\hat{\lambda}_i] = 0$); 3) constant treatment effect over gain-loss attitudes (Condition High * $E[\hat{\lambda}_i] = 0$) $F$-statistics and two-sided $p$-values reported.

Our replication study was conducted to assure confidence in our previously obtained heterogeneous treatment effects. The registration of our pre-analysis plan, including power calculations, can be found at `https://www.socialscienceregistry.org/trials/3124`. The analysis proposed there carries one important difference to that conducted here: our proposed methodology for identifying gain-loss attitudes was based on standard logit, rather than mixed logit methods. This was the methodology used in our initial draft posted at `https://papers.ssrn.com/sol3/papers.cfm?abstractid=3170670`. Advice from an anonymous referee highlighted the value of the mixed logit methods that we currently conduct. For completeness, in Appendix B.6 we provide the pre-registered replication analysis.

There, as well, we find a striking consistency between the results obtained in our initial and replication samples.

Table A10: Exchange Behavior and Probabilistic Forced Exchange, Clustered SE

| | *Dependent Variable: Exchange (=1)* | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Condition High | -0.004 | -0.340 | -0.004 |
| | (0.027) | (0.076) | (0.026) |
| $E[\hat{\lambda}_i]$ | | -0.136 | |
| | | (0.036) | |
| Condition High * $E[\hat{\lambda}_i]$ | | 0.225 | |
| | | (0.046) | |
| Reduced form ($\hat{l}_i$) | | | -0.050 |
| | | | (0.014) |
| Condition High * Reduced form ($\hat{l}_i$) | | | 0.077 |
| | | | (0.018) |
| Constant (Condition Low) | 0.380 | 0.584 | 0.380 |
| | (0.020) | (0.061) | (0.019) |
| R-Squared | 0.000 | 0.017 | 0.014 |
| # Observations | 1024 | 1024 | 1024 |
| # Clusters | 53 | 53 | 53 |
| $H_0$ : Zero Endowment Effect in L | $F_{1,52} = 34.96$ | $F_{1,52} = 1.87$ | $F_{1,52} = 38.26$ |
| | $(p < 0.01)$ | $(p = 0.18)$ | $(p < 0.01)$ |
| $H_0$ : Zero Treatment Effect (H-L) | $F_{1,52} = .02$ | $F_{1,52} = 20.07$ | $F_{1,52} = 0.02$ |
| | $(p = 0.89)$ | $(p < 0.01)$ | $(p = 0.89)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Exchange in B | | $F_{1,52} = 13.98$ | $F_{1,52} = 13.19$ |
| | | $(p < 0.01)$ | $(p < 0.01)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,52} = 24.03$ | $F_{1,52} = 19.48$ |
| | | $(p < 0.01)$ | $(p < 0.01)$ |

*Notes*: Ordinary least square regression. Standard errors clustered at session level in parentheses. Null hypotheses tested for 1) zero baseline endowment effect regression (Constant coefficient = 0.5); 2) zero treatment effect (Condition High coefficient= 0); 3) no relationship between gain-loss attitudes and behavior in Condition Low behavior ($E[\hat{\lambda}_i]$ or $\hat{l}_i = 0$); 4) constant treatment effect over gain-loss attitudes (Condition High * $E[\hat{\lambda}_i]$ or Condition High * $\hat{l}_i = 0$). $F$-statistics and two-sided $p$-values reported.

Table A11: Stage 2 Behavior and Stage 1 Experience, Clustered SE

| | *Dependent Variable: Exchange (=1)* | | | |
|---|---|---|---|---|
| | Stage 1 Object Not Replaced | | Stage 1 Object Replaced | |
| | (1) | (2) | (3) | (4) |
| Condition High | 0.013 | -0.255 | -0.019 | -0.418 |
| | (0.035) | (0.120) | (0.044) | (0.124) |
| $E[\hat{\lambda}_i]$ | | -0.121 | | -0.153 |
| | | (0.053) | | (0.064) |
| Condition High * $E[\hat{\lambda}_i]$ | | 0.176 | | 0.272 |
| | | (0.071) | | (0.077) |
| Constant (Condition Low) | 0.386 | 0.569 | 0.374 | 0.600 |
| | (0.027) | (0.092) | (0.032) | (0.104) |
| R-Squared | 0.000 | 0.011 | 0.000 | 0.024 |
| # Observations | 511 | 511 | 513 | 513 |
| # Clusters | 53 | 53 | 53 | 53 |
| $H_0$ : Zero Endowment Effect in L | $F_{1,52} = 17.82$ | $F_{1,52} = 0.57$ | $F_{1,52} = 15.78$ | $F_{1,52} = 0.92$ |
| | ($p < 0.01$) | ($p = 0.45$) | ($p < 0.01$) | ($p = 0.34$) |
| $H_0$ : Zero Treatment Effect (F-B) | $F_{1,52} = 0.13$ | $F_{1,52} = 4.51$ | $F_{1,52} = 0.18$ | $F_{1,52} = 11.31$ |
| | ($p = 0.72$) | ($p = 0.04$) | ($p = 0.67$) | ($p < 0.01$) |
| $H_0$ : Gain-Loss Attitudes $\perp$ Exchange in B | | $F_{1,52} = 5.25$ | | $F_{1,52} = 5.81$ |
| | | ($p = 0.03$) | | ($p = 0.02$) |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,52} = 6.19$ | | $F_{1,52} = 12.62$ |
| | | ($p = 0.02$) | | ($p < 0.01$) |

*Notes*: Ordinary least square regression. Standard errors clustered at session level in parentheses. Null hypotheses tested for 1) zero baseline endowment effect regression (Constant coefficient = 0.5); 2) zero treatment effect (Condition High coefficient= 0); 3) no relationship between gain-loss attitudes and behavior in Condition Low behavior ($E[\hat{\lambda}_i]$ coefficient = 0); 4) constant treatment effect over gain-loss attitudes (Condition High * $E[\hat{\lambda}_i] = 0$). $F$-statistics and two-sided $p$-values reported.

Table A12: Replication Consistency, Clustered SE

| | *Dependent Variable: Exchange (=1)* | | | | |
|---|---|---|---|---|---|
| | Initial Study | | Replication Study | | |
| | (1) | (2) | (3) | (4) | (5) |
| Condition High | 0.004 | -0.409 | -0.010 | -0.239 | -0.805 |
| | (0.034) | (0.111) | (0.044) | (0.102) | (0.411) |
| $E[\hat{\lambda}_i]$ | | -0.159 | | -0.103 | -0.116 |
| | | (0.053) | | (0.053) | (0.053) |
| Condition High * $E[\hat{\lambda}_i]$ | | 0.266 | | 0.161 | 0.174 |
| | | (0.065) | | (0.064) | (0.064) |
| Constant (Condition Low) | 0.365 | 0.616 | 0.399 | 0.542 | 0.917 |
| | (0.028) | (0.093) | (0.030) | (0.081) | (0.343) |
| Additional Controls | No | No | No | No | Yes |
| Additional Interactions | No | No | No | No | Yes |
| R-Squared | 0.000 | 0.023 | 0.000 | 0.008 | 0.060 |
| # Observations | 607 | 607 | 417 | 417 | 417 |
| # Clusters | 31 | 31 | 22 | 22 | 22 |
| $H_0$ : Zero Endowment Effect in B | $F_{1,30} = 23.85$ | $F_{1,30} = 1.53$ | $F_{1,21} = 11.73$ | $F_{1,21} = 0.26$ | $F_{1,21} = 1.48$ |
| | $(p < 0.01)$ | $(p = 0.23)$ | $(p < 0.01)$ | $(p = 0.61)$ | $(p = 0.24)$ |
| $H_0$ : Zero Treatment Effect (F-B) | $F_{1,30} = 0.01$ | $F_{1,30} = 13.44$ | $F_{1,21} = 0.05$ | $F_{1,21} = 5.51$ | $F_{1,21} = 3.84$ |
| | $(p = 0.90)$ | $(p < 0.01)$ | $(p = 0.82)$ | $(p = 0.03)$ | $(p = 0.06)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Exchange in B | | $F_{1,30} = 9.09$ | | $F_{1,21} = 3.79$ | $F_{1,21} = 4.78$ |
| | | $(p < 0.01)$ | | $(p = 0.07)$ | $(p = 0.04)$ |
| $H_0$ : Gain-Loss Attitudes $\perp$ Treatment Effect | | $F_{1,30} = 16.61$ | | $F_{1,21} = 6.32$ | $F_{1,21} = 7.47$ |
| | | $(p < 0.01)$ | | $(p < 0.01)$ | $(p = 0.01)$ |

*Notes*: Ordinary least square regression. Standard errors clustered at session level in parentheses. Null hypotheses tested for 1) zero baseline endowment effect regression (Constant coefficient = 0.5); 2) zero treatment effect (Condition High coefficient= 0); 3) no relationship between gain-loss attitudes and behavior in Condition Low behavior ($E[\hat{\lambda}_i] = 0$); 4) constant treatment effect over gain-loss attitudes (Condition High * $E[\hat{\lambda}_i] = 0$).

## B.6 Replication Exchange Study and Reconciliation with Pre-Analysis Plan

In this section we report the methodology and corresponding analyses from earlier versions of this paper (`https://papers.ssrn.com/sol3/papers.cfm?abstractid=3170670` and `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3589906`) as specified in the pre-registration plan of our replication study (`https://www.socialscienceregistry.org/trials/3124`). The key difference is that while our approach in the present version of the paper relies on a mixed-logit methodology following a suggestion of an anonymous referee, our previous approach employed standard logit methods. All our previous results are closely in line with those obtained using the new methodology. Here we provide a summary of the central exercises conducted in prior versions of the manuscript. For the complete analysis please see `https://papers.ssrn.com/sol3/papers.cfm?abstractid=3170670` and `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3589906`.

### B.6.1 Stage 1: Identifying Gain-Loss Attitudes

Our previous methodology relied on the same preference statements that we introduced in Section 3.1, but focused only on the liking preference statements. Instead of residualizing the first principal component of the preference statements as described in Section 3.2.1, in our previous analyses we constructed a simple structural model of the liking preference statement based upon standard random utility methods (McFadden, 1974) with the objective of capturing the source of both of these features: gain-loss attitudes and differences in intrinsic utility for the two objects.

