WORKING PAPERS

Pay Every Subject or Pay Only Some?

Lisa R. Anderson
Beth A. Freeborn
Patrick McAlvanah

WORKING PAPER NO. 342

July 2019

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

Pay Every Subject or Pay Only Some?

Lisa R. Anderson
Department of Economics
College of William and Mary
lisa.anderson@wm.edu

Beth A. Freeborn
Bureau of Economics
Federal Trade Commission
bfreeborn@ftc.gov

Patrick McAlvanah
Bureau of Economics
Federal Trade Commission
pmcalvanah@ftc.gov

July 2019*

Abstract

One technique employed by budget-conscious researchers is to pay only some of the subjects for their choices in an experiment. We test the effect of paying some subjects versus paying all subjects in the context of risk preferences, controlling for the difference in stakes induced by paying only some subjects. Over two experiments, we demonstrate that paying only some subjects yields lower levels of risk aversion than does paying all subjects, though it yields more risk aversion than paying all subjects lower stakes with expected values equivalent to the "pay some" condition. We also demonstrate that paying only some subjects not only changes the level of risk aversion but also impacts the ordering of subjects by elicited risk aversion. Neither probability weighting nor standard experimental demographics were correlated with subjects' differences between these conditions. We exploit our multiple measurements of risk aversion to estimate a simple structural model of latent risk aversion, and use these results to derive a correction factor in order to approximate the results as if all subjects were paid high stakes. Our findings imply that probabilistically paying some subjects high stakes meaningfully impacts the elicited level of risk aversion, although it better approximates the experimental ideal of paying all subjects high stakes compared to paying all subjects lower stakes.

JEL Classification: C90, D81

*We thank Devesh Raval, David Schmidt, William Violette, and Daniel H. Wood for helpful comments. We thank the Charles Koch Foundation for financial support for this research. The views expressed in this article are those of the authors and do not necessarily represent those of the Federal Trade Commission or any of its Commissioners.

1 Introduction

Though the gold standard in experimental economics is to pay all subjects for decisions that reflect economically meaningful stakes, in some experiments or surveys a researcher cannot feasibly afford to do so. One technique employed by budget-conscious researchers is to pay only some of the subjects for their choices in an experiment, rather than pay every subject. For example, a researcher measuring time preferences may have all subjects choose between \$40 today versus \$50 in one year, and then pay one out of ten subjects chosen at random for their choice, rather than pay all subjects for their choice between \$4 today versus \$5 in one year. One rationale for paying only some subjects is mental accounting. Stakes that are sufficiently low, such as \$4 today versus \$5 in one year, may fall below an attention or perception threshold, and subjects' choices over such low stakes may not accurately reflect their true preferences. Paying one out of ten subjects \$50, rather than all subjects \$5, retains the expected budget of \$5 per subject but involves stakes that subjects might "take seriously." Another justification for paying only some subjects is to economize on transaction costs associated with payments.

A number of papers (Starmer and Sugden, 1991; Cubitt et al., 1998; Laury, 2006; see Charness et al. (2016) for a review) have tested the validity of choosing at random a subset of *questions* to count for payment in an experiment with multiple decisions, rather than paying for every decision in an experiment. The majority of these papers have found no difference in responses from paying for only a subset of questions versus paying for every question. However, less attention has been devoted to the effects, if any, of paying only some *subjects* for their choices, which is surprising given the relative frequency of this practice in economic experiments.

We test the effect of paying some subjects versus paying every subject in the context of eliciting risk preferences. Our main test involves three treatments. In the first treatment, subjects choose between lotteries with high stakes, and all subjects are paid for their choices. In the second treatment, subjects choose between lotteries with identical stakes as in the first treatment, but now only one out of eight subjects chosen at random is paid for their choice. In the third treatment, all subjects are paid for their choices, but subjects choose between lotteries with lower stakes. Specifically, the lotteries have expected values that are one-eighth that of the first treatment, which equalizes the expected values between the second and third treatment.

Our contribution is fourfold. We present the first study with the main focus of testing the effect of paying some versus paying every subject. Though some previous studies (Beaud and Willinger, 2015; Brokesova, et al., 2017) have examined this, these tests were included as secondary tests or robustness checks and were not the main focus of these studies. As such, these tests used smaller samples and between-subjects identification, so previous findings of no significant differences between paying some versus paying every subject are unsurprising. Further, some previous studies (Baltussen et al., 2012; Beaud and Willinger, 2015) did not control for the differences in expected values between conditions. By contrast, we employ larger samples, within-subject identification, and control for the expected value differences between conditions. Over two experiments, we demonstrate that randomly paying only some subjects yields less risk aversion than does paying every subject. However, paying some subjects yields more risk aversion than paying all subjects stakes with expected values equivalent to the "pay some" condition. Our results imply a mean risk aversion parameter such that an individual would be indifferent between a certain \$22.74 and a 50% chance of \$100 in our treatment in which all subjects are paid relatively higher stakes. In contrast, for the same gamble, this same subject would be indifferent to \$27.24 in the second treatment, in which only one of eight subjects is paid these higher stakes, and \$30.14 if all subjects were paid stakes with one-eighth the expected value.

Given the frequency of the practice of paying only some subjects, our results have implications for experimental design in economics, as well as other social sciences and survey design. In some instances, a researcher may be unconcerned if a particular elicitation method introduces directional bias, as long as all observations are biased similarly and the individual ordering is preserved. For example, a researcher may measure risk preferences to include as a control during analysis of another main parameter of interest, but be unconcerned with the actual levels of risk aversion. Our second contribution is to examine whether the ordering of subjects by risk aversion differs between our treatment conditions. We find (Spearman) rank correlations of between 0.54 and 0.76, suggesting that paying only some subjects does affect the level as well as the ordering of subjects' risk aversion. However, we find that the condition in which only some subjects are paid high stakes has a greater rank order correlation with the condition in which all subjects are paid high stakes than does paying all subjects lower stakes. That is, probabilistically paying some subjects high stakes elicits risk aversion levels that are more similar (in terms of both levels and rank ordering) to the experimental ideal of paying all subjects high stakes than does paying all subjects lower stakes.

Third, we explore potential mechanisms for differences in subjects' evaluations between treatments. In our second experiment, we run an additional treatment which enables us to fit a probability weighting parameter for each subject. We examine if subjects who exhibit a larger degree of probability weighting also display particularly large differences between our payment treatments. That is, we examine if the subjects who most over-weight small probabilities also over-weight the "one out of eight subjects will be paid" probability in our second treatment, relative to the third treatment in which every subject is paid with smaller stakes. We find that probability weighting is not significantly related to the difference in subjects' evaluations between the pay every subject versus the pay some subject conditions. We also test whether demographics and alternative hypothetical measures of risk-taking are related to subjects' responses in the various payment conditions. All subjects completed an exit survey, providing demographic characteristics such as gender, college major, and information about whether a subject financed their education through a job or student loans. The survey also included a series of hypothetical questions that are reflective of risk preferences, such as frequency of gambling and seat belt use, as well as an alternate measure of risk preferences developed by Barsky et al. (1997). None of the demographic variables or alternate measures of risk-taking were consistently predictive of the difference in subjects' responses between the different payment conditions.

Our experiment contains multiple measurements of an individual's risk aversion. We assume that the most costly treatment, where all individuals are paid high stakes, serves as a better measure of an individual's underlying true risk aversion than the lower cost methods in which either only some subjects are paid high stakes or all subjects are paid low stakes. We exploit our multiple measurements to estimate a simple structural model in which an individual's response to each treatment is a product of latent risk aversion plus measurement error. For researchers who cannot afford to pay all subjects high stakes, our final contribution is to use these structural estimates to construct a correction factor to apply to similarly-sized gambles in order to better approximate the results as if all subjects were paid high stakes.

2 Experimental Design

A total of 192 subjects participated in two experiments consisting of 24 sessions with eight subjects in each. All subjects were undergraduate students at the College of William and Mary and were recruited in a variety of large introductory-level courses. Upon arrival at the session, each subject received a $5 show up payment and was provided with paper, pencils and a calculator. Subjects were separated by dividers and all of the instructions in the Appendix were read aloud. Subjects completed a series of multiple price lists (hereafter, MPL) tasks that were variations of the Holt and Laury (2002) lottery choice experiment. These tasks are commonly used to assess risk tolerance, and the general finding is that subjects tend to be slightly risk-averse on average. We balanced each group of eight subjects with four male and four female students, as some studies have found that female subjects are more risk-averse than male subjects[1]. The 24 sessions were split equally between two experiments, so that each experiment contained 96 subjects.

Each MPL consisted of several rows of choices between two risky lotteries, A and B.

[1] See Croson and Gneezy (2009) for a survey of the literature on gender differences in experiments.

Lottery B always had a higher maximum payoff than lottery A, but also had a higher variance in payoffs. With each successive row in the MPL, the probability of winning the higher prize amount increased by ten percentage points. The expected value of the riskier lottery B thus increases relative to lottery A as one moves down the rows and the probability of the larger option in each lottery increases. Typically, a respondent chooses the safer lottery for a portion of the choices and switches to the riskier lottery for the remainder of choices. Within an MPL, a higher fraction of choices for the safer lottery A indicates a greater level of risk aversion.

2.1 Experiment 1: Three Treatments

Experiment 1 contained three different MPL treatments. The "*AllHigh*" treatment, presented in Appendix Table 1, consisted of relatively high payoffs, and all eight subjects in a session were paid. In this treatment, Option A has lower possible payoffs ($32 or $25.60) than Option B ($61.60 or $1.60). Decision 10 is a choice between certain amounts of money: $32 for Option A and $61.60 for Option B. This decision serves as a rationality check; we would expect a rational subject to always choose Option B in Decision 10.[2]

Subjects in Experiment 1 faced two additional MPL treatments that varied with respect to the payoffs associated with each option or with the number of subjects who were paid for the treatment. The "*SomeHigh*" treatment (Appendix Table 2) was identical to the *AllHigh* treatment with one exception. Although subjects faced high payoffs, only one subject out of the eight in each session was paid for the *SomeHigh* treatment. That subject was randomly chosen in each session. The "*AllLow*" treatment (Appendix Table 3) contained payoffs that were 1/8th of the payoffs in the *SomeHigh* and *AllHigh* treatments, $4 or $3.20 for Option A and $7.70 or $0.20 for Option B. As suggested by the label, all subjects were paid for the *AllLow* treatment. Thus, expected payoffs were identical across the *SomeHigh* treatment and the *AllLow* treatment, since one in eight subjects was paid in the *SomeHigh* treatment. Once

[2]There were 672 subject-treatment conditions: 3 treatments for 96 subjects in Experiment 1 and 4 treatments for 96 subjects in Experiment 2 Out of these 672 responses, seven subjects chose the dominated response for decision 10. If a subject chose the dominated response, we dropped his or her response for that treatment.

subjects made all ten decisions in all three treatments, one of the decisions was randomly selected by a 10-sided die throw in each treatment. A second 10-sided die throw determined the payout for the selected decision, and a final ten-sided die throw selected which of the eight subjects was chosen for payment in the *SomeHigh* condition, with the die re-rolled if the result was a 9 or 10. Importantly, subjects completed their MPL choices over all treatments before any of the random numbers to determine the decisions used for payoffs and the outcome of the lotteries were generated, in order to minimize any wealth effects from one lottery result affecting choices in other lotteries.