Consider an individual endowed with $X$ that is asked to provide ratings statements for both $X$ and $Y$. Under the KR model, an individual evaluates their endowment, $X$, based upon $U(X,0|X,0)$. Given that the agent is endowed with $X$ and is uninformed of the possibility of confiscation at the time of the ratings, they plausibly evaluate $Y$ based upon $U(0,Y|X,0)$. With standard logit shocks, $\epsilon_X$ and $\epsilon_Y$, the parameters associated with these

KR utilities are easily estimated. We assume subjects will provide a higher rating for their endowed object, $X$, if

$$U(X,0|X,0) + \epsilon_X > U(0,Y|X,0) + \epsilon_Y + \delta,$$

where $\delta$ is a discernibility parameter which accounts for the fact that the goods may be given identical ratings (for use of such methods, see, e.g., Cantillo et al., 2010). Similarly, subjects provide a higher rating for the alternative object, $Y$, if

$$U(0,Y|X,0) + \epsilon_Y > U(X,0|X,0) + \epsilon_X + \delta,$$

and provide the same rating if the difference in utilities falls within the range of discernibility,

$$|U(X,0|X,0) + \epsilon_X - (U(0,Y|X,0) + \epsilon_Y)| \leq \delta.$$

Under the functional form assumptions of Section 2 with $\eta = 1$, for someone endowed with object $X$, we obtain familiar logit probabilities for the ranking of ratings $R(X)$ and $R(Y)$,

$$\begin{aligned} P(R(X) > R(Y)) &= \frac{\exp(U(X,0|X,0))}{\exp(U(X,0|X,0)) + \exp(U(0,Y|X,0)+\delta)} = \frac{\exp(X)}{\exp(X) + \exp(2Y - \lambda X + \delta)} \\ P(R(Y) > R(X)) &= \frac{\exp(U(0,Y|X,0))}{\exp(U(0,Y|X,0)) + \exp(U(X,0|X,0)+\delta)} = \frac{\exp(2Y - \lambda X)}{\exp(X+\delta) + \exp(2Y - \lambda X)} \\ P(R(X) = R(Y)) &= 1 - P(R(X) > R(Y)) - P(R(Y) > R(X)), \end{aligned}$$

where the intrinsic utility values, $X$ and $Y$, the discernibility parameter $\delta$, and the gain-loss parameter, $\lambda$, are the desired estimands.[34] We normalize one of the good's values to be $Y = 1$, and estimate the remaining parameters via maximum likelihood.

Table A13 provides aggregate estimates of intrinsic utilities, $\lambda$ and $\delta$, separately for each pair of goods in both the initial study and our replication. In each case we find aggregate support for loss aversion, $\lambda > 1$, though less pronounced in our replication study.

Table A13: Prior Analysis: Aggregate Parameter Estimates

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8 ) |
|---|---|---|---|---|---|---|---|---|
| | *Initial Study* | | | | *Replication Study* | | | |
| | Est. | (Std. Err.) | Est. | (Std. Err.) | Est. | (Std. Err.) | Est. | (Std. Err.) |
| | *Pair 1* | | *Pair 2* | | *Pair 1* | | *Pair 2* | |
| *Gain-Loss Attitudes:* | | | | | | | | |
| $\hat{\lambda}$ | 1.56 | (0.14) | 1.29 | (0.12) | 1.18 | (0.15) | 1.12 | (0.13) |
| *Utility Values:* | | | | | | | | |
| $\hat{X}_1$ *(Pen Set)* | 0.63 | (0.05) | | | 0.66 | (0.06) | | |
| $\hat{Y}_1$ *(USB Stick)* | 1 | - | | | 1 | - | | |
| $\hat{X}_2$ *(Picnic Mat)* | | | 0.84 | (0.05) | | | 1.05 | (0.07) |
| $\hat{Y}_2$ *(Thermos)* | | | 1 | - | | | 1 | - |
| *Discernibility:* | | | | | | | | |
| $\hat{\delta}$ | 0.55 | (0.06) | 0.45 | (0.05) | 0.45 | (0.06) | 0.62 | (0.07) |

*Notes*: Maximum likelihood estimates. Robust standard errors in parentheses.

[34] For someone endowed with the alternative object, $Y$, these same probabilities are

$$\begin{aligned}
P(R(X) > R(Y)) &= \frac{\exp(U(X,0|0,Y))}{\exp(U(X,0|0,Y)) + \exp(U(0,Y|0,Y) + \delta)} = \frac{\exp(2X - \lambda Y)}{\exp(Y + \delta) + \exp(2X - \lambda Y)} \\
P(R(Y) > R(X)) &= \frac{\exp(u(0,Y|0,Y))}{\exp(U(0,Y|0,Y)) + \exp(U(X,0|0,Y) + \delta)} = \frac{\exp(Y)}{\exp(Y) + \exp(2X - \lambda Y + \delta)} \\
P(R(X) = R(Y)) &= 1 - P(R(X) > R(Y)) - P(R(Y) > R(X)).
\end{aligned}$$

### B.6.2 Stage 1: Individual Gain-Loss Attitudes

The aggregate estimates show evidence of loss aversion. To construct bounds for estimates of individual gain-loss attitudes, we evaluate individual choices assuming average utility and discernibility values. For example, consider an individual endowed with the pen set in Pair 1 in the initial study. At the aggregate estimates of $\delta$ and $X$ for Pair 1, if this individual were to state a higher rating for the pen set than for the USB stick, it would imply $0.632 > 2 - \hat{\lambda} * 0.632 + 0.549$ or $\hat{\lambda} > 3.03$. Similarly, stating a higher rating for the USB stick would imply $\hat{\lambda} < 1.30$,[35] and stating the same rating implies $\hat{\lambda} \in [1.30, 3.03]$. Of these three possible cases, two demonstrate evidence of loss aversion $\hat{\lambda} > 1$, while the other case is plausibly loss neutral as $\hat{\lambda} = 1$ can rationalize the ratings.[36] In total, there exist twelve cases of endowments and relative liking statements.

Overall, in our initial study 217 subjects (35.7 percent) are categorized as loss-averse, 240 (39.5 percent) are categorized as potentially loss-neutral, and 150 (24.7 percent) are categorized as gain-seeking. In our replication study, 124 subjects (29.7 percent) are categorized as loss-averse, 185 (44.4 percent) are categorized as potentially loss-neutral, and 108 (25.9 percent) are categorized as gain-seeking. These are the taxonomies of individual gain-loss types used in our previous analysis.

### B.6.3 Stage 2: Heterogeneous Treatment Effects

Table A14 presents linear probability models for Stage 2 behavior with dependent variable *Exchange (=1)*. Panels A and B provide separate results for our initial and replication studies. All of these results leverage our initial methodology described above that only relies on the liking preference statements. Beginning with the initial study, we find a null average treatment effect in Column (1). In Condition Low, 36.5 percent of subjects choose

[35] To state a higher rating for the USB implies $2 - \hat{\lambda} * 0.632 > 0.632 + 0.549$ or $\hat{\lambda} < 1.30$.

[36] It may seem prima-facie surprising that providing the same rating in this case is consistent with loss aversion. The logic is simple: given that the pen set has substantially lower intrinsic utility than the USB stick, one must be loss-averse to rate them equally.

to exchange, demonstrating a significant endowment effect relative to the null hypothesis of 50 percent exchange, $F_{1,605} = 18.32,\ (p < 0.01)$. Probabilistic forced exchange in Condition High has a null average treatment effect, increasing trading probabilities by only 0.4 percentage points on aggregate. Columns (2) through (4) conduct the same regressions separately for subjects categorized as loss-averse, loss-neutral, and gain-seeking, based on their Stage 1 liking statements. Panel A of Table A14 shows a dramatic heterogeneous treatment effect. Loss-averse subjects exhibit a statistically significant endowment effect in Condition Low, and grow more approximately 16 percentage points more willing to exchange in Condition High. Gain-seeking subjects exhibit no endowment effect in Condition Low, and grow approximately 25 percentage points less willing to exchange in Condition High. The heterogeneous treatment effect over gain-seeking and loss-averse subjects of roughly 40 percentage points closely follows our theoretical development on the sign of comparative statics, and is significant at all conventional levels, $F_{1,363} = 15.76,\ (p < 0.01)$.

As detailed in the main text, we registered and conducted an exact replication in the summer of 2018 with 417 subjects, again at the University of Bonn. The registration of our pre analysis plan, including power calculations, can be found at `https://www.socialscienceregistry.org/trials/3124`. The number of subjects for the replication was guided by a requirement of 80 percent power for the 40 percentage point difference in treatment effect between gain-seeking and loss-averse subjects noted above. Ex-post, our initial study turned out to be slightly over-powered and the replication was thus conducted with around 400 subjects. Panel B of Table A14 provides the replication analysis analogous to that presented in Panel B. The null average treatment effect, positive treatment effect for loss-averse subjects, and negative treatment effect for gain-seeking subjects are all reproduced with accuracy. Indeed, the 40 percentage point heterogeneous treatment effect in our initial study is echoed in a 37 percentage point difference between gain-seeking and loss-averse subjects in our replication study.