Our *AllHigh* treatment thus represents the ideal experimental condition of paying all subjects high stakes. Our *SomeHigh* treatment enables us to test whether subjects treat probabilistic payments differently than the certain payment in the *AllHigh* condition, and our *AllLow* treatment allows us to test if any difference between the *AllHigh* and *SomeHigh* conditions was merely due to the expected value differences between those conditions.

2.2 Experiment 2: Four Treatments

When eliciting risk preferences, the structure of paying some subjects naturally induces a compound lottery. In our *SomeHigh* treatment, subjects face a probabilistic first stage as to whether they are randomly selected for payment, and then a second stage in which they face the underlying simple lottery. The reduction of compound lotteries axiom of Expected Utility Theory states that individuals should evaluate a compound lottery identically to its corresponding simple lottery that generates the same probability distribution over outcomes. Researchers have long questioned the validity of this reduction of compound lotteries axiom (Bar-Hillel, 1973; Kahneman and Tversky; 1979). For example, one of the components of Kahneman and Tversky's (1979) prospect theory is that when evaluating risky choices, individuals simplify the comparison by editing out common features between the risky gambles, known as the isolation effect. The authors compared subjects' preferences to two different hypothetical gambles. In the first scenario, subjects faced a two-stage game with a 75%

chance of ending the game with nothing, and a 25% chance of proceeding to the second stage. In the second stage, subjects chose between a 80% chance of $4,000 versus a certain $3,000.[3] In the second scenario, subjects chose between a simple lottery of a 20% chance of $4,000 versus a 25% chance of $3,000. Although these outcomes have the same final probabilities in both scenarios, 78% of respondents preferred the $3,000 in the first scenario which was framed as a compound lottery, whereas only 35% of respondents preferred the $3,000 option in the second scenario. The authors suggest that individuals do not fully account for the 75% chance of ending the game in the first scenario as it is common to both options, and therefore isolated out during the utility evaluations, leading to the preference reversal between two otherwise equivalent outcomes.

A number of papers have tested whether individuals evaluate compound lotteries in accordance with the Expected Utility axiom (Bar-Hillel, 1973; Bernasconi and Loomes, 1992; Miao and Zhong, 2012; Abdellaoui et al., 2015; Harrison et al., 2015; Hajimoladarvish, 2018), with the majority finding that individuals do not reduce compound lotteries in adherence to Expected Utility. If individuals do not treat compound lotteries equivalently to their corresponding simple lotteries, then it seems natural for individuals to evaluate an experiment in which all subjects are paid higher stakes probabilistically as different than one in which all subjects are paid lower stakes.

Several theories have been proposed to account for individuals' failures to reduce compound lotteries (Kreps and Porteus, 1978; Kahneman and Tversky, 1979; Segal, 1990). We consider one avenue for failure to evaluate compound lotteries in accordance with Expected Utility: improper weighting of the first stage of the gamble due to probability weighting. We test whether an individual's degree of probability weighting (whereby individuals overweight small probabilities and under-weight large probabilities) is associated with differences in elicited risk aversion between our treatments. To test this hypothesis, we ran a second experiment which includes the same three treatments above, as well as a fourth treatment,

[3] In actuality, the outcomes were denominated in Israeli currency, not dollars.

"*AllLowProb.*" Whereas the *AllLow* treatment was constructed by multiplying the monetary prizes of the *AllHigh* condition by 1/8th, the *AllLowProb* condition is constructed by multiplying all probabilities in the *AllHigh* condition by 1/8th. The payoff options for this *AllLowProb* treatment, in which all subjects are paid but the options within each MPL have lower probabilities, are shown in Appendix Table 4.

Note that the non-zero payoff amounts for the *AllLowProb* treatment match the payoff amounts in the *AllHigh* and the *SomeHigh* treatments. However, the probabilities of receiving these payoffs are 1/8th the probabilities in the *AllHigh* treatment. For instance, in Decision 1 of the *AllHigh* treatment, subjects who pick Option A have a 1/10 (10%) chance of winning \$32 and a 9/10 (90%) chance of winning \$25.60. In Decision 1 of the *AllLowProb* treatment, subjects who pick Option A have a 1/80 (1.25%) chance of winning \$32 and a 9/80 (11.25%) chance of winning \$25.60. Subjects also were explicitly informed in the *AllLowProb* treatment that there is a substantial chance (87.5% in Decision 1) they will earn \$0.[4]

This *AllLowProb* treatment allows us to examine whether subject probability weighting is a potential mechanism to explain any differences in risk parameters observed between the *SomeHigh* and *AllLow* treatments. For the analysis below, we used variation in decision-making between *AllHigh* and *AllLowProb* to identify a subject-level probability weighting parameter.[5] In addition to identifying and testing probability weighting, Experiment 2 also serves as a robustness check, as our *AllLowProb* condition has the same expected value as our *SomeHigh* condition, but due to varying the probability by 1/8th as compared to altering the amounts by 1/8th.

The focus of our study is an examination of how payment procedures affect behavior. A consistent finding across MPL-based studies is that risk tolerance varies widely across

[4]Because this treatment involved probabilities that could not be implemented with a die roll, we used a random number generator to determine which decision row was used. We then generated another random number to determine the payoffs dependent on the option chosen by the subject.

[5]Prior to implementing Experiment 2, we ran simulations to determine how many additional treatments at various low probabilities were necessary to identify the probability weighting parameter. There was very little change in the fitted probability parameters when additional blocks of questions were added. Given concerns regarding subject fatigue and budget, we chose to add a single low probability treatment.

individuals.[6] For this reason, we employed a within-subject identification rather than the between-subject approach used in some other studies. One concern that arises when using a within-subject approach is that questions or experiences in earlier treatments (e.g., die rolls) might have an effect on a subject's decisions in later treatments. To minimize order effects, the instructions that were read aloud contained payoffs that were different from the actual treatments used to determine earnings. To further account for the possibility that early treatments may influence responses in subsequent treatments, we divided the twelve sessions in Experiment 1 into six "order groups" representing every possible order in which the three treatments could be presented. For example, Order 1 presented the *AllHigh* treatment first, followed by the *SomeHigh* treatment, and finally the *AllLow* treatment. Order 2 presented the *AllHigh* treatment first, the *AllLow* treatment second, and finally the *SomeHigh.* We rotated these six different order conditions for two sessions each of our twelve sessions. Experiment 2 included four treatments which resulted in 24 possible orders in which to present the MPL decision sheets. Because we could not balance across all 24 possible orders, we used four orders such that each treatment was presented as first, second, third, or fourth exactly one time. These four orders were repeated three times.

At the end of each session, subjects completed a survey that included a series of demographic and behavioral questions pertinent to risk-behavior (see Appendix B). Among these were age, gender, and questions such as how the subject paid for college, seat belt use, speeding habits, and gambling habits. In addition to a $5 show up payment, subjects earned an average of $50.04 in Experiment 1 and $51.79 in Experiment 2 for their MPL choices. Sessions lasted an average of 45 minutes.

3 Data and Analysis

One difficulty in calculating measures of risk preferences via multiple price lists is the assumption of a functional form, if any, for the utility function. Another difficulty is how to treat

[6] See, for example, Choi et al. (2007).

individuals with multiple switching rows within an MPL, which violates preference monotonicity. For robustness, we calculate three different measures of individuals' risk tolerance; each measure utilizes different assumptions and yields different estimates for non-monotonic lottery choices.

Our first measure, Number of Safe Choices, is the number of safe choices an individual chose across the ten rows within each block of MPL questions. The advantage of this measure is that it is non-parametric, with no assumptions about the form of the utility function, and requires no assumptions about how to treat individuals with multiple switching rows.

Our second measure of risk preference, Switch CRRA, calculates individuals' risk preferences via their "switching row," the row in which the individual switched from safe to risky choices within the MPL. Switching between the safer to the riskier option implies that the subject was indifferent at some point along the interval between the two rows, which defines a range of values for a risk preference parameter. Since some subjects switch back and forth between the safer and riskier lottery more than once, we follow Harrison et al. (2007) in using the first switching point as the bottom of the range of values and the last switching point as the top of the range of values for the risk aversion parameter. We assume that an individual's utility for outcome x is the constant relative risk aversion (CRRA) utility form of $u(x) = \frac{x^{1-r}}{1-r}$, where larger values of r denote greater risk aversion.

Our third measure of risk tolerance, MLE CRRA, provides the most rigorous treatment of response error or missing responses. We assume that individuals treat each row of a MPL as a separate discrete choice task with error, and that an individual is more likely to choose lottery A as its expected utility increases relative to that of lottery B. We model expected utility conditional on the individual's risk aversion, and estimate via maximum likelihood the CRRA parameter that best explains an individual's lottery choices. We again assume that an individual's utility is given by the CRRA functional form as above. For the simplest specification, with no probability weighting, we assume that when faced with the MPL choice between the two risky lotteries A and B, individuals calculate the expected utility of each lottery, EU_A and EU_B, conditional on their risk aversion r. For example, the expected utility of lottery A is given by $EU_A = p_1 * u(x_1) + \dots p_n * u(x_n)$, where p_n denotes the probability of x_n. Individuals then choose lottery A or B based upon the difference between the two expected utilities, $EU_A - EU_B$. We assume that the probability that an individual chooses lottery A is $\Pr(A) = \Phi(EU_A - EU_B)$, where Φ represents the standard cumulative normal distribution function. The conditional log-likelihood of an individual's observed responses y_i in the lottery task depends upon risk aversion r and is thus:

$$\ln L(r, y) = \sum_i [(\ln(\Pr(A))|y_i = A) + (\ln(1 - \Pr(A))|y_i = B)]$$

3.1 Summary Statistics

We first present the percentage of subjects choosing the safe option in each decision row across our three treatments in Table 1. The left half of the table presents statistics for all subjects for completeness. The right half of the table presents statistics for our ultimate analysis sample, the 173 subjects with a valid MLE CRRA estimate for the *AllHigh*, *SomeHigh*, and *AllLow* treatments.[7] Subjects obeyed preference monotonicity in aggregate: for all treatments, the fraction of subjects choosing the safe option (weakly) decreases moving down the lottery decision rows, as the expected value of the riskier lottery increases relative to the safer lottery.

We next present the summary statistics for each of our three risk measures for each experiment separately in Table 2. We first test if the results are similar between Experiment 1 and Experiment 2. In our analysis sample, subjects in Experiment 1 made an average of 6.48 safe choices in *AllHigh*, whereas subjects in Experiment 2 made an average of 6.12 safe choices. For *SomeHigh*, subjects made an average of 6.01 safe choices in Experiment 1 and 6.00 safe choices in Experiment 2. For *AllLow*, subjects made an average of 5.83 and 5.67 safe choices in Experiments 1 and 2, respectively. We conduct a Wilcoxon rank-sum test (the non-parametric analogue version of a t-test for unmatched data) for each measure of risk

[7] Our main results and conclusions also still hold if we instead drop all subjects who ever switched rows more than once (violating monotonicity), which would yield an analysis sample of 161 subjects.

aversion, and for each treatment, to test if the different subjects responded similarly between Experiment 1 and Experiment 2. Table 2 presents the p-values from each of these tests. None of the risk aversion measures are significantly different at even the 10% level between Experiment 1 and Experiment 2, for any treatment, for either the full or the analysis sample.[8]

As a further test of data integrity and poolability, and an interesting question in its own right, we test if the number of switches in each lottery task varies between the experiments and between treatment conditions. Experiment 2 contains an additional treatment, *AllLowProb*, so subjects in Experiment 2 faced a longer experimental task and higher cognitive load, which may have led to more response errors and a greater number of monotonicity violations. Table 3 presents the number of switches in each lottery task, as well as the p-values from a Wilcoxon rank-sum test for differences between the two experiments. For each of the treatment conditions, there is no significant difference in the number of switches between the two experiments. In summary, despite the additional treatment in Experiment 2, and the different format of the *AllLowProb* treatment, we find no evidence that the responses in the common treatments differed significantly between the two experiments. As such, we pool the data between the two experiments and conduct analysis on this pooled data unless otherwise noted.

3.2 Impact of Payoff Treatments on Risk Aversion

The first scatterplot in Figure 1 depicts an individual's choice in the *AllHigh* treatment against his or her choice in the *SomeHigh* treatment, using the MLE CRRA risk measure, with the linear best fit line plotted through the data.[9] More mass lies below the 45 degree line, indicating that subjects on average displayed more risk aversion in the *AllHigh* condition than in the *SomeHigh* condition. The second scatterplot depicts an individual's choice in the *AllHigh* treatment against the *AllLow* treatment, and the third scatterplot depicts *SomeHigh*

[8] We obtain similar results with t-tests as well; none of the risk aversion measures were significantly different for any treatment at the 10% level, in either the full or analysis sample.

[9] Our data are jittered for graphical clarity. As our risk measures can only take on a small number of possible values, multiple subjects had identical values and thus overlap within the same graphing point. Jittering adds random noise to data to make clearer on a graph the true frequency of the number of individuals at each possible value. This random noise is not present during the actual data analysis.

against *AllLow*. An identical pattern holds: subjects made more risk-averse choices in *AllHigh* than in *AllLow*, and in *SomeHigh* compared to *AllLow*.

Table 4 presents the pooled means of Experiments 1 and 2 together for *AllHigh*, *SomeHigh*, and *AllLow*. For example, subjects chose an average of 6.31 safe choices in *AllHigh*, 6.01 safe choices in *SomeHigh*, and 5.75 safe choices in *AllLow*. Next, we present the main results of our paper, a test of whether subjects' estimated risk aversion differs by payoff treatment condition. For each of the three risk tolerance measures, we conduct a Wilcoxon signed-rank test (the non-parametric analog of a paired t-test) for each of the pairwise combinations of our three treatments (*AllHigh* vs. *SomeHigh*, *AllHigh* vs. *AllLow*, and *SomeHigh* vs. *AllLow*). For each of the three risk tolerance measures, we find that subjects are significantly more risk-averse in *AllHigh* than in *SomeHigh*. That is, subjects are more risk-averse over otherwise identical lottery questions when all subjects are guaranteed to receive a payment, compared to when only one in eight subjects will receive a payment. Similarly, subjects are significantly more risk-averse, on all three risk measures, in treatment *AllHigh* than in *AllLow*. Subjects are more risk-averse in the "high stakes" treatment in which all subjects are paid than in the "low stakes" treatment in which everyone is paid. These two results are not surprising, as the expected value of the lotteries are EV(*AllHigh*) $>$ EV(*SomeHigh*) $=$ EV(*AllLow*), and previous work has demonstrated that risk aversion increases as stakes rise (Binswanger, 1981; Holt and Laury, 2002; Fehr-Duda et al., 2010).

In the final column of Table 4, we present p-values testing for differences in risk preferences between *SomeHigh* and *AllLow*. These two lotteries have the same expected value, yet we find that risk aversion measures differ between these two treatments. Specifically, subjects are significantly more risk-averse in *SomeHigh* than in *AllLow* for all three risk measures. Although these two conditions have identical expected values, subjects are more risk-averse in the higher stakes lottery in which one out of eight subjects will be paid compared to a lottery with 1/8th the stakes sizes in which all subjects are paid. In summary, the evidence is overwhelmingly consistent that paying a fraction of subjects meaningfully impacts elicited risk aversion.

Another interesting result is that, for all three risk tolerance measures, subjects are most risk-averse in *AllHigh*, intermediate in *SomeHigh*, and display the least risk aversion in *AllLow*. That is, the condition in which only some subjects will be paid for high stakes leads to risk aversion that is intermediate compared to when all subjects are paid for high stakes versus when all subjects are paid for low stakes. Does the risk aversion elicited in *SomeHigh* more closely resemble *AllHigh*, which had identical questions save for the fact that all subjects were paid in *AllHigh*? Or does the risk aversion in *SomeHigh* more closely resemble *AllLow*, the condition in which all subjects are paid, but for lower stakes which equates the expected value to *SomeHigh*? To quantify this, for each of our risk aversion measures, we calculate the θ that solves *SomeHigh*$= \theta$*AllLow*$+(1-\theta)$*AllHigh*. For the number of safe choices; $\theta = 0.54$; for the switching row CRRA, $\theta = 0.51$; and for the MLE CRRA, $\theta = 0.59$. That is, for the first two risk aversion measures, *SomeHigh* lies equidistant between the *AllHigh* and *AllLow* conditions, whereas for the MLE CRRA risk measure *SomeHigh* more closely resembles the *AllLow* condition. Figure 2 depicts the distribution of responses for MLE CRRA across the three treatments. The distribution of responses in the *AllHigh* condition is "shifted right," indicating the greater risk aversion, compared to the intermediate risk aversion of the *SomeHigh* condition, or the *AllLow* condition.

Though not our main focus, we also present a comparison of the above three payment conditions to the *AllLowProb* condition in Table 5. Interestingly, subjects displayed the most risk aversion in the *AllLowProb* condition. The *SomeHigh* and *AllLowProb* are equivalent, except the *SomeHigh* condition is phrased as a compound gamble whereas the *AllLowProb* explicitly multiplies out the compound probabilities into simple lotteries. In the *AllLow* condition, the probabilistic chance of subject payment is effectively "multiplied through" into lower monetary amounts whereas the *AllLowProb* conditions the probabilistic payment is multiplied through the simple lotteries' probability amounts. Despite the fact that the *SomeHigh*, *AllLow*, and *AllLowProb* conditions have the same expected values, subjects were significantly more risk-averse in the *AllLowProb* condition. Our finding of more risk aversion in the *AllLowProb* condition compared to the *SomeHigh* condition is contrary to the isolation effect example of Prospect Theory above, which found less risk aversion when the compound probabilities were multiplied through into equivalent simple lotteries.

3.3 Impact of Payoff Treatments on the Rank Ordering of Risk Aversion

In some situations, a particular elicitation method may introduce bias, but this bias may be unconcerning for certain analyses. For example, a researcher may wish to control for risk aversion in a regression with a different main parameter of interest, but due to data availability, survey fatigue, or cost reasons, use a simple measure of risk preferences that is upwardly biased. In this setting, even though the measured levels are "incorrect" relative to the values from a more elaborate elicitation method, the bias is irrelevant if all observations are biased similarly and the individual ordering of risk aversion is preserved.

We have demonstrated that paying subjects probabilistically affects the elicited level of risk aversion. We now test whether the relative ordering of subjects by risk aversion is preserved between our treatment conditions. For each of the risk aversion measures, we compute the Spearman rank correlations between each of the treatment conditions. Table 6 presents these results. We find rank correlations ranging between 0.53 to 0.74 across our treatments and risk measures, indicating that paying some versus paying all subjects not only impacts the level of elicited risk aversion, but also the ordering of subjects' risk aversion to a meaningful degree. Interestingly, for all three risk measures, there is a higher rank correlation between the *AllHigh* and the *SomeHigh* conditions than between the *AllHigh* and *AllLow* conditions. Of the two conditions with lower expected value and cost to the researcher, paying some subjects better approximates the rank ordering of paying all subjects high stakes than does paying all subjects low stakes.

3.4 Correction Factor Using Multiple Measurements and Latent Risk Aversion

Our experiment contains multiple measures of an individual's risk aversion, with the *AllHigh* condition likely serving as the "best" measurement of true risk aversion. We estimate a simple structural model which depicts an individual's observed risk aversion measure in each treatment ($AllHigh_i$, $SomeHigh_i$, $AllLow_i$) as a function of an individual's unobserved latent risk aversion ($LatentRisk_i$) plus measurement error ϵ. Specifically, we simultaneously estimate:

$$\begin{aligned}
AllHigh_i &= LatentRisk_i + \epsilon_{1i} \\
SomeHigh_i &= \alpha_2 + \beta_2 LatentRisk_i + \epsilon_{2i} \\
AllLow_i &= \alpha_3 + \beta_3 LatentRisk_i + \epsilon_{3i}
\end{aligned}$$

We assume that the *AllHigh* is the purest measure of true risk aversion by setting this intercept to zero and the coefficient on *LatentRisk* to one. An individual's response in the *SomeHigh* and *AllHigh* condition is also a function of latent risk aversion, albeit a noisier measure than in the *AllHigh* condition, with α capturing any bias in levels and β capturing any slope shift to measured risk aversion from using one of the lower-cost elicitation methods. We estimate the coefficients to this system of equations using maximum likelihood, with the standard errors clustered by subject.