Our replication study reproduces with precision the heterogeneous treatment effect over gain-loss types obtained in our initial study under our prior methods. Subjects classified as

Table A14: Prior Analysis: Exchange Behavior and Probabilistic Forced Exchange

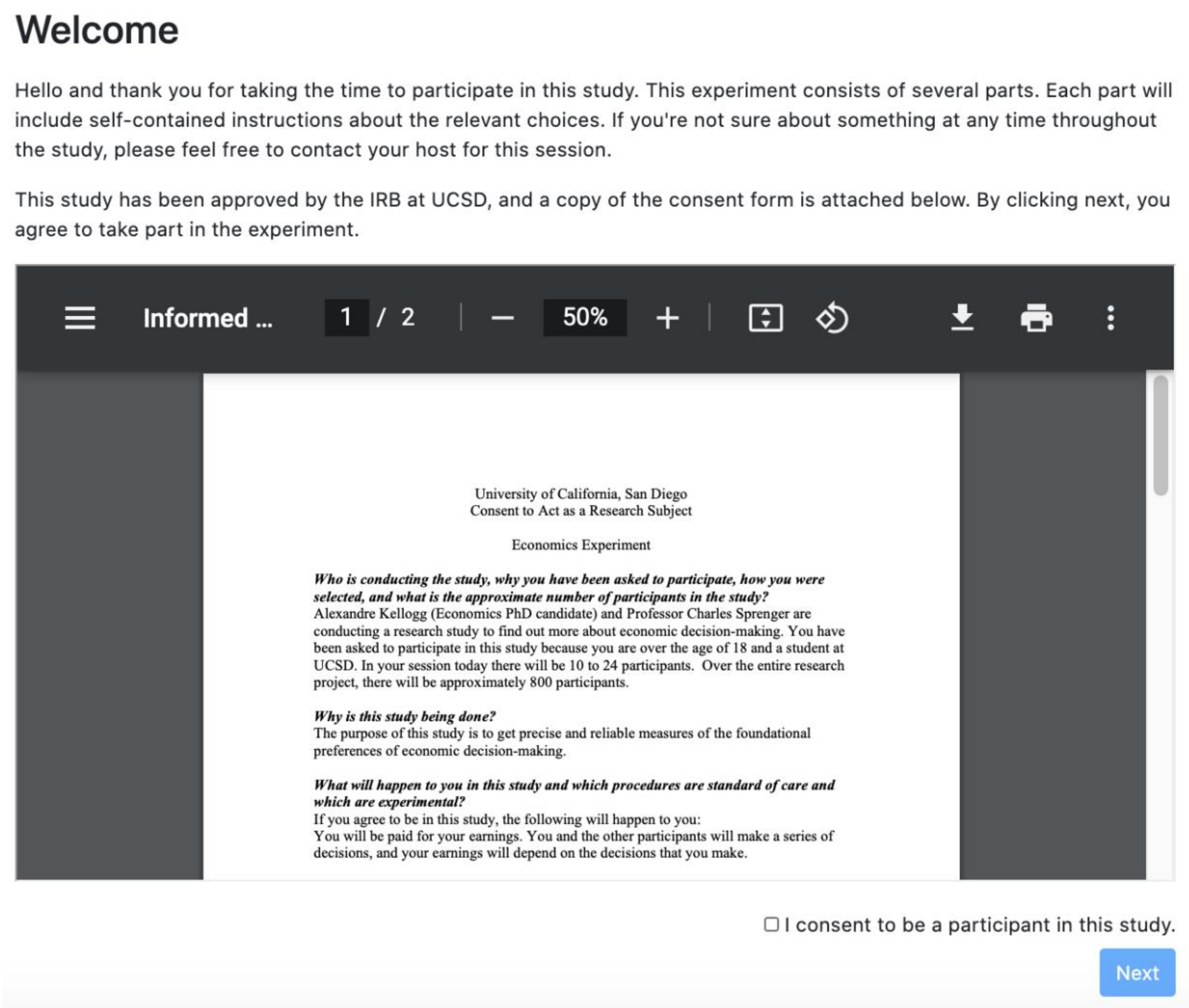
| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | *Dependent Variable: Exchange (=1)* | | | |
| | Full Sample | Loss Averse | Loss Neutral | Gain Seeking |
| *Panel A: Initial Study* | | | | |
| Condition High | 0.004 | 0.158 | 0.027 | -0.248 |
| | (0.034) | (0.067) | (0.066) | (0.078) |
| Constant (Condition Low) | 0.365 | 0.330 | 0.361 | 0.429 |
| | (0.028) | (0.049) | (0.053) | (0.067) |
| R-Squared | 0.000 | 0.025 | 0.001 | 0.072 |
| # Observations | 607 | 217 | 240 | 150 |
| $H_0$: Zero Endowment Effect in B | $F_{1,605}$=18.32 | $F_{1,215}$=12.21 | $F_{1,238}$=6.85 | $F_{1,148}$=1.15 |
| | ($p < 0.01$) | ($p < 0.01$) | ($p < 0.01$) | ($p = 0.29$) |
| $H_0$: Zero Treatment Effect (F-B) | $F_{1,605} = 0.01$ | $F_{1,215} = 5.64$ | $F_{1,238} = 0.17$ | $F_{1,148} = 10.18$ |
| | ($p = 0.90$) | ($p = 0.02$) | ($p = 0.68$) | ($p < 0.01$) |
| $H_0$: Constant (col. 2) = Constant (col. 4) | | | | $F_{1,363} = 1.44$ |
| | | | | ($p = 0.23$) |
| $H_0$: Condition High (col. 2) =Condition Low (col. 4) | | | | $F_{1,363} = 15.76$ |
| | | | | ($p < 0.01$) |
| *Panel B: Replication Study* | | | | |
| Condition High | -0.010 | 0.206 | -0.073 | -0.160 |
| | (0.044) | (0.085) | (0.075) | (0.094) |
| Constant (Condition Low) | 0.399 | 0.271 | 0.444 | 0.474 |
| | (0.030) | (0.058) | (0.059) | (0.067) |
| R-Squared | 0.000 | 0.045 | 0.005 | 0.027 |
| # Observations | 417 | 124 | 185 | 108 |
| $H_0$: Zero Endowment Effect in B | $F_{1,415}$=7.97 | $F_{1,122}$=15.40 | $F_{1,183}$=0.89 | $F_{1,106}$=0.16 |
| | ($p < 0.01$) | ($p < 0.01$) | ($p = 0.35$) | ($p = 0.69$) |
| $H_0$: Zero Treatment Effect (F-B) | $F_{1,415} = 0.05$ | $F_{1,122} = 5.79$ | $F_{1,183} = 0.95$ | $F_{1,106} = 2.92$ |
| | ($p = 0.83$) | ($p = 0.02$) | ($p = 0.33$) | ($p = 0.09$) |
| $H_0$: Constant (col. 2) = Constant (col. 4) | | | | $F_{1,228} = 5.22$ |
| | | | | ($p = 0.02$) |
| $H_0$: Condition High (col. 2) =Condition Low (col. 4) | | | | $F_{1,228} = 8.33$ |
| | | | | ($p < 0.01$) |

*Notes*: Ordinary least square regression. Robust standard errors in parentheses. Null hypotheses tested for 1) zero endowment effect in Condition Low, regression (Constant = 0.5); 2) zero treatment effect (F-B); 3) Identical Condition Low behavior across loss-averse and gain-seeking subjects (Constant (col. 2) = Constant (col. 4)); 4) Identical treatment effects of forced exchange across loss-averse and gain-seeking subjects (High condition (col. 2) = High condition (col. 4)). Hypotheses 3 and 4 tested via interacted regression with observations from columns (2) and (4).

loss-averse respond to Condition High by increasing their willingness to exchange; subjects classified as gain-seeking respond by decreasing their willingness to exchange.

# Appendix C Instructions - Labor supply experiment

The following set of screenshots demonstrates a demo version of our experiment, designed on oTree (Chen et al., 2016).



# Welcome

Hello and thank you for taking the time to participate in this study. This experiment consists of several parts. Each part will include self-contained instructions about the relevant choices. If you're not sure about something at any time throughout the study, please feel free to contact your host for this session.

This study has been approved by the IRB at UCSD, and a copy of the consent form is attached below. By clicking next, you agree to take part in the experiment.

Informed ... 1 / 2 50%

University of California, San Diego
Consent to Act as a Research Subject

Economics Experiment

***Who is conducting the study, why you have been asked to participate, how you were selected, and what is the approximate number of participants in the study?***
Alexandre Kellogg (Economics PhD candidate) and Professor Charles Sprenger are conducting a research study to find out more about economic decision-making. You have been asked to participate in this study because you are over the age of 18 and a student at UCSD. In your session today there will be 10 to 24 participants. Over the entire research project, there will be approximately 800 participants.

***Why is this study being done?***
The purpose of this study is to get precise and reliable measures of the foundational preferences of economic decision-making.

***What will happen to you in this study and which procedures are standard of care and which are experimental?***
If you agree to be in this study, the following will happen to you:
You will be paid for your earnings. You and the other participants will make a series of decisions, and your earnings will depend on the decisions that you make.

☐ I consent to be a participant in this study.

Next

## Receiving Payment and Your Identity

All payments will be sent to you via Venmo/Zelle. In order to receive payment, we will need to collect an email address linked to this form of payment. This information will only be seen by the PIs in this study. As soon as your payments are made, the link between the choices you made and your payment will be destroyed, and the record with your email address will be deleted. Your identity will not be a part of the subsequent data analysis.

Please enter the email associated with your Zelle account:

Next

## Experiment Overview

In the main part of this experiment, you will have the possibility to earn money by completing a number of tasks. Each task consists of transcribing a line of blurry Greek letters from a Greek text. The experiment is divided into five parts, which we will explain in turn.

### Part 1

In **Part 1**, you will make 32 decisions. In each decision, you will be offered a rate for each task that you complete, and you will be asked to decide how many tasks you want to complete at that rate. For example, in one of the decisions you could be offered $0.20 for every task that you complete, and you will have to decide how many tasks (from 0 to 100) you want to complete at that rate.

Before you make any decisions about the number of tasks you wish to complete for each rate, you will be asked to complete 2 practice tasks. This will allow you to become acquainted with the task and give you a sense of how long a task takes you to complete.

**Part 1** takes approximately 20 minutes.

### Part 2

In **Part 2**, you will make 42 decisions. In each decision, you will be presented with two options: Option A will be a lottery, paying (for example) $10 with 20% chance and $0 with 80% chance and Option B will be a sure amount of $5. For each choice, the probabilities in Option A will vary, and you will be asked to indicate which option you prefer.

**Part 2** takes approximately 10 minutes.

### Part 3

After you have made your decisions in **Part 1** and **Part 2**, a computer will randomly determine which decision from **Part 1** or **Part 2** is chosen to be the *decision-that-counts*. The *decision-that-counts* will determine which of your prior choices will be implemented, and will thus determine part of your earnings for this study. Each of the choices you make are equally likely to be selected.

Next, everyone will be required to complete 10 tasks. Completing these mandatory tasks is required in order to earn your completion fee of $7.00 as well as the earnings determined from the *decision-that-counts*.

**Part 3** can take anywhere from 5 minutes to 20 minutes, depending on how long it takes you to complete a task. On average, it takes about 42 seconds to complete one task.

### Part 4

Once you have completed the 10 tasks, you may be asked to complete additional tasks as determined by your prior choices from the *decision-that-counts*.

**Part 4** may take anywhere from 1 minute to 2 hours depending on which decision was selected as the *decision-that-counts*, the number of tasks that you selected, and the time it takes you to complete each task.

### Part 5

Once you have finished with **Part 4**, you will then be asked to solve a few puzzles and answer a few demographic questions. There are a total of 5 puzzles to solve within 10 minutes. For each correct answer you submit, you will receive an additional $1.

After filling out your responses, you will receive your final payment via the account information you provided in the previous page. Your final payment will consist of a $7.00 completion fee + your earnings from completing each of the parts. If you do not complete all of the tasks you had previously chosen during **Part 4**, you will not receive the completion fee nor payment for any of the tasks, and will instead receive a show-up fee of $5.00.

### Summary

To review, this experiment consists of 5 distinct parts. In **Part 1** and **Part 2**, you will make a series of choices, each of which is equally likely to determine your payment. In **Part 3**, the *decision-that-counts* will be revealed and you will be asked to complete the required 10 tasks. In **Part 4**, you will be asked to complete any additional tasks determined by your prior choices (if relevant). Finally, **Part 5** has a brief set of puzzles and survey questions prior to concluding the experiment.

More detailed instructions will be presented prior to each part. If you have questions or want clarification, remember that you can always contact your host for this session. On the next page, you will learn more about the Greek transcription tasks you will be asked to complete.

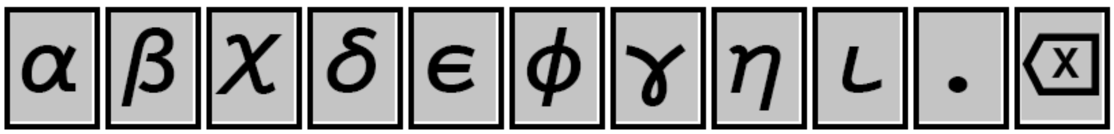
Next

# The Task

To complete a task, you will have to transcribe a line of blurry letters from a Greek text. For each task, Greek text will appear on your screen. You will be asked to transcribe these letters by finding and clicking on the corresponding letter, which will insert that letter into the completion box. If you would like to delete the most recently added character, please click on the backspace image. One task is one row of Greek text. For the task to be complete, your transcription must be 80% accurate or better. The following is an example of a row of blurry Greek text and the row of letters you will be asked to select from, along with its solution:

δηιφα.βαιδφι.ηβεχφηφιεδ.γηειφφγφαγ.

Please select from the following characters to enter your transcription.



The correct transcription for the example task is provided above; recall, however, that you do not need to be 100% accurate. If you submit a transcription that is 80% accurate or better (defined as requiring 7 or fewer insertions, deletions, or substitutions to achieve the perfect transcription), you will have completed the task. If your answer is incorrect (less than 80% accurate), you will have 2 more tries (for a total of 3) to submit a correct answer, after which you will be presented with a new task and the incorrect submissions will not be counted as a completed task.

As part of the task, several auditory "beeps" will sound randomly while you are completing the tasks. Please put on your headphones and/or turn your volume up so that you can hear the beeping noises. After each time you hear this beeping noise, you must press the “Noise” button at the bottom left of the screen. If you do not press the “Noise” button within five seconds of hearing the beeping noise, your transcription will be reset. If you press the noise button erroneously (when there was no beeping sound), your transcription will be reset. Note that resetting the transcription does not count as a try; your current progress will simply be deleted and you will have to re-enter the transcription.

The time it takes to complete a task will vary from person to person. **On average, however, each task takes about 42 seconds to complete.**

Before you make your decisions, we will present you with 1 tasks so that you better understand what it means to complete a task as you make your choices. This will also provide you a chance to ensure that you can hear the beeping noises correctly. If you have any trouble with the task or any questions about it, please contact your host!

Next

## Sample of Task

Please transcribe the row of Greek letters by selecting the appropriate letters. Press **Next** when you wish to submit your response. Remember to press the **Noise** button within 5 seconds of hearing the beeps, otherwise your responses will be removed and you will have to start over. You have 3 chances to complete each task. If you fail to complete the task within three tries you will simply be shown a different one.

### Number completed 0/1. Attempt: 1 of 3.

Please select from the following characters to enter your transcription.

| α | β | χ | δ | ϵ | ϕ | γ | η | ι | . | ⌫ |
|---|---|---|---|---|---|---|---|---|---|---|

Noise

Next

# Part 1

Throughout the following screens, you will make a total of 32 decisions. We will begin by explaining the kind of decisions that you will make for the first 30 decisions of the study. After you make these 30 decisions, you will receive a new set of instructions regarding the last 2 decisions.

## Decisions 1 to 30

In the next six screens, you will have to decide how many tasks you are willing to complete for a given rate. *As a reminder, one task means a correct transcription of a blurry line of Greek text*. The rates will be presented in lists of 5 at a time, and all rates within a list will either be *deterministic*, for example $0.15/task, or *stochastic*, for example a 50% chance of $0.10/task and a 50% chance of $0.20/task. The rates per task will range from $0.00 to $0.60 per task.

An example of your choice environment is provided below.

| Low Wage (50%) | High Wage(50%) | | Chosen Tasks | |
|---|---|---|---|---|
| $0.0/task | $0.1/task | (50% Chance of $0.00<br>50% Chance of $2.20) | 22 tasks<br>(~16mins) | |

Each rate will have a corresponding slider where you can choose, for that rate, how many tasks you are willing to complete. As you move the slider, you will see a subtotal next to the rate, as well as the estimated time to complete the number of Greek tasks indicated. This time is estimated based on the average of 42 seconds per task at the bottom of the page, but you may enter your own estimated time given what you learned in the practice tasks.

Recall that each of the decisions that you will make throughout **Part 1** and **Part 2** of this study is equally likely to be the decision that counts. **Thus, it is very important that you think carefully about each decision you make, as it could be the one selected for payment.** If one of these 30 decisions is randomly selected to be the *decision-that-counts* , you will be asked to complete the number of Greek tasks that you indicated and you will be compensated at the rate specified by the *decision-that-counts*. Recall that everyone will be required to complete their 10 tasks before continuing on to the number of tasks that you indicated in the *decision-that-counts*.

If the rate for the *decision-that-counts* is deterministic, say $0.15/task, and you said you would work 50 tasks at that rate, you will be paid a $7.00 completion fee + $7.50 for your tasks for a total of $14.50 (plus an additional $1 for each puzzle you correctly solve in **Part 5**).

If the *decision-that-counts* involves a stochastic rate, say $0.10/task with 50% chance and $0.20/task with 50% chance, and you chose to work 50 tasks, then you will be asked to complete the 50 tasks after the mandatory 10. After you have completed all of these tasks, the computer will reveal which of these two rates applies by flipping a coin. Once the rate is determined, say the computer selects $0.20/task ($0.10/task), you will be paid a total of $17.00 ($12.00), $7.00 for the completion fee + $10.00 ($5.00) for your 50 tasks (plus an additional $1 for each puzzle you correctly solve in **Part 5**).

Over the next six pages, you will be presented with a series of 5 wages per page and asked to indicate the amount of tasks you wish to complete at the given rates.

Next

## Effort Choices

Please indicate the number of tasks you are willing to do at each of the following wages by adjusting the slider below. Recall that, in addition to the number of tasks you select here, everyone will be required to complete 1 tasks paid for by the completion fee. There is an equal chance that each of these wages will be selected as the decision-that-counts, so please make your decisions carefully. **Make sure to adjust all the sliders before continuing, otherwise you will be asked to re-enter your choices.**

| Low Wage (50%) | High Wage(50%) | | Chosen Tasks | |
|---|---|---|---|---|
| \$0.0/task | \$0.1/task | (50% Chance of \$0.00<br>50% Chance of \$2.20) | 22 tasks<br>(~16mins) | |
| \$0.0/task | \$0.2/task | (50% Chance of \$0.00<br>50% Chance of \$13.40) | 67 tasks<br>(~47mins) | |
| \$0.025/task | \$0.225/task | (50% Chance of \$0.48<br>50% Chance of \$4.28) | 19 tasks<br>(~14mins) | |
| \$0.05/task | \$0.25/task | (50% Chance of \$3.65<br>50% Chance of \$18.25) | 73 tasks<br>(~52mins) | |
| \$0.075/task | \$0.275/task | (50% Chance of \$1.95<br>50% Chance of \$7.15) | 26 tasks<br>(~19mins) | |

Hourly wage and time computed using task time of (sec): 42

☐ I confirm my final choices for all 5 sliders.

Next

# Effort Choices

Please indicate the number of tasks you are willing to do at each of the following wages by adjusting the slider below. Recall that, in addition to the number of tasks you select here, everyone will be required to complete 1 tasks paid for by the completion fee. There is an equal chance that each of these wages will be selected as the decision-that-counts, so please make your decisions carefully. **Make sure to adjust all the sliders before continuing, otherwise you will be asked to re-enter your choices.**

| Wage | Chosen Tasks |
|---|---|
| $0.2/task | |
| $0.225/task | |
| $0.25/task | |
| $0.275/task | |
| $0.3/task | |

Hourly wage and time computed using task time of (sec): 42

☐ I confirm my final choices for all 5 sliders.

Next

## Effort Choices

Please indicate the number of tasks you are willing to do at each of the following wages by adjusting the slider below. Recall that, in addition to the number of tasks you select here, everyone will be required to complete 1 tasks paid for by the completion fee. There is an equal chance that each of these wages will be selected as the decision-that-counts, so please make your decisions carefully. **Make sure to adjust all the sliders before continuing, otherwise you will be asked to re-enter your choices.**

| Wage | Chosen Tasks |
|---|---|
| $0.05/task | |
| $0.1/task | |
| $0.125/task | |
| $0.15/task | |
| $0.175/task | |

Hourly wage and time computed using task time of (sec): 42

☐ I confirm my final choices for all 5 sliders.