Table 7 presents these results. Across the three different risk measures, the coefficient on *LatentRisk* for the *SomeHigh* condition ranges from 0.72 to 0.77. The coefficient on *LatentRisk* in the *AllLow* condition is noticeably lower, ranging from 0.53 to 0.63, echoing our previous findings that the *SomeHigh* condition does a better job approximating the *AllHigh* results than does the *AllLow* condition. Overall, we find that a model of a common component of underlying risk aversion supports our results across the three treatments.

We use these results to construct a correction factor for researchers who can only afford to pay some subjects high stakes, but wish to approximate the results of the more expensive design in which all subjects are paid higher stakes. We first benchmark the prediction error from simply assuming that the results from paying only some subjects high stakes are equivalent to paying all subjects. For the MLE CRRA measure, treating an individual's response to the *SomeHigh* condition as the correct estimate for the *AllHigh* condition generates a root mean square error (a measure of model fit) of 0.278, relative to the sample mean of 0.532 for the *AllHigh* treatment. We next consider the model fit from observing an individual's choices in the *SomeHigh* condition, and using these results to predict the individual's choices in the *AllHigh* condition via the simple regression of $AllHigh_i = \alpha + \beta\ SomeHigh_i + \epsilon_i$. For MLE CRRA, this regression generates the prediction of $\widehat{AllHigh}_i = 0.217{+}0.676{*}SomeHigh_i$, and a root mean square error of 0.247, a slight improvement over the previous model. We now consider the prediction of $AllHigh$ using the results from the above structural model. Solving the second equation for the latent risk yields $LatentRisk_i = (SomeHigh_i - \alpha_2 - \epsilon_{2i})$ / β_2. Assuming that ϵ_{2i} has a mean of zero, and substituting into the first equation yields $\widehat{AllHigh}_i = (SomeHigh_i - \alpha_2)$ / β_2. Using the values from Table 7 of α_2 and β_2 for MLE CRRA yields $\widehat{AllHigh}_i = (SomeHigh_i - 0.100)$ / 0.773, which leads to a root mean square error of 0.147.[10] Thus, our correction generates almost a 50% reduction in the root mean square error relative to simply assuming that paying some subjects yields identical results to paying all subjects. Our correction also improves upon the correction generated from a simple regression of $AllHigh$ on $SomeHigh$ by accounting for the measurement error between these elicitations.

3.5 Tests for Order Effects

One concern with our identification is that our experiment relied upon a within-subject design. The identifying assumption of a within-subject design is that early treatments do not impact a

[10] Correspondingly, the improved correction to generate values as if all subjects had been paid high stakes for the Number of Safe Choices measure would be $\widehat{AllHigh}_i = (SomeHigh_i - 1.51)$ / 0.724. If all subjects had been paid stakes similar to our *AllLow* condition, the correction would be $\widehat{AllHigh}_i = (AllLow_i - 2.44)$ / 0.534 for the Number of Safe Choices and $\widehat{AllHigh}_i = (AllLow_i - 0.122)$ / 0.633 for MLE CRRA.

subject's responses in later treatments, otherwise a within-subject design may lead to spurious effects.[11] We varied the presentation order of treatments in our experiments, and test if responses varied by order condition. Experiment 1 rotated all possible orders of the three treatments across six order conditions, and Experiment 2 rotated its four treatments across four different order conditions, such that each treatment came first, second, third, or fourth exactly once. We first test if the levels of elicited risk aversion varied across all different order combinations in each experiment. For example, we tested whether the elicited MLE CRRA in the $AllHigh$ condition of Experiment 1 varied between {Order 1 and 2, Order 1 and 3, ... Order 4 and 6, and Order 5 and 6}; we then repeated these tests for the $SomeHigh$ and $AllLow$ conditions. For the MLE CRRA measure, in Experiment 1, 1 out of the 45 order combinations displayed a significant difference; in Experiment 2, 0 out of the 18 conditions were significant.[12] We next consider if order affected the differences in risk aversion between treatment conditions. For the MLE CRRA measure, we test if each treatment difference, ($\{AllHigh - SomeHigh\}$, $\{AllHigh - AllLow\}$, $\{SomeHigh - AllLow\}$), varied across the order combinations. For Experiment 1, 2 out of the 45 order combinations were significantly different, and for Experiment 2, 0 out of 18 order combinations were significantly different.

One concern with these order tests is that we had a small number of subjects in each order condition. For example, Experiment 1 had approximately 15 subjects in each order conditions, so we only have the power to rule out large differences by order. To address this concern, we pool all instances in which treatment i came immediately before j, and compare these instances against the other orders. For example, in Experiment 1, in both Order 1 ({*AllHigh, SomeHigh, AllLow*}) and Order 6 ({*AllLow, AllHigh, SomeHigh*} the *SomeHigh* condition came immediately after the *AllHigh* condition. We pool these two conditions, and test if the level of the *SomeHigh* condition varied when it immediately followed the *AllHigh* condition compared to when *SomeHigh* immediately preceded the *AllHigh* condition,

[11] See Charness et al. (2012) for a discussion of the benefits and drawbacks of between-subject versus within-subject experimental designs.

[12] For brevity's sake, we report the order effects solely for the MLE CRRA measure; identical conclusions follow from using the Number of Safe Choices or the Switch CRRA measures.

and compared to when these two treatments were separated temporally. We repeat this methodology for the other treatment conditions and their corresponding temporal placement relative to the alternative treatments. Pooling these instances yields comparisons with more than 30 subjects in each group, increasing our power to detect potential order differences. For the MLE CRRA, 0 of 18 of these pooled order comparisons were significantly different between orders.

In summary, we tested for multiple combinations of order effects, and find no evidence of any order contamination due to our within-subject design.

3.6 Effect Size

Table 4 reveals that paying only some subjects high stakes yields risk aversion estimates that are one-fifth of a standard deviation lower than the estimates from paying all subjects high stakes, where the standard deviation measures the between-subject heterogeneity in the *AllHigh* treatment. At first glance, an impact of one-fifth of a standard deviation is a small effect size. However, effects sizes on risk aversion may be bounded if risk-neutrality serves as a censor for how large a treatment effect can be. To put our effect size into perspective, we compare our effect size to the estimated differences between men and women in risk-taking. Byrnes et al. (1999) report a meta-analysis of 150 studies, representing 322 effects, on differences in risk-taking between men and women. This meta-analysis found a mean effect size of 0.13 with a 95% confidence interval of 0.12 to 0.14. Our demonstrated effect size of 0.19 from paying some subjects rather than all subjects is thus 50% larger than the gap in estimated risk-taking between men and women.

Our relatively small effect size may explain why some previous studies examining the effect of paying only subjects found no effect. Detecting an effect size of this magnitude would require a large number of subjects using between-subject identification, further justifying our within-subject design.

4 Comparison to Previous Work

We now compare our results to previous research examining the effect of probabilistically paying some subjects in studies measuring risk preferences. Baltussen et al. (2012) elicit risk preferences in an experimental task mimicking the Deal or No Deal game show, and investigate the impact of three different payment treatments. In their first-treatment, labelled the Basic treatment, each subject played Deal or No Deal once and was paid for the outcome. In the second treatment, the Within-Subject Random Incentive System, each subject played the game ten times, and one of the games' outcomes was randomly selected for payment. In the third treatment, the Between-Subject Random Incentive System, each subject played once, and one in ten subjects was paid the outcome. Each of the subjects was in only one of these three treatments. The authors find no significant difference in risk aversion between the first and second treatments, but do find less risk aversion in the third treatment, in which only one in ten subjects were paid. As the authors acknowledge, however, the face values of the gambles remained constant between treatments, and thus the stakes are equal in the first two treatments but lower in the third treatment.

Brokesova et al. (2017) elicit risk preferences in a laboratory setting, and compare these preferences to field-generated preferences in response to a bank's offering either a certain amount or a risky amount to newly opened savings accounts during a promotional campaign. Within the lab task, the authors compare the elicited risk preferences in a baseline task in which all subjects were paid for their choices to those in a treatment condition in which only one subject in each session received payment, and found no significant differences in risk preferences.

Beaud and Willinger (2015) measured risk preferences in a portfolio choice experiment. In their first experiment, subjects were endowed with wealth and completed a portfolio allocation task via pencil and paper, with 10% of subjects receiving a payment. In contrast, subjects in their computerized second experiment had to work to receive earnings, and then chose what fraction of their earnings to allocate to an investment. All subjects were paid for their choices in the second experiment.

Unlike our findings, Beaud and Willinger (2015) found no difference in elicited risk preferences between their "pay some" versus "pay all" subjects conditions. However, the main focus of Beaud and Willinger (2015) was a measure other than the effect of probabilistic payments on risk preferences. Risk preferences in their "pay some" condition were measured amidst conditions differing on several dimensions from their "pay all" condition, confounding the effect and limiting comparability to our study. The experimental manipulations in Baltussen et al. (2012) and Brokesova et al. (2017) do present a clean test of differences in risk preferences from paying subjects probabilistically; the manipulations in each of these experiments correspond exactly to our *AllHigh* and *SomeHigh* conditions. Brokesova et al. (2017) do not find a significant difference between their "pay all" versus "pay some subjects" conditions, in contrast to our findings. However, their "pay all subjects" condition had 56 subjects and their "pay some subjects" had 51 subjects, with between-subject identification. Given this, their identification could rule out only large effect sizes and their finding of no result is unsurprising. Baltussen et al. (2012) did find a significant difference between their "pay some" versus "pay all subjects" conditions, as do we. However, the face values of the gambles were identical between these two conditions, as in our *AllHigh* and *SomeHigh* condition. As such, the expected value of the gamble is lower in their "pay some" condition, and hence their finding of greater risk aversion in the "pay all" condition may simply be due to stake effects. Our study is the first study which adds the *AllLow* condition, which has equal expected value to the *SomeHigh* condition, and therefore controls for stake effects. Subjects are significantly more risk-averse in *SomeHigh* than in *AllLow*; this pattern of results suggests that subjects do not fully discount for the fact that only some subjects will receive payment.

5 Possible Explanations for Differences in Risk Aversion between Treatments

5.1 Individual Correlates

We next test if individual characteristics are associated with differences in risk aversion between the treatment conditions. For each of the three risk measures, we regress:

$$RiskMeasure_i^{Tx} - RiskMeasure_i^{Ty} = \alpha_0 + \boldsymbol{\alpha}_1 \mathbf{DEMOG}_i + \boldsymbol{\alpha}_2 \mathbf{RISK}_i + \epsilon_i$$

where i indexes individuals. $RiskMeasure$ denotes each of the three different risk measures {Number of Safe Choices, Switch CRRA, MLE CRRA} and T_x and T_y denote different treatment conditions. That is, our dependent variable is the difference in risk measures between the pairwise combinations for each of our three different risk measures. $\mathbf{DEMOG}_i$ is a set of demographic controls including subject age, and indicators for male subjects, White subjects, Asian subjects, Business majors, Social Science majors, and whether a subject is partly self-financing college through either loans or part-time employment. $\mathbf{RISK}_i$ is a set of alternate measures of risk preferences, comprised of categorical variables for frequencies of seat belt usage, driving over the speed limit, gambling, and the number of safe choices an individual made in each of the two Barsky et al. (1997) questions. Though occasionally a particular variable was significant for a particular specification, no variable was consistently significant across multiple specifications, and we conclude that the differences between our conditions are unexplained by the demographic and risk variables that we considered.[13]

5.2 Probability Weighting

We now examine if individuals' estimated risk aversion between the treatment conditions is associated with their estimated probability weighting. That is, we examine whether individuals

[13] Results available upon request.

who most over-weight small probabilities also display the greatest difference in risk aversion between the conditions with equivalent expected values, *SomeHigh* and *AllLow*.

To estimate probability weighting, we utilize the one-parameter probability weighting function popularized by Prelec (1998).[14] In the expected utility calculation of each lottery, p_n is now replaced by $w(p_n) = \exp(-(-\ln(p_n))^{\gamma})$. We again assume that individuals choose between lottery A versus B based upon the difference in the expected utilities, with the expected utilities now a function of the probability weighting parameter γ in addition to risk aversion r. We jointly estimate via maximum likelihood the parameters r and γ which best explain each individual's observed responses. Intuitively, the probability weighting parameter γ is identified by any difference in an individual's responses to *AllHigh* versus *AllLowProb*. *AllLowProb* is equivalent to *AllHigh* but with all probabilities multiplied by the common factor of 1/8. If an individual switches from A to B at the same row in *AllHigh* and *AllLowProb*, then risk aversion r is identified by the switching row in *AllHigh* and the probability weighting parameter γ would be 1. Conversely, if an individual selects a different row for *AllLowProb* than for *AllHigh*, then the parameter γ will differ from 1.

Figure 3 presents the distribution of our estimated values of probability weighting parameter γ. 85% of our sample have a γ less than 1, consistent with previous findings of systematic bias towards over-weighting small probabilities, suggesting that our measure of γ is not simply capturing noise or response error between the *AllHigh* and *AllLowProb* conditions. We find a mean value of 0.53 for γ, which comports with previous estimates from the literature.[15]

We first consider the simple association between the difference in risk aversion between the *SomeHigh* and *AllLow* condition, and our estimated probability weighting parameter γ. We next add the set of demographics and risk controls used above, and in a separate specification

[14] There are several one- and two-parameter versions of functional forms to model probability weighting. We employed a one-parameter function to reduce the amount of questions necessary for identification and reduce survey fatigue. Stott (2006) compared various combinations of one- and two-parameter probability weighting functions paired with different value functions. The most predictive model had power utility, as we use here, and the single parameter Prelec (1998) probability weighting function.

[15] Of papers employing the same one-parameter probability weighting function as we used, Wu and Gonzalez (1996) estimated a mean of 0.74 and Bleichrodt and Pinto (2000) estimated a mean of 0.53.

we add the risk aversion measure from *AllHigh*, to control for the interaction of probability weighting and risk aversion. In our fullest specification, we add the demographic and risk controls as well as the *AllHigh* risk aversion measure. In no specification was the coefficient on the probability weighting parameter γ ever statistically significant.[16] Our findings are thus similar to Barseghyan et al. (2011), who examine the concordance of risk preferences between individuals' choices of home and automobile deductibles, and find that probability weighting cannot explain individuals' different risk tolerance between the two.

6 Conclusion

We examined the impact on elicited risk preferences of the relatively common technique of paying only some subjects for their choices, as compared to paying all subjects for their choices. We elicited subjects' risk preferences in three conditions: a high-stakes condition in which all subjects were paid; a high-stakes condition in which only one out of eight subjects were paid; and a low-stakes condition in which all subjects were paid. This lower stakes condition had an expected value equal to one-eighth of the high condition, enabling us to examine if any change in risk preferences between the conditions is simply due to a stakes effect. Subjects were significantly more risk-averse in the high-stakes condition in which all subjects were paid compared to when subjects were paid probabilistically. However, subjects were also significantly more risk-averse when subjects probabilistically received higher stakes than when all subjects received lower stakes, despite the equality of stakes. Our results demonstrate that the practice of paying subjects probabilistically affects the level of elicited risk aversion. Moreover, paying some versus paying all subjects also affects the relative ordering of risk aversion by subjects; the most risk-averse subjects in the condition in which all subjects are paid are not necessarily the most risk-averse subjects under a probabilistic payment mechanism. For researchers who cannot feasibly afford to pay all subjects high stakes, we provide a correction factor for elicitations that pay only some subjects to approximate the risk levels of elicitations

[16] Results available upon request.

that did pay all subjects.

We elicited subjects' probability weighting parameters to test if subjects who most overweight small probabilities displayed the largest differences between our "pay all" versus "pay some subjects" conditions. Probability weighting was not significantly related to the differences in risk preferences between these conditions. Standard experimental demographics, as well as alternative measures of risk preferences, were also not reliably predictive of differences between conditions.

Our experiments were a mixed payment design; in every treatment subjects were paid for a randomly selected question, and then in some treatments only a randomly selected subject was paid. Future work could examine if our results hold if all questions count for payment. We explored the impact on risk preferences from paying subjects probabilistically, as risk preferences are the most directly relevant and affected preference from probabilistic payment. Future work could also examine if our findings regarding probabilistic payment by subjects extends to other preferences as well, such as time preferences or social preferences. We tested the effect of paying all subjects higher stakes versus one out of eight subjects receiving payments; future work could explore whether our finding of differing risk aversion holds across other levels of probabilistic subject payment.

Our results reveal that subjects discount, albeit not fully, probabilistic payment mechanisms. Our *SomeHigh* and *AllLow* condition had identical expected values to subjects, and therefore nearly identical costs to the researcher, but our *SomeHigh* condition had closer correlation, in terms of both elicited absolute levels of risk aversion as well as the rank ordering of risk aversion, with the *AllHigh* condition than did the *AllLow* condition with the *AllHigh* condition. For budget-constrained researchers who cannot feasibly afford to pay all subjects high stakes, paying some subjects probabilistically for higher stakes elicits risk preferences that more closely approximate the ideal condition of paying all subjects high stakes than does paying all subjects lower stakes.

References

Abdellaoui, Mohammed, Peter Klibanoff, Laetitia Placido. "Experiments on compound risk in relation to simple risk and to ambiguity," *Management Science*, 61 (6) (2015), pp. 1306-1322.

Baltussen, Guido, Thierry Post, Martin J. van den Assem, and Peter P. Wakker. "Random Incentive Systems in a Dynamic Choice Experiment," *Experimental Economics* (2012) 15: 418-443. https://doi.org/10.1007/s10683-011-9306-4

Bar-Hillel, Maya. "On the Subjective Probability of Compound Events," *Organizational Behavior and Human Performance*, Volume 9, Issue 3,1973, Pages 396-406.

Barseghyan, L., J. Prince, and J. C. Teitelbaum (2011): "Are Risk Preferences Stable Across Contexts? Evidence from Insurance Data," *American Economic Review*, 101(2), 591–631.

Barsky, R.B., Juster, F.T., Kimball, M.S., Shapiro, M.D. (1997), "Preference Parameters and Behavioral Heterogeneity: an Experimental Approach in the Health and Retirement Study", *The Quarterly Journal of Economics* 112, 729-758.

Beaud, M., Willinger, M. "Are people risk vulnerable?" *Management Science* 61 (2015), 624–636.

Bernasconi, M. & Loomes, G. "Failures of the Reduction Principle in an Ellsburg-Type Problem," *Theory and Decision* (1992) 32: 77. https://doi.org/10.1007/BF00133629

Binswanger, H. P. (1981): "Attitudes toward Risk: Theoretical Implications of an Experiment in Rural India," *Economic Journal*, 91, 867-890.

Bleichrodt, Hans and Jose Luis Pinto (2000). "A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis," *Management Science*, 46(11), 1485-1496.

Brokesova, Zuzana, Cary Deck, Jana Peliova. "Comparing a Risky Choice in the Field and Across Lab Procedures," *Journal of Economic Psychology*, Volume 61, 2017, Pages 203-212.

Byrnes, James P., David C. Miller, William D. Schafer. "Gender Differences in Risk Taking: A Meta-Analysis," *Psychological Bulletin*, Vol 125(3), May 1999, 367-383.

Charness, Gary, Uri Gneezy, and Brianna Halladay. "Experimental Methods: Pay One or Pay All," *Journal of Economic Behavior & Organization* 131 (2016) 141–150.

Charness, Gary, Uri Gneezy, and Michael A. Kuhn. "Experimental Methods: Between-subject and Within-subject Design," *Journal of Economic Behavior & Organization* 2012, vol. 81, issue 1, 1-8.

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv. 2007. "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review*, 97 (5): 1921-1938.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences," *Journal of Economic Literature*, 47 (2): 448-74.

Cubitt, R.P., Starmer, C. & Sugden, R. "On the Validity of the Random Lottery Incentive System," *Experimental Economics* (1998) 1: 115. https://doi.org/10.1007/BF01669298.

Fehr-Duda, Helga, Bruhin, Adrian, Epper, Thomas and Schubert, Renate, (2010). "Rationality on the Rise: Why Relative Risk Aversion Increases with Stake Size," *Journal of Risk and Uncertainty*, 40, issue 2, p. 147-180.

Hajimoladarvish, Narges. "How do people reduce compound lotteries?," *Journal of Behavioral and Experimental Economics*, Volume 75, 2018, Pages 126-133, ISSN 2214-8043, https://doi.org/10.1016/j.socec.2018.06.002.

Harrison, Glenn W., Jimmy Martínez-Correa, J. Todd Swarthout. "Reduction of compound lotteries with objective probabilities: Theory and evidence," *Journal of Economic Behavior & Organization*, Volume 119, 2015, Pages 32-55, ISSN 0167-2681, https://doi.org/10.1016/j.jebo.2015.07.012.

Harrison, Glenn, List, John & Towe, Charles (2007). "Naturally Occurring Preferences and Exogenous Laboratory Experiments: A Case Study of Risk Aversion." *Econometrica*. 75. 433-458. 10.2139/ssrn.698541

Holt, C. and S. Laury: Risk Aversion and Incentive Effects, *American Economic Review*, 92(5), (2002).