Next

# Effort Choices

Please indicate the number of tasks you are willing to do at each of the following wages by adjusting the slider below. Recall that, in addition to the number of tasks you select here, everyone will be required to complete 1 tasks paid for by the completion fee. There is an equal chance that each of these wages will be selected as the decision-that-counts, so please make your decisions carefully. **Make sure to adjust all the sliders before continuing, otherwise you will be asked to re-enter your choices.**

| Low Wage (50%) | High Wage(50%) | Chosen Tasks | |
|---|---|---|---|
| $0.1/task | $0.3/task | | |
| $0.125/task | $0.325/task | | |
| $0.15/task | $0.35/task | | |
| $0.175/task | $0.375/task | | |
| $0.2/task | $0.4/task | | |

Hourly wage and time computed using task time of (sec): 42

☐ I confirm my final choices for all 5 sliders.

Next

# Effort Choices

Please indicate the number of tasks you are willing to do at each of the following wages by adjusting the slider below. Recall that, in addition to the number of tasks you select here, everyone will be required to complete 1 tasks paid for by the completion fee. There is an equal chance that each of these wages will be selected as the decision-that-counts, so please make your decisions carefully. **Make sure to adjust all the sliders before continuing, otherwise you will be asked to re-enter your choices.**

| Low Wage (50%) | High Wage(50%) | Chosen Tasks |
|---|---|---|
| \$0.075/task | \$0.375/task | |
| \$0.05/task | \$0.45/task | |
| \$0.0/task | \$0.5/task | |
| \$0.1/task | \$0.5/task | |
| \$0.0/task | \$0.6/task | |

Hourly wage and time computed using task time of (sec): 42

☐ I confirm my final choices for all 5 sliders.

Next

# Part 1 Continued

## Decisions 31 and 32

Next, you will be asked to make your final two decisions for **Part 1** of this study. Each of the two decisions will be presented on their own page, so **please make sure you carefully review the rates for each decision**. As in the previous decisions, you will have to decide how many tasks to complete at different rates. The only difference in these two decisions is the structure of the rates: with 50.0% chance, you will get $0.20/task, with 5.0% chance you will get a fixed payment of $X *regardless of the number of tasks that you decided to do*, and with 45.0% chance you will get a fixed payment of $Y *regardless of the number of tasks that you decided to do*. For example, if you select to complete 30 tasks, then after you complete the 30 tasks you will either be paid $0.20/task, $X, or $Y.

Recall that at the end of the study, one of the 32 decisions you've made in **Part 1**, including these final two, may be randomly selected for payment. This means that each decision is equally likely to be the decision-that-counts. **Thus, it is very important that you think carefully about each decision you make, as it could be the one selected for payment.**

Next

# Effort Choices

Please indicate the number of tasks you are willing to do at each of the following wages by adjusting the slider below. Recall that there is an equal chance that each of these wages will be selected as the decision-that-counts. Make sure to adjust all the sliders before continuing, otherwise you will be asked to re-enter your choices.

| **Fixed (L) (5.0%)** | **Fixed (H) (45.0%)** | **Wage (50.0%)** | | **Chosen Tasks** |
|---|---|---|---|---|
| $0 | $20 | $0.2/task | (50% Chance of $3.20<br>5% Chance of $0.00<br>45% Chance of $20.00) | 16 tasks<br>(~12mins) |

Hourly wage and time computed using task time of (sec): 42

Next

## Effort Choices

Please indicate the number of tasks you are willing to do at each of the following wages by adjusting the slider below. Recall that there is an equal chance that each of these wages will be selected as the decision-that-counts. Make sure to adjust all the sliders before continuing, otherwise you will be asked to re-enter your choices.

| **Fixed (L) (45.0%)** | **Fixed (H) (5.0%)** | **Wage (50.0%)** | **Chosen Tasks** |
|---|---|---|---|
| $0 | $20 | $0.2/task | |

Hourly wage and time computed using task time of (sec): 42

Next

# Instructions for Part 2

On the following pages, you will be asked to make 21 choices per page. In each choice, you will be presented with two options -- "Option A" and "Option B" -- and asked to indicate which of the two you prefer.

"Option A" will be a lottery that pays either $10.00 ($3.00) with probability varying from 0.0% to 100.0%, or $0.00 (-$3.50) otherwise; "Option B" yields a payoff of $5.00 ($0.00) for sure, i.e. with a probability of 100%.

On each page, the first and last choice will be selected by default to help demonstrate that Option B is initially the preferred option, but Option A grows more desirable in each row; by the last choice, Option A should clearly be preferred. You will not be able to change these choices.

For the remaining choices, please select your preference between Option A and Option B. Once you have switched from Option B to Option A, all subsequent choices will be automatically switched to Option A. This is intended to help maintain consistency due to the ordering of the choices: if you prefer Option A to Option B in choice number 10 (for example), then you should prefer Option A to Option B in choice 11, 12, and so on. An example of this for a set of potential choices is shown below. Someone who prefers a 10% chance of $10 (and 90% chance of $0) to $5 for sure should also prefer a 15% chance of $10 (and 85% chance of $0) to $5 for sure, because a 15% chance is strictly better than a 10% chance and the $5 for sure never changes.

| **Option A** | | **Option B** |
|---|---|---|
| $10.00 with a probability of 0.0%,$0.00 otherwise | ○ Option A ◉ Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 5.0%,$0.00 otherwise | ○ Option A ◉ Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 10.0%,$0.00 otherwise | ◉ Option A ○ Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 15.0%,$0.00 otherwise | ◉ Option A ○ Option B | $5.00 with a probability of 100.0% |

**After you have made all of your choices, please review the page prior to submitting these decisions. Recall that one of these decisions may be randomly chosen for your payment.**

If you indicated that you prefer Option A (the lottery) for the relevant decision, a random number between 1 and 100 will be generated to determine the outcome of the lottery. For instance, if the **decision-that-counts** is $10.00 with 20% chance and $0.00 with 80% chance, a random number between 1-80 will result in payment of $0.00, but a random number between 81-100 will result in payment of $10.00.

If you indicated that you prefer Option B for the relevant decision, you would receive $5.00 in this example.

Recall that, along with your choices from **Part 1**, each of the choices you make in **Part 2** are equally likely to be the *decision-that-counts*. Please carefully consider each choice as they are all equally likely to determine your final payment.

Next

## Part 2 Decisions

| Option A | | Option B |
|---|---|---|
| $3.00 with a probability of 0.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 5.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 10.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 15.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 20.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 25.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 30.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 35.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 40.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 45.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 50.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 55.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 60.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 65.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 70.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 75.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 80.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 85.0%,-$3.50 otherwise | ○ Option A ● Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 90.0%,-$3.50 otherwise | ● Option A ○ Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 95.0%,-$3.50 otherwise | ● Option A ○ Option B | $0.00 with a probability of 100.0% |
| $3.00 with a probability of 100.0%,-$3.50 otherwise | ● Option A ○ Option B | $0.00 with a probability of 100.0% |

Next

# Part 2 Decisions

| Option A | | Option B |
|---|---|---|
| $10.00 with a probability of 0.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 5.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 10.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 15.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 20.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 25.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 30.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 35.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 40.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 45.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 50.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 55.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 60.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 65.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 70.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 75.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 80.0%,$0.00 otherwise | ○ Option A ● Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 85.0%,$0.00 otherwise | ● Option A ○ Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 90.0%,$0.00 otherwise | ● Option A ○ Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 95.0%,$0.00 otherwise | ● Option A ○ Option B | $5.00 with a probability of 100.0% |
| $10.00 with a probability of 100.0%,$0.00 otherwise | ● Option A ○ Option B | $5.00 with a probability of 100.0% |

Next

# Results

The following decision was randomly chosen for your payment:

| | Option A | | | Option B |
|---|---|---|---|---|
| | $10.00 with a probability of 85.0%, $0.00 otherwise | ◉ | ○ | $5.00 with a probability of 100.0% (sure payoff) |

As shown above, you indicated that you prefer Option A in this decision.

For the lottery, one of the two possible outcomes has been randomly realized based on the corresponding probabilities.

Your payoff in this task equals **$10.00**.

Next

# The Task

Recall that for each task, you will have to transcribe a line of blurry Greek letters from a Greek text. One task is one row of Greek text. For the task to be complete, your transcription must be 80% accurate or better. The following is an example of a row of blurry Greek text and the dictionary:

δηιφα.βαιδφι.ηβεχφηφιεδ.γηειφφγφαγ.

Please select from the following characters to enter your transcription.

α β χ δ ε φ γ η ι . ⌫

The correct transcription for the example task is provided; recall, however, that you do not need to be 100% accurate. If you submit a transcription that is 80% accurate or better (defined as requiring 7 or fewer insertions, deletions, or substitutions to achieve the perfect transcription), you will have completed the task. If your answer is incorrect (less than 80% accurate), you will have 2 more tries (for a total of 3) to submit a correct answer, after which you will be presented with a new task and the incorrect submissions will not be counted as a completed task.

As part of the task, several auditory "beeps" will sound randomly throughout the transcription process. Please put on your headphones and/or adjust the volume so that you can hear the beeping noises. After each time you hear this beeping noise, you must press the “Noise” button at the bottom left of the screen. If you do not press the “Noise” button within five seconds of hearing the beeping noise, your transcription will be reset. If you press the noise button erroneously (when there was no beeping sound), your transcription will be reset. Note that resetting the transcription does not count as a try; your current progress will simply be deleted and you will have to re-enter the transcription.

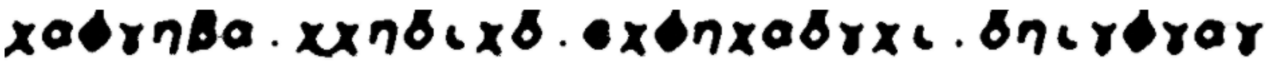
Once you click the next button, you will be presented with the 1 required tasks. After you complete these, you will be continue onto **Part 5** and attempt to solve several puzzles ($1 per correct submission) and answer a few demographic questions. Then, you will receive compensation based on your *decision-that-counts*. Recall that you were randomly selected to be paid $10.00 from the lottery task.
Once you finish all of these tasks, you will receive a completion fee of $7.00 in addition to your lottery payout, for a total of $17.00 (plus $1 per correct puzzle entry).

Next

# Mandatory Tasks

Please transcribe the row Greek letters by selecting the appropriate letters. Press **Next** when you wish to submit your response. Remember to press the **Noise** button within 5 seconds of hearing the beep, otherwise your responses will be removed and you will have to start over. You will have 3 chances to complete each task. If you fail to complete the task within three tries you will simply be shown a different one.