Kahneman, Daniel and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, Vol. 47, No. 2 (Mar., 1979), pp. 263-292.

Kreps, David M., and Evan L. Porteus. "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica* 46, no. 1 (1978): 185-200. doi:10.2307/1913656.

Laury, Susan, (2006), Pay One or Pay All: Random Selection of One Choice for Payment, No 2006-24, Experimental Economics Center Working Paper Series, Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University, https://EconPapers.repec.org/RePEc:exc:wpaper:2006-24.

Miao, Bin and Zhong, Songfa. "An Experimental Study of Attitude towards Compound Lottery" (November 16, 2012). Available at SSRN: https://ssrn.com/abstract=2039216 or http://dx.doi.org/10.2139/ssrn.2039216

Prelec, Drazen. "The Probability Weighting Function." *Econometrica*, vol. 66, no. 3, 1998, pp. 497–527. *JSTOR*, JSTOR, www.jstor.org/stable/2998573.

Segal, Uzi. "Two-Stage Lotteries without the Reduction Axiom." *Econometrica* 58, no. 2 (1990): 349-77. doi:10.2307/2938207.

Starmer, C., Sugden, R., 1991. "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation," *American Economic Review* 81, 971–978.

Stott, Henry P. "Cumulative Prospect Theory's Functional Menagerie," *Journal of Risk and Uncertainty* (2006) 32: 101.

Wu. George and Richard Gonzalez. (1996). "Curvature of the Probability Weighting Function," *Management Science*, 42(12), 1676-1690.

Table 1 – Percent of Subjects Choosing Safe Option at each Decision Row

		Percent of Subjects Choosing the Safe Option							
		All Subjects				Valid MLE			
Decision	CRRA if Indifferent	*AllHigh*	*SomeHigh*	*AllLow*	*AllLowProb*	*AllHigh*	*SomeHigh*	*AllLow*	*AllLowProb*
1	-1.71	99.5	95.8	99.0	98.0	100	100	100	98.8
2	-0.95	97.9	95.3	98.4	95.8	100	100	100	96.4
3	-0.49	97.9	93.2	95.8	93.8	100	100	97.7	94.0
4	-0.14	93.8	90.1	91.1	88.5	96.5	96.0	93.1	91.7
5	0.15	79.2	73.4	75.5	78.1	83.2	78.0	78.6	79.8
6	0.41	67.7	59.9	56.8	63.5	68.8	62.4	59.5	66.7
7	0.68	46.4	38.5	28.6	56.3	48.6	39.9	28.9	58.3
8	0.97	23.4	17.7	11.5	33.3	23.1	18.5	11.0	36.9
9	1.37	10.4	7.3	6.8	22.9	10.4	5.8	6.3	23.8
10	N/A	1.6	1.0	0	2.1	0	0	0	0
Observations		192	192	192	96	173	173	173	84

Table 2 – Summary Statistics for Risk Measures By Experiment 1 and 2

	All Subjects				Valid MLE			
	AllHigh	*SomeHigh*	*AllLow*	*AllLowProb*	*AllHigh*	*SomeHigh*	*AllLow*	*AllLowProb*
Num Safe Choices								
Experiment 1	6.33□	5.87□	5.72□	–□	6.48□	6.01□	5.83□	–□
	(1.55)	(1.54)	(1.54)		(1.43)	(1.44)	(1.39)	
Experiment 2	6.02□	5.60□	5.55□	6.31□	6.12□	6.00□	5.67□	6.43□
	(1.73)	(2.14)	(1.65)	(2.04)	(1.73)	(1.64)	(1.65)	(2.04)
Rank-Sum p-value	.131□	.555□	.382□	–□	.168□	.819□	.451□	–□
Observations	189□	190□	192□	94□	173□	173□	173□	84□
Switch CRRA								
Experiment 1	.643□	.517□	.465□	–□	.688□	.557□	.501□	–□
	(.448)	(.443)	(.442)		(.406)	(.412)	(.408)	
Experiment 2	.561□	.441□	.411□	.614□	.590□	.555□	.448□	.659□
	(.493)	(.583)	(.474)	(.578)	(.491)	(.461)	(.478)	(.576)
Rank-Sum p-value	.196□	.413□	.276□	–□	.187□	.782□	.389□	–□
Observations	189□	190□	192□		173□	173□	173□	84□
MLE CRRA								
Experiment 1	.583□	.492□	.477□	–□	.582□	.471□	.439□	–□
	(.319)	(.366)	(.367)		(.294)	(.326)	(.318)	
Experiment 2	.512□	.499□	.404□	.570□	.479□	.461□	.405□	.573□
	(.401)	(.400)	(.408)	(.426)	(.373)	(.358)	(.382)	(.390)
Rank-Sum p-value	.232□	.953□	.216□	–□	.104□	.807□	.546□	–□
Observations	188□	181□	189□	96□	173□	173□	173□	84□

Standard deviations in parentheses.

Table 3 – Summary Statistics for Number of Switches By Experiment 1 and 2

	All Subjects				Valid MLE			
	AllHigh	*SomeHigh*	*AllLow*	*ALP*	*AllHigh*	*SomeHigh*	*AllLow*	*ALP*
Number of Switches								
Experiment 1	1.15□ (.598)	1.18□ (.725)	1.14□ (.573)	–□	1.09□ (.417)	1.09□ (.514)	1.04□ (.298)	–□
Experiment 2	1.16□ (.744)	1.16□ (.886)	1.18□ (.665)	1.19□ (.799)	1.04□ (.243)	1.07□ (.486)	1.09□ (.428)	1.07□ (.460)
Rank-Sum p-value	.793□	.246□	.754□	–□	.812□	.910□	.429□	–□
	AllHigh	*SomeHigh*	*AllLow*	*ALP*	*AllHigh*	*SomeHigh*	*AllLow*	*ALP*
Experiment 1 and 2 (Pooled)	1.15□ (.673)	1.17□ (.808)	1.16□ (.620)	1.19□ (.799)	1.06□ (.344)	1.08□ (.499)	1.07□ (.367)	1.07□ (.460)
	AH vs *SH*	*AH* vs *AL*	*SH* vs *AL*		*AH* vs *SH*	*AH* vs *AL*	*SH* vs *AL*	
Rank-Sum p-value	.415□	.814□	.530□		.438□	.703□	.368□	
Observations	192	192	192	96	173	173	173	84

AH denotes *AllHigh*; *SH* denotes *SomeHigh*; *AL* denotes *AllLow*; and *ALP* denotes *AllLowProb*. Standard deviations in parentheses.

Table 4 – Differing Risk Preferences by Treatment

	AllHigh	*SomeHigh*	*AllLow*
Num Safe Choices	6.31□	6.01□	5.75□
	(1.59)	(1.53)	(1.52)
	AH vs *SH*	*AH* vs *AL*	*SH* vs *AL*
Signed-Rank p-value	0.0008□	0.0000□	0.0064□
	AllHigh	*SomeHigh*	*AllLow*
Switch CRRA	.640□	.556□	.475□
	(.451)	(.435)	(.436)
	AH vs *SH*	*AH* vs *AL*	*SH* vs *AL*
Signed-Rank p-value	0.0019□	0.0000□	0.0029□
	AllHigh	*SomeHigh*	*AllLow*
MLE CRRA	.532□	.467□	.422□
	(.337)	(.341)	(.350)
	AH vs *SH*	*AH* vs *AL*	*SH* vs *AL*
Signed-Rank p-value	0.0011□	0.0000□	0.028□
Observations	173	173	173

AH denotes *AllHigh*; *SH* denotes *SomeHigh*; *AL* denotes *AllLow*; and *ALP* denotes *AllLowProb*. Standard deviations in parentheses.

Table 5 – Differing Risk Preferences to *AllLowProb* in Experiment 2

	AllHigh	*SomeHigh*	*AllLow*	*AllLowProb*
Num Safe Choices	6.12□	6.00□	5.67□	6.43□
	(1.73)	(1.64)	(1.65)	(2.04)
	AH vs ALP	SH vs ALP	AL vs ALP	
Signed-Rank p-value	0.140□	.050□	0.0001□	
	AllHigh	*SomeHigh*	*AllLow*	*AllLowProb*
Switch CRRA	.590□	.555□	.448□	.659□
	(.491)	(.461)	(.478)	(.576)
	AH vs ALP	SH vs ALP	AL vs ALP	
Signed-Rank p-value	0.223□	.0678□	0.0003□	
	AllHigh	*SomeHigh*	*AllLow*	*AllLowProb*
MLE CRRA	.479□	.461□	.405□	.573□
	(.373)	(.358)	(.382)	(.390)
	AH vs ALP	SH vs ALP	AL vs ALP	
Signed-Rank p-value	0.0328□	0.0077□	0.0001□	
Observations	84□	84□	84□	84□

AH denotes *AllHigh*; SH denotes *SomeHigh*; AL denotes *AllLow*;and ALP denotes *AllLowProb.* Standard deviations in parentheses.

Table 6 – Rank Correlations of Risk Measures Across Treatments

Num Safe Choices			
	AllHigh	*SomeHigh*	*AllLow*
AllHigh	1.00	–	–
SomeHigh	0.76	1.00	–
AllLow	0.56	0.55	1.00

Switch CRRA			
	AllHigh	*SomeHigh*	*AllLow*
AllHigh	1.00	–	–
SomeHigh	0.75	1.00	–
AllLow	0.55	0.54	1.00

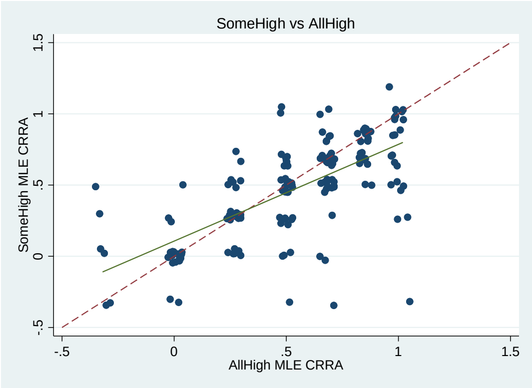
MLE CRRA			
	AllHigh	*SomeHigh*	*AllLow*
AllHigh	1.00	–	–
SomeHigh	0.69	1.00	–
AllLow	0.56	0.56	1.00

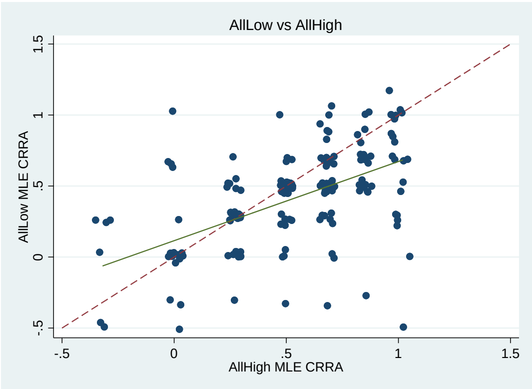
Table 7 – Structural Equation Model of Latent Risk Aversion

	Num Safe Choices	Switch CRRA	MLE CRRA
AllHigh			
Constant	0□	0□	0□
	–	–	–
LatentRisk□	1	1□	1□
	–	–	–
SomeHigh			
Constant	1.51□	.101□	.100□
	(.310)	(.036)	(.036)
LatentRisk□	.724□	.779□	.773□
	(.049)	(.053)	(.056)
AllLow			
Constant	2.44□	.141□	.122□
	(.471)	(.054)	(.047)
LatentRisk□	.534□	.572□	.633□
	(.075)	(.076)	(.082)

Standard Errors in parentheses.