Number completed 0/1. Attempt: 1 of 3.



Please select from the following characters to enter your transcription.

$\alpha$ $\beta$ $\chi$ $\delta$ $\epsilon$ $\phi$ $\gamma$ $\eta$ $\iota$ .

Noise

Next

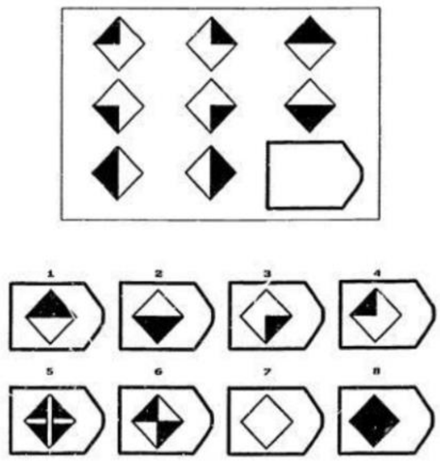
# Part III Solving puzzles

## Instructions

You will be presented with 5 problems, each showing a pattern with a bit cut out of it. Look at the pattern, think what piece is needed to complete the pattern correctly both along the rows and down the columns, BUT NOT THE DIAGONALS.

You will be paid $1 for each correct problem you solve. You have 10 minutes to complete all problems.

For example, for the following matrix, the correct pattern is 8.




Click next to start solving the problems.

Next

# Make your choice

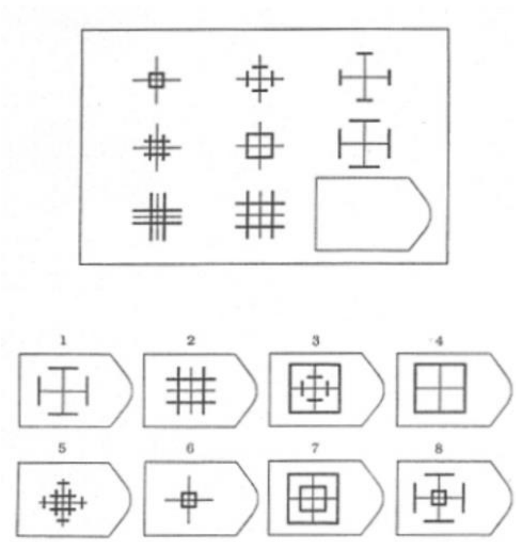
Time left to complete this section: **9:52**

## Question 1 of 5

## Instructions

You will be presented with 5 problems, each showing a pattern with a bit cut out of it. Look at the pattern, think what piece is needed to complete the pattern correctly both along the rows and down the columns, BUT NOT THE DIAGONALS.

You will be paid $1 for each correct problem you solve. You have 10 minutes to complete all problems.




Please choose an item that best fits the pattern:

---------

Next

# Make your choice

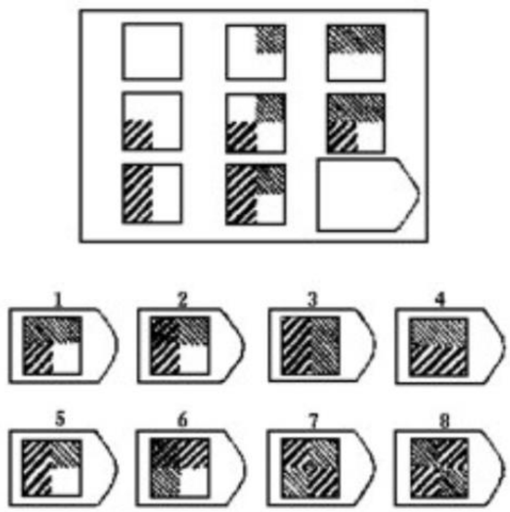
Time left to complete this section: **9:37**

## Question 2 of 5

## Instructions

You will be presented with 5 problems, each showing a pattern with a bit cut out of it. Look at the pattern, think what piece is needed to complete the pattern correctly both along the rows and down the columns, BUT NOT THE DIAGONALS.

You will be paid $1 for each correct problem you solve. You have 10 minutes to complete all problems.




Please choose an item that best fits the pattern:

---------

Next

## Make your choice

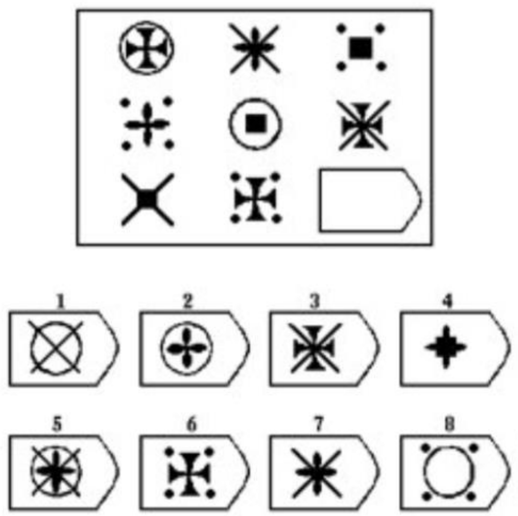
Time left to complete this section: **9:21**

## Question 3 of 5

## Instructions

You will be presented with 5 problems, each showing a pattern with a bit cut out of it. Look at the pattern, think what piece is needed to complete the pattern correctly both along the rows and down the columns, BUT NOT THE DIAGONALS.

You will be paid $1 for each correct problem you solve. You have 10 minutes to complete all problems.




Please choose an item that best fits the pattern:

4

Next

# Make your choice

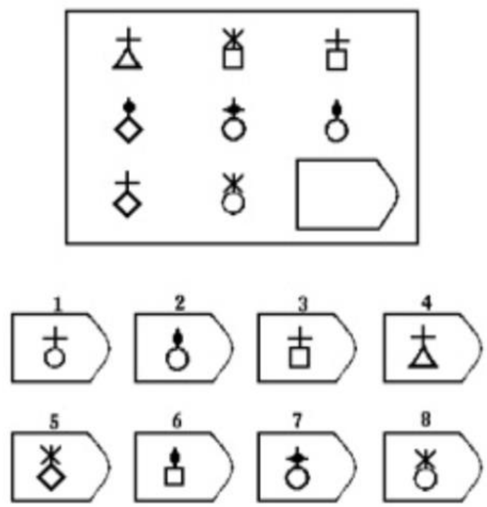
Time left to complete this section: **9:06**

## Question 4 of 5

## Instructions

You will be presented with 5 problems, each showing a pattern with a bit cut out of it. Look at the pattern, think what piece is needed to complete the pattern correctly both along the rows and down the columns, BUT NOT THE DIAGONALS.

You will be paid $1 for each correct problem you solve. You have 10 minutes to complete all problems.




Please choose an item that best fits the pattern:

---------

Next

# Make your choice

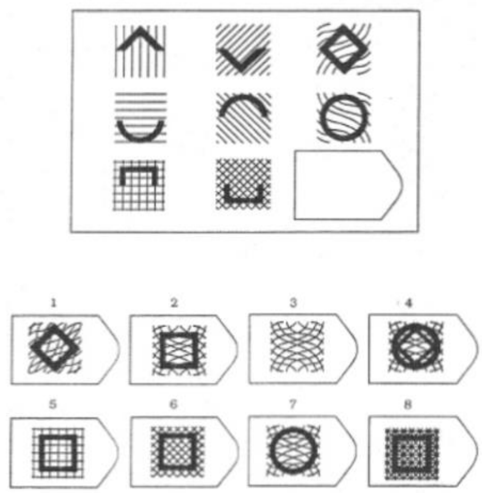
Time left to complete this section: **8:43**

## Question 5 of 5

## Instructions

You will be presented with 5 problems, each showing a pattern with a bit cut out of it. Look at the pattern, think what piece is needed to complete the pattern correctly both along the rows and down the columns, BUT NOT THE DIAGONALS.

You will be paid $1 for each correct problem you solve. You have 10 minutes to complete all problems.




Please choose an item that best fits the pattern:

---------

Next

## Results

You have completed all problems.

You have correctly solved 0 problems.

Your total payment for this part is $0.

Next

# Thank You for Participating

Before we finalize your earnings, please answer the following short survey.

What year of your undergraduate education are you in?

○ First ○ Second ○ Third ○ Fourth ○ Other

What is your major or intended major?

What is your gender?

○ Male ○ Female ○ Other ○ Decline to Answer

Which of the following income brackets do your parents fall into?

○ Below 50k ○ 50k to 100k ○ Above 100k ○ Decline to Answer

How do you evaluate yourself: Are you in general a more risk-taking (risk-prone) person (10) or do you try to avoid risks (0, risk-averse)?

○ 0 (Risk averse) ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 (Fully prepared to takes risks)

Next

# Thank You for Participating

As a reminder, your lottery payoff was randomly determined to be $10.00.
Your total earnings, including the completion fee of $7.00 and the earnings from the Raven Matrices of $0.00, are $17.00.

Next

# Appendix D Instructions - Exchange experiment

## D.1 Images of Objects Presented to subjects

The following images were projected to the wall of the lecture room at the beginning of the respective stage. For the displayed example, the Stage 1 pair consisted of the USB stick and erasable pens, but this was counter-balanced at the session level.


Figure A10: Image 1 Projected on the Wall to Present Objects.
For Stage 1 with objects pair consisting of USB stick and erasable pens.

## D.2 Original instructions in German (computer-based)

**Willkommen in Teil 1 von 2 in diesem Experiment!**


Figure A11: Image 2 Projected on the Wall to Present Objects.
For Stage 2 with objects pair consisting of thermos and picnic mat.

Bitte schließen Sie den Vorhang Ihrer Kabine und lesen die folgenden Informationen. Alle Eingaben, die Sie in diesem Experiment am Computer machen, sind völlig anonym und können nicht mit Ihrer Person in Verbindung gebracht werden. Es geht an keiner Stelle in diesem Experiment um Schnelligkeit. Bitte nehmen Sie sich stets ausreichend Zeit, um die Anweisungen zu lesen und zu verstehen.

Sie besitzen nun das Produkt vor Ihnen. Sie können es jederzeit anfassen und inspizieren. Bitte öffnen Sie jedoch noch nicht die Verpackung und benutzen das Produkt nicht.

Die beiden Ihnen vorgestellten Produkte wurden zufällig und in gleichen Mengen auf die Kabinen verteilt. Ihre Kabinennummer hat sich ebenfalls rein zufällig aus der Wahl Ihres Sitzplatzes im Präsentationsraum ergeben.

Klicken Sie OK, wenn Sie diese Informationen gelesen haben. Falls Sie Fragen haben, rufen Sie bitte den Leiter des Experiments.

**Bitte beantworten Sie die Fragen.**
[ USB stick / Thermoskanne ]
Wie gut gefällt Ihnen das Produkt?
Wie gern würden Sie dieses Produkt mitnehmen?

[ Radierbare Kugelschreiber / Picknick-Matte ]
Wie gut gefällt Ihnen das Produkt?
Wie gern würden Sie dieses Produkt mitnehmen?

Wenn Sie sich für ein Produkt entscheiden müssten, welches würden Sie lieber behalten?

[ USB stick / Thermoskanne ] [ Radierbare Kugelschreiber / Picknick-Matte

**Bitte lesen Sie die folgenden Informationen aufmerksam.**
Der Leiter des Experiments wird gleich mit einer Bingo-Trommel eine Zufallszahl zwischen 1 und 20 ziehen.Die gezogene Zahl wird danach laut durchgesagt. Wenn die gezogene Zahl eine Zahl [ von 11 bis 20 / von 1 bis 10 ] ist, werden/wird [ Ihr USB-Stick / Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] weggenommen und Sie erhalten stattdessen eine/einen [ USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ]. Wenn die gezogene Zahl eine Zahl [ von 1 bis 10 / von 11 bis 20 ] ist, behalten Sie [ Ihr USB-Stick / Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] und es passiert nichts. Nachdem die Zahl gezogen wurde und gegebenenfalls ein Austausch der Produkte vollzogen wurde, passiert nichts mehr in diesem Teil des Experiments. Sie können das Produkt dann endgültig behalten.
Bitte bestätigen Sie erst, wenn Sie alles verstanden haben. Falls Sie Fragen haben, rufen

Sie bitte den Leiter des Experiments und warten, bis er zu Ihnen kommt.

**[ Mood elicitation 1 ]**

Bitte beantworten Sie die folgenden Fragen dazu, wie Sie sich gerade fülen. Welche Ausdrücke treffen auf Sie jetzt im Moment eher zu? Positionieren Sie den Schieberegler entsprechend. "Unglücklich, Wütend, Unzufrieden, Traurig, Verzweifelt"—"Glücklich, Begeistert, Zufrieden, Frühlich"

**Es ist soweit! Bitte warten Sie, bis die Zahl gezogen wurde.**

Zur Erinnerung: Wenn die gezogene Zahl [ von 11 bis 20 / von 1 bis 10 ] ist, verlieren Sie [ Ihr USB-Stick / Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] und erhalten stattdessen eine/einen [ USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ]. Wenn die gezogene Zahl [ von 1 bis 10 / von 11 bis 20 ] ist, behalten Sie [ Ihr USB-Stick / Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ].

**Die gezogene Zahl ist [ 1 / 2 / ... / 20 ].**

Dies ist eine Zahl [ von 1 bis 10 / von 11 bis 20 ]. Daher [ verlieren Sie [ Ihren USB-Stick / Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] und erhalten stattdessen eine/einen [ USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] / können Sie [ Ihren USB-Stick"/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] behalten ]. Bitte warten Sie, während der Leiter des Experiments den Austausch in den Kabinen durchführt.

**[ Mood elicitation 2 and control question. ]**

Bitte beantworten Sie die folgenden Fragen dazu, wie Sie sich gerade fülen. Welche

Ausdrücke treffen auf Sie jetzt im Moment eher zu? Positionieren Sie den Schieberegler entsprechend. "Unglücklich, Wütend, Unzufrieden, Traurig, Verzweifelt"—"Glücklich, Begeistert, Zufrieden, Frühlich"

In der Lottoziehung die eben stattgefunden hat: Wie hoch war die Wahrscheinlichkeit (in Prozent), dass Sie Ihr ursprüngliches Produkt verlieren würden? Bitte geben Sie eine Zahl zwischen 0 und 100 ein. Please enter a number between 0 and 100.

**Teil 1 des Experiments ist vorbei!**

Bitte befolgen Sie die Anweisungen.

- Prägen Sie sich die Nummer Ihrer Kabine ein.
- Sie können jetzt zurück in den Präsentationsraum gehen.
- Bitte lassen Sie [ Ihren USB-Stick/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] in der Kabine liegen. Sie werden in wenigen Minuten zurück in der gleichen Kabine sein.
- Zur Erinnerung: Das Produkt gehört nun endgültig Ihnen und Sie werden es mit aus dem Experiment nehmen.

**Willkommen in Teil 2 in diesem Experiment!**

Bitte schließen Sie den Vorhang Ihrer Kabine und lesen die folgenden Informationen.

Sie besitzen nun den/die [ USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] vor Ihnen. Sie können es jederzeit anfassen und inspizieren. Bitte öffnen Sie jedoch noch nicht die Verpackung und benutzen das Produkt nicht.

Die beiden für Teil 2 vorgestellten Produkte ( [ USB Stick und radierbare Kugelschreiber ] / [ Thermoskanne und Picknick-Matte ]) wurden erneut zufällig und in gleichen Mengen auf die Kabinen verteilt.

Klicken Sie OK, wenn Sie diese Informationen gelesen haben. Falls Sie Fragen haben, rufen Sie bitte den Leiter des Experiments.

**[ Instructions Stage 2—ONLY BASELINE (p=0.0) ]**

Bitte lesen Sie die folgenden Informationen aufmerksam. Der/Die [ USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] aus Teil 2 des Experiments gehört nun Ihnen und Sie können es behalten. Wenn Sie möchten, können Sie [ Ihren USB-Stick/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] freiwillig gegen ein/eine [ USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] tauschen. Wie auch immer Sie sich entscheiden, Ihre Wahl ist endg¸ltig und Sie werden Ihr gewähltes Produkt danach mit aus dem Experiment nehmen. Bitte bestätigen Sie erst, wenn Sie alles verstanden haben. Falls Sie Fragen haben, rufen Sie bitte den Leiter des Experiments und warten, bis er zu Ihnen kommt

**[ Instructions Stage 2—ONLY FORCED EXCHANGE (p=0.5) ]**

Bitte lesen Sie die folgenden Informationen aufmerksam. Sie haben ein neues Produkt in Teil 2 des Experiments erhalten ( [ einen USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] ).

Sie erhalten gleich die Gelegenheit, [ Ihren USB-Stick/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] freiwillig gegen [ einen USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] zu tauschen. Wenn Sie sich für einen Tausch entscheiden, erhalten Sie wie gewüscht [ einen USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] für [ Ihren USB-Stick/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] und können [ Ihren USB-Stick/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] endgültig behalten. Das Experiment ist damit abgeschlossen.

Wenn Sie sich gegen einen Tausch entscheiden, besteht danach eine Wahrscheinlichkeit von 50%, dass der Austausch dennoch erzwungen wird und sie trotzdem tauschen müssen.

Folgendes passiert konkret im Fall, dass Sie sich gegen einen freiwilligen Tausch entscheiden: Der Leiter des Experiments wird (wie in Teil 1 des Experiments) mit einer Bingo-Trommel eine Zufallszahl zwischen 1 und 20 ziehen. Die gezogene Zahl wird danach laut durchgesagt. Wenn die gezogene Zahl eine Zahl [ von 11 bis 20 / von 1 bis 10 ] ist, wird/werden Ihnen [ Ihr USB-Stick/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] weggenommen und Sie erhalten stattdessen [ einen USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ]. Wenn die gezogene Zahl eine Zahl [ von 1 bis 10 / von 11 bis 20 ] ist, behalten Sie [ Ihren USB-Stick/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] und es passiert nichts. Nachdem die Zahl gezogen wurde und gegebenenfalls ein Austausch der Produkte vollzogen wurde, passiert nichts mehr in diesem Teil des Experiments. Sie können das Produkt dann endgültig behalten.

Bitte bestätigen Sie erst, wenn Sie alles verstanden haben. Falls Sie Fragen haben, rufen Sie bitte den Leiter des Experiments und warten, bis er zu Ihnen kommt

**[ Mood elicitation 3 ]**

Bevor Sie die Möglichkeit erhalten, Ihr Produkt zu tauschen, beantworten Sie bitte die folgenden Fragen dazu, wie Sie sich gerade fülen. Welche Ausdrücke treffen auf Sie jetzt im Moment eher zu? Positionieren Sie den Schieberegler entsprechend. "Unglücklich, Wütend, Unzufrieden, Traurig, Verzweifelt"—"Glücklich, Begeistert, Zufrieden, Frühlich"

**Möchten Sie [ Ihren USB-Stick/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] gegen [ einen USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] tauschen?**

Ja, ich möchte tauschen.

Nein, ich möchte nicht tauschen.

**[ ONLY BASELINE (p=0.0) ]**

Sie haben sich [ für / gegen ] einen freiwilligen Tausch enschieden. Bitte warten Sie, während der Leiter des Experiments den Austausch in den Kabinen durchführt.

**[ ONLY FORCED EXCHANGE (p=0.5) ]**

Sie haben sich [ für / gegen ] einen freiwilligen Tausch enschieden. Bitte warten Sie, während der Leiter des Experiments den Austausch in den Kabinen durchführt.

[ ONLY NON-TRADERS ] Danach entscheidet sich, ob Sie trotzdem tauschen müssen.

[ ONLY TRADERS ] Bitte warten Sie, bis das Experiment weitergeht. Es wird nun eine Zufallszahl für diejenigen gezogen, die sich gegen den freiwilligen Austausch entschieden haben. Danach geht das Experiment für Sie weiter.

[ ONLY NON-TRADERS ] Zur Erinnerung: Wenn die gezogene Zahl [ von 11 bis 20 / von 1 bis 10 ] ist, verlieren Sie [ Ihr USB-Stick"/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] und erhalten stattdessen eine/einen [ USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ]. Wenn die gezogene Zahl [ von 1 bis 10 / von 11 bis 20 ] ist, behalten Sie [ Ihr USB-Stick"/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ].