Figure 1: Scatterplots of MLE CRRA by Treatment





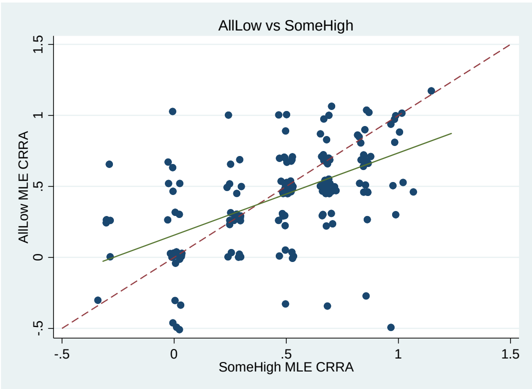


Figure 2: Distribution of MLE CRRA by Treatment

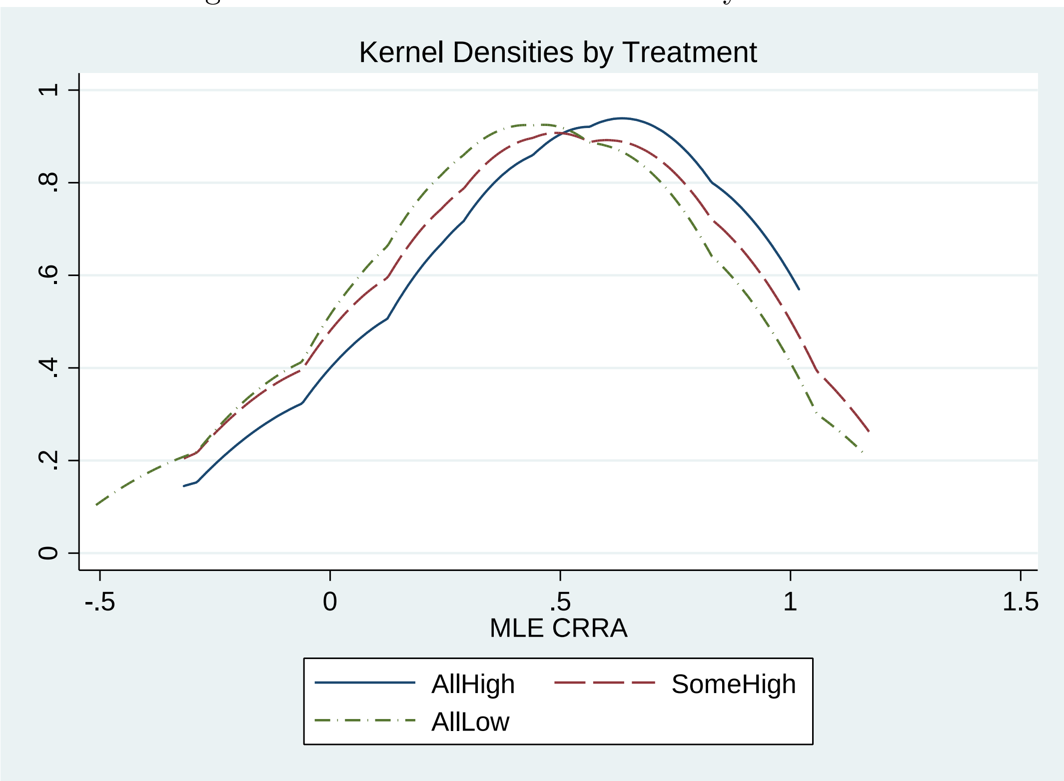
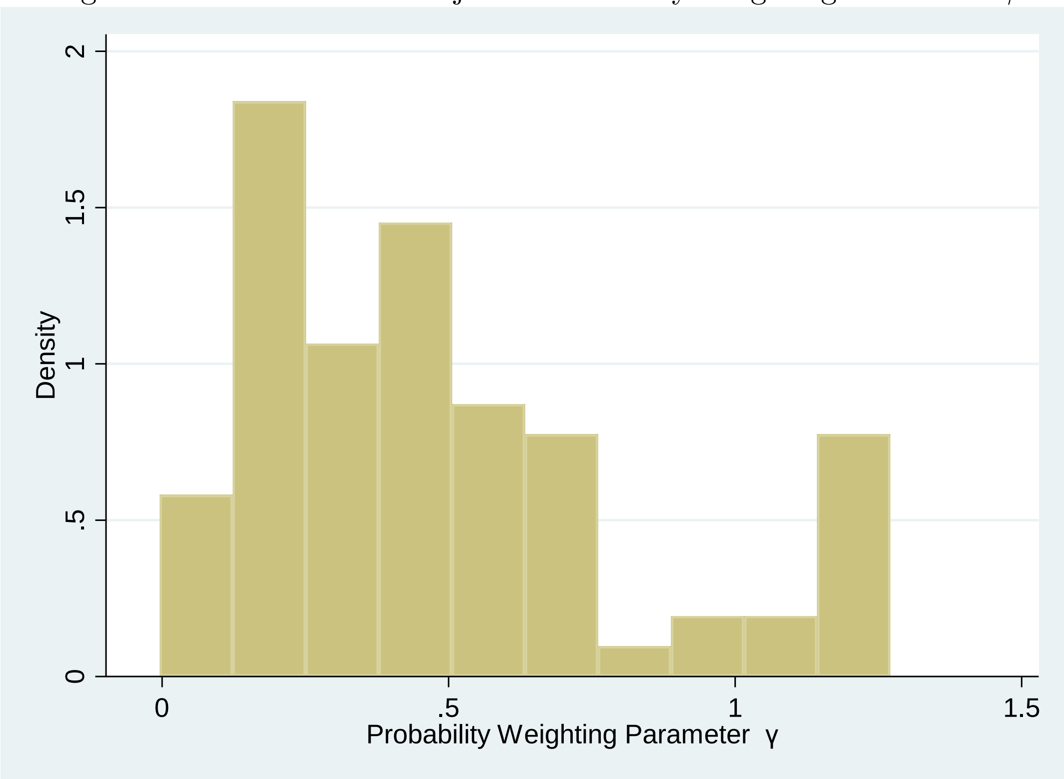


Figure 3: Distribution of Subjects' Probability Weighting Parameter γ



APPENDIX A: EXPERIMENTAL INSTRUMENT

INSTRUCTIONS (for Experiment 1)

You will be making choices between two lotteries, such as those represented as "Option A" and "Option B" below. Note that the actual payoffs amounts for your decisions will differ from those listed in these instructions. The money prizes are determined by throwing a ten-sided die. Each outcome, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, is equally likely. Thus if you choose Option A, you will have a 1 in 10 chance of earning $2.00 and a 9 in 10 chance of earning $1.60. Similarly, Option B offers a 1 in 10 chance of earning $3.85 and a 9 in 10 chance of earning $0.10.

Decision	**Option A**	**Option B**	**Your Choice** **Circle One**
1st	$2.00 if the die is 1 $1.60 if the die is 2 - 10	$3.85 if the die is 1 $0.10 if the die is 2 - 10	**A** or **B**

Each row of the decision table contains a pair of choices between **Option A** and **Option B**.

You make your choice by circling either "A" or "B" in the far right hand column of the table. Only one option in each row (i.e. for each Decision) can be circled.

Decision	**Option A**	**Option B**	**Your Choice** **Circle One**
1st	$2.00 if the die is 1 $1.60 if the die is 2 - 10	$3.85 if the die is 1 $0.10 if the die is 2 - 10	**A** or **B**
2nd . .	$2.00 if the die is 1 - 2 $1.60 if the die is 3 - 10	$3.85 if the die is 1 - 2 $0.10 if the die is 3 - 10	**A** or **B**

Even though you will make ten decisions, **only one** of these will end up being used. The selection of the one to be used depends on the throw of a ten-sided die. No decision is any more likely to be used than any other, and you will not know in advance which one will be selected, so please think about each one carefully. The first throw of the ten-sided die fixes the row (i.e. the Decision) that will be used to determine your earnings. For example, suppose that you make all ten decisions and the throw of the die is 9, then your choice, A or B, for decision 9 below would be used and the other decisions would not be used.

Decision	Option A	Option B	Your Choice Circle One
. . 9th	\$2.00 if the die is 1 - 9 \$1.60 if the die is 10	\$3.85 if the die is 1 - 9 \$0.10 if the die is 10	**A** or **B**

After the random die throw fixes the Decision row that will be used, we need to make a second die throw to determine the earnings for the Option you chose for that row. In Decision 9 below, for example, a throw of 1, 2, 3, 4, 5, 6, 7, 8, or 9 will result in the higher payoff for the option you chose, and a throw of 10 will result in the lower payoff.

Decision	Option A	Option B	Your Choice
9th	\$2.00 if the die is 1 - 9 \$1.60 if the die is 10	\$3.85 if the die is 1 - 9 \$0.10 if the die is 10	**A** or **B**
10th	\$2.00 if the die is 1 - 10	\$3.85 if the die is 1 - 10	**A** or **B**

For decision 10, the random die throw will not be needed, since the choice is between amounts of money that are fixed: \$2.00 for Option A and \$3.85 for Option B.

Making Ten Decisions: At the end of these instructions you will see tables with 10 decisions in 10 separate rows, and you choose by circling one choice (A or B) in the far right hand column for each of the 10 rows. You may make these choices in any order.

The Relevant Decision: One of the 10 rows (i.e. Decisions) is then selected at random, and the Option (A or B) that you chose in that row will be used to determine your earnings. Note: Please think about each decision carefully, since each row is equally likely to end up being the one that is used to determine payoffs.