[ ONLY NON-TRADERS ]

Die gezogene Zahl ist [ 1 / 2 / ... / 20 ].

Dies ist eine Zahl [ von 1 bis 10 / von 11 bis 20 ]. Daher [ verlieren Sie [ Ihren USB-Stick / Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] und erhalten stattdessen eine/einen [ USB-Stick / radierbare Kugelschreiber / eine Picknick-Matte / eine Thermoskanne ] / können Sie [ Ihren USB-Stick"/ Ihre radierbaren Kugelschreiber / Ihre Picknick-Matte / Ihre Thermoskanne ] behalten ]. Bitte warten Sie, während der

Leiter des Experiments den Austausch in den Kabinen durchführt.

**[ Mood elicitation 4 ]**

Bitte beantworten Sie die folgenden Fragen dazu, wie Sie sich gerade fülen. Welche Ausdrücke treffen auf Sie jetzt im Moment eher zu? Positionieren Sie den Schieberegler entsprechend. "Unglücklich, Wütend, Unzufrieden, Traurig, Verzweifelt"—"Glücklich, Begeistert, Zufrieden, Frühlich"

**Das Experiment ist zu Ende!**

Sie können beide Produkte behalten. Zudem erhalten Sie gleich eine Teilnahmevergütung von 4 Euro. Bitte warten Sie noch kurz in Ihrer Kabine, bis Sie der Experimentator herausruft. Vielen Dank für Ihre Teilnahme!

## D.3 English translation of instructions

**Welcome to part 1 of 2 in this experiment!**

Please close the curtain of you cabin and read the following information. All computer entries that you make in this experiment are fully anonymous and cannot be traced back to you. Speed is not important at any point in this experiment. Please always take sufficient time to read and understand the instructions.

You are currently in possession the product in front of you. You may touch it and inspect it anytime. However, please do not open the packaging and do not use the product The two objects presented to you ( [ USB stick and erasable pens / thermos and picnic mat ] ) have been randomly allocated to the cabins in equal quantities. Your cabin number was also randomly determined based on your choice of seat in the presentation room.

Please click on OK when you have read these information. If you have questions, please call an experimenter.

**Please answer the questions.**

[ USB stick / thermos ]

How much do you like this product?

How much would you want to have this product?

[ Erasable pens / picnic mat ]

How much do you like this product?

How much would you want to have this product?

If you had to choose one of the objects, which one would you prefer to keep?

[ Erasable pens / picnic mat ] [ USB stick / thermos ]

**Please read the following information carefully.**

The experimenter will soon draw a random number between 1 and 20 using a lotto drum. The drawn number will then be announced loudly. If the drawn number is a number [ from 11 to 20 / from 1 to 10 ], your [ USB stick / erasable pens / thermos / picnic mat ] will be taken away from you and you instead receive [ USB stick / erasable pens / thermos / picnic mat ]. If the drawn number is a number [ from 1 to 10 / from 11 to 20 ], you will keep your [ USB stick / erasable pens / thermos / picnic mat ] and nothing happens. After the number has been drawn and the exchange of objects has taken place (if applicable), nothing else happens in this part of the experiment. You can then keep your object for good.

Please only confirm below once you have understood everything. If you have questions, please call the experimenter and wait until he comes to your cabin.

**[ Mood elicitation 1 ]**

Please answer the following questions about how you currently feel. Which expressions

better apply to you at the moment?
"Unhappy, Angry, Unsatisfied, Sad, Desperate"—"Happy, Thrilled, Satisfied, Content, Hopeful"

**The time has come. Please wait until the number has been drawn.**
Remember: If the drawn number is a number [ from 11 to 20 / from 1 to 10 ], your [ USB stick / erasable pens / thermos / picnic mat ] will be taken away from you and you instead receive [ USB stick / erasable pens / thermos / picnic mat ]. If the drawn number is a number [ from 1 to 10 / from 11 to 20 ], you will keep your [ USB stick / erasable pens / thermos / picnic mat ].

**The drawn number is [ 1 / 2 / ... / 20 ].**
This number is a number [ from 1 to 10 / from 11 to 20 ]. Therefore [ you can keep your [ USB stick / erasable pens / thermos / picnic mat ] / your [ USB stick / erasable pens / thermos / picnic mat ] will be taken away from you and you instead receive [ USB stick / erasable pens / thermos / picnic mat ] ]. Please wait while the experimenter carries out the exchange in all cabins.

**[ Mood elicitation 2 and control question. ]**
Please answer the following questions about how you currently feel. Which expressions better apply to you at the moment?
"Unhappy, Angry, Unsatisfied, Sad, Desperate"—"Happy, Thrilled, Satisfied, Content, Hopeful"

Regarding the lottery draw, that has just taken place: What was the probability (in percent) that you would lose your initial object? Please enter a number between 0 and 100.

**Part 1 of the experiment is over!**

Please follow the instructions.

- Memorize your cabin number.
- You can no go back to the presentation room.
- Please leave your [ USB stick / erasable pens / thermos / picnic mat ] in the cabin. You will be back in the same cabin in a few minutes.
- Remember: The object now belongs to you for good and you will take it away from this experiment.

**Welcome to part 2 in this experiment!**

Please close the curtain of you cabin and read the following information. You are now also in possession of the [ USB stick / erasable pens / thermos / picnic mat ] in front of you. You can touch and inspect it at any time. However, please do not yet open the packaging and do not use the object yet. The two objects presented to you for part 2 ( [ USB stick and erasable pens / thermos and picnic mat ] ) have again been randomly allocated to the cabins in equal quantities.

Please click on OK when you have read these information. If you have questions, please call an experimenter.

**[ Instructions Stage 2—ONLY BASELINE (p=0.0) ]**

Please read the following information carefully. The [ USB stick / erasable pens / thermos / picnic mat ] from part 2 of the experiment now belongs to you and you can keep it for good. If you like, you can exchange your [ USB stick / erasable pens / thermos / picnic mat ] voluntarily for [ USB stick / erasable pens / thermos / picnic mat ]. Whichever way you decide, your choice is final and you will take your selected object with you from this experiment.

Please only confirm below once you have understood everything. If you have questions, please call the experimenter and wait until he comes to your cabin.

**[ Instructions Stage 2—ONLY FORCED EXCHANGE (p=0.5) ]**

Please read the following information carefully. You have received a new object in part 2 of the experiment ( [ USB stick / erasable pens / thermos / picnic mat ] ). You will soon get the opportunity to exchange your [ USB stick / erasable pens / thermos / picnic mat ] voluntarily for [ USB stick / erasable pens / thermos / picnic mat ].

If you decide to exchange, you will receive [ USB stick / erasable pens / thermos / picnic mat ] as requested for your [ USB stick / erasable pens / thermos / picnic mat ] and you can then keep your [ USB stick / erasable pens / thermos / picnic mat ] for good. The experiment is then finished.

If you decide against an exchange, there will be a probability of 50 percent that the exchange will be forced anyways and you have to exchange nevertheless.

Concretely, the following happens in the case that you decide against a voluntary exchange: The experimenter will draw a random number between 1 and 20 using a lotto drum (as in part 1 of the experiment). The drawn number will then be announced loudly. If the drawn number is a number [ from 11 to 20 / from 1 to 10 ], your [ USB stick / erasable pens / thermos / picnic mat ] will be taken away from you and you instead receive [ USB stick / erasable pens / thermos / picnic mat ]. If the drawn number is a number [ from 1 to 10 / from 11 to 20 ], you will keep your [ USB stick / erasable pens / thermos / picnic mat ] and nothing happens. After the number has been drawn and the exchange of objects has taken place (if applicable), nothing else happens in this part of the experiment. You can then keep your object for good.

Please only confirm below once you have understood everything. If you have questions, please call the experimenter and wait until he comes to your cabin.

**[ Mood elicitation 3 ]**

Before you get the opportunity to exchange your object, please answer the following questions about how you currently feel. Which expressions better apply to you at the moment? "Unhappy, Angry, Unsatisfied, Sad, Desperate"—"Happy, Thrilled, Satisfied, Content, Hopeful"

**Do you want to exchange your [ USB stick / erasable pens / thermos / picnic mat ] for a [ USB stick / erasable pens / thermos / picnic mat ]?**

Yes, I want to exchange.

No, I do not want to exchange.

**[ ONLY BASELINE (p=0.0) ]**

You have decided [ for / against ] a voluntary exchange. Please wait while the experimenter carries out the exchange in all cabins.

**[ ONLY FORCED EXCHANGE (p=0.5) ]**

You have decided [ for / against ] a voluntary exchange. Please wait while the experimenter carries out the exchange in all cabins.

[ ONLY NON-TRADERS ] After this, it will be determined whether you have to exchange anyways.

[ ONLY TRADERS ] Please wait until the experiment continues. A random number will now be drawn for those who decided against a voluntary exchange. After that the experiment continues for you.

[ ONLY NON-TRADERS ] Remember: If the drawn number is a number [ from 11 to 20 / from 1 to 10 ], your [ USB stick / erasable pens / thermos / picnic mat ] will be taken away from you and you instead receive [ USB stick / erasable pens / thermos / picnic mat ]. If the drawn number is a number [ from 1 to 10 / from 11 to 20 ], you will keep your [ USB stick / erasable pens / thermos / picnic mat ].

[ ONLY NON-TRADERS ]

The drawn number is [ 1 / 2 / ... / 20 ]

This number is a number [ from 1 to 10 / from 11 to 20 ]. Therefore [ you can keep you [ USB stick / erasable pens / thermos / picnic mat ] / your [ USB stick / erasable pens / thermos / picnic mat ] will be taken away from you and you instead receive [ USB stick / erasable pens / thermos / picnic mat ]. Please wait while the experimenter carries out the exchange in all cabins.

**[ Mood elicitation 4 ]**

Please answer the following questions about how you currently feel. Which expressions better apply to you at the moment?

“Unhappy, Angry, Unsatisfied, Sad, Desperate”—“Happy, Thrilled, Satisfied, Content, Hopeful”

**The experiment is over!**

You can keep both your objects. You will also receive a show-up fee of 4 euros. Please wait shortly in you cabin until the experimenter calls you out. Thank you for your participation!