Determining the Payoff for Each Round: After one of the decisions has been randomly selected, we will throw the ten-sided die a second time. The number is equally likely to be 1, 2, 3, ... 10. This number determines your earnings for the Option (A or B) that you previously selected for the decision being used.

Determining Who Gets Paid: In some cases, there will be a third die throw to determine which person in the room will be paid for the set of decisions on a particular sheet. The top of each decision sheet explains who will be paid for that particular decision sheet.

Determining the Final Payoff: There will be 3 decision sheets, each with 10 rows. You will find out your earnings for each of these 3 sheets after you have made all of your decisions today.

Instructions Summary

To summarize, you will indicate an option, A or B, for each of the rows by circling one choice in the far right hand column.

Then the throw of a ten-sided die fixes which row of the table (i.e. which Decision) is relevant for your earnings.

In that row, your decision fixed the choice for that row, Option A or Option B, and a final throw of the ten-sided die will determine the money payoff for the decision you made.

In addition, in some cases, there will be a third die throw to determine which person in the room will be paid for the set of decisions on a particular sheet. The top of each decision sheet explains who will be paid for that particular decision sheet.

This whole process will be repeated, but the prize amounts may change from one sheet to the next, so look at the prize amounts carefully before you start making decisions.

INSTRUCTIONS (for Experiment 2)

ID Number:

You will be making choices between two lotteries, such as those represented as "Option A" and "Option B" below. Note that the actual payoffs amounts for your decisions will differ from those listed in these instructions. The money prizes are determined by the computer equivalent of throwing a ten-sided die. Each outcome, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, is equally likely. If you choose Option A in the row shown below, you will have a 1 in 10 chance of earning $2.00 and a 9 in 10 chance of earning $1.60. Similarly, Option B offers a 1 in 10 chance of earning $3.85 and a 9 in 10 chance of earning $0.10.

Decision	**Option A**	**Option B**	**Your Choice Circle One**
1st	$2.00 if the die is 1 $1.60 if the die is 2 - 10	$3.85 if the die is 1 $0.10 if the die is 2 - 10	**A** or **B**

Each row of the decision table contains a pair of choices between **Option A** and **Option B**.

You make your choice by circling either "A" or "B" in the far right hand column of the table. Only one option in each row (i.e. for each Decision) can be circled.

Decision	**Option A**	**Option B**	**Your Choice Circle One**
1st	$2.00 if the die is 1 $1.60 if the die is 2 - 10	$3.85 if the die is 1 $0.10 if the die is 2 - 10	**A** or **B**
2nd . .	$2.00 if the die is 1 - 2 $1.60 if the die is 3 - 10	$3.85 if the die is 1 - 2 $0.10 if the die is 3 - 10	**A** or **B**

Even though you will make ten decisions, **only one** of these will end up being used. The selection of the one to be used depends on the "throw of the die" that is the determined by a random number generator. No decision is any more likely to be used than any other, and you will not know in advance which one will be selected, so please think about each one carefully. This random selection of a decision fixes the row (i.e. the Decision) that will be used.

For example, suppose that you make all ten decisions and the random number is 9, then your choice, A or B, for decision 9 below would be used and the other decisions would not be used.

Decision	Option A	Option B	Your Choice Circle One
. . 9th	$2.00 if the die is 1 - 9 $1.60 if the die is 10	$3.85 if the die is 1 - 9 $0.10 if the die is 10	**A** or **B**

After the random number generator fixes the Decision row that will be used, we need to generate a second random number to determine the earnings for the Option you chose for that row. In Decision 9 below, for example, a throw of 1, 2, 3, 4, 5, 6, 7, 8, or 9 will result in the higher payoff for the option you chose, and a throw of 10 will result in the lower payoff.

Decision	Option A	Option B	Your Choice
9th	$2.00 if the die is 1 - 9 $1.60 if the die is 10	$3.85 if the die is 1 - 9 $0.10 if the die is 10	**A** or **B**
10th	$2.00 if the die is 1 - 10	$3.85 if the die is 1 - 10	**A** or **B**

In some cases, such as decision 10 above, the random number generator will not be needed, since the choice is between amounts of money that are fixed: $2.00 for Option A and $3.85 for Option B.

Making Ten Decisions: At the end of these instructions you will see tables with 10 decisions in 10 separate rows, and you choose by circling one choice (A or B) in the far right hand column for each of the 10 rows. You may make these choices in any order.

The Relevant Decision: One of the rows is then selected at random, and the Option (A or B) that you chose in that row will be used to determine your earnings. Note: Please think about each decision carefully, since each row is equally likely to end up being the one that is used to determine payoffs.

Determining the Payoff for Each Round: After one of the decisions has been randomly selected, we will generate another random number to determine your earnings for the Option (A or B) that you previously selected for the decision being used.

Determining Who Gets Paid: In some cases, there will be a die throw to determine which person in the room will be paid for the set of decisions on a particular sheet. The top of each decision sheet explains who will be paid for that particular decision sheet.

Determining the Final Payoff: There will be 4 decision sheets, each with 10 rows. You will find out your earnings for each of these 4 sheets after you have made all of your decisions today.

Instructions Summary

To summarize, you will indicate an option, A or B, for each of the rows by circling one choice in the far right hand column.

Then a random number fixes which row of the table (i.e. which decision) is relevant for your earnings.

In that row, your decision fixed the choice for that row, Option A or Option B, and a final random number will determine the money payoff for the decision you made.

In addition, in some cases, there will be a die throw to determine which person in the room will be paid for the set of decisions on a particular sheet. The top of each decision sheet explains who will be paid for that particular decision sheet.

This whole process will be repeated, but the prize amounts may change from one sheet to the next, so look at the prize amounts carefully before you start making decisions.

APPENDIX Table 1: *AllHigh* Condition

<u>EVERYONE IN THE ROOM</u> WILL BE PAID FOR 1 OF THE 10 DECISIONS ON THIS SHEET.

Decision	Option A	Option B	Your Decision Circle One
1	$32.00 if the die is 1 $25.60 if the die is 2-10	$61.60 if the die is 1 $1.60 if the die is 2-10	**A** or **B**
2	$32.00 if the die is 1 -2 $25.60 if the die is 3-10	$61.60 if the die is 1-2 $1.60 if the die is 3-10	**A** or **B**
3	$32.00 if the die is 1-3 $25.60 if the die is 4-10	$61.60 if the die is 1-3 $1.60 if the die is 4-10	**A** or **B**
4	$32.00 if the die is 1-4 $25.60 if the die is 5-10	$61.60 if the die is 1-4 $1.60 if the die is 5-10	**A** or **B**
5	$32.00 if the die is 1-5 $25.60 if the die is 6-10	$61.60 if the die is 1-5 $1.60 if the die is 6-10	**A** or **B**
6	$32.00 if the die is 1-6 $25.60 if the die is 7-10	$61.60 if the die is 1-6 $1.60 if the die is 7-10	**A** or **B**
7	$32.00 if the die is 1-7 $25.60 if the die is 8-10	$61.60 if the die is 1-7 $1.60 if the die is 8-10	**A** or **B**
8	$32.00 if the die is 1-8 $25.60 if the die is 9-10	$61.60 if the die is 1-8 $1.60 if the die is 9-10	**A** or **B**
9	$32.00 if the die is 1-9 $25.60 if the die is 10	$61.60 if the die is 1-9 $1.60 if the die is 10	**A** or **B**
10	$32.00 if the die is 1-10	$61.60 if the die is 1-10	**A** or **B**

Result of first random number generated (to determine Decision): ___________

Result of second random number generated (to determine Payoff): ___________

Payoff: ___________

APPENDIX Table 2: *SomeHigh* Condition

<u>ONE PERSON OUT OF THE 8 PEOPLE IN THE ROOM</u> WILL BE PAID FOR 1 OF THE 10 DECISIONS ON THIS SHEET.

Decision	Option A	Option B	Your Decision Circle One
1	$32.00 if the die is 1 $25.60 if the die is 2-10	$61.60 if the die is 1 $1.60 if the die is 2-10	**A** or **B**
2	$32.00 if the die is 1 -2 $25.60 if the die is 3-10	$61.60 if the die is 1-2 $1.60 if the die is 3-10	**A** or **B**
3	$32.00 if the die is 1-3 $25.60 if the die is 4-10	$61.60 if the die is 1-3 $1.60 if the die is 4-10	**A** or **B**
4	$32.00 if the die is 1-4 $25.60 if the die is 5-10	$61.60 if the die is 1-4 $1.60 if the die is 5-10	**A** or **B**
5	$32.00 if the die is 1-5 $25.60 if the die is 6-10	$61.60 if the die is 1-5 $1.60 if the die is 6-10	**A** or **B**
6	$32.00 if the die is 1-6 $25.60 if the die is 7-10	$61.60 if the die is 1-6 $1.60 if the die is 7-10	**A** or **B**
7	$32.00 if the die is 1-7 $25.60 if the die is 8-10	$61.60 if the die is 1-7 $1.60 if the die is 8-10	**A** or **B**
8	$32.00 if the die is 1-8 $25.60 if the die is 9-10	$61.60 if the die is 1-8 $1.60 if the die is 9-10	**A** or **B**
9	$32.00 if the die is 1-9 $25.60 if the die is 10	$61.60 if the die is 1-9 $1.60 if the die is 10	**A** or **B**
10	$32.00 if the die is 1-10	$61.60 if the die is 1-10	**A** or **B**

Result of first random number generated (to determine Decision): ___________

Result of second random number generated (to determine Payoff): ___________

Result of die throw (to determine which person in the room will be paid): _________

Payoff: ___________

APPENDIX Table 3: *AllLow* Condition

<u>EVERYONE IN THE ROOM</u> WILL BE PAID FOR 1 OF THE 10 DECISIONS ON THIS SHEET.

Decision	Option A	Option B	Your Decision Circle One
1	$4.00 if the die is 1 $3.20 if the die is 2-10	$7.70 if the die is 1 $0.20 if the die is 2-10	**A** or **B**
2	$4.00 if the die is 1 -2 $3.20 if the die is 3-10	$7.70 if the die is 1-2 $0.20 if the die is 3-10	**A** or **B**
3	$4.00 if the die is 1-3 $3.20 if the die is 4-10	$7.70 if the die is 1-3 $0.20 if the die is 4-10	**A** or **B**
4	$4.00 if the die is 1-4 $3.20 if the die is 5-10	$7.70 if the die is 1-4 $0.20 if the die is 5-10	**A** or **B**
5	$4.00 if the die is 1-5 $3.20 if the die is 6-10	$7.70 if the die is 1-5 $0.20 if the die is 6-10	**A** or **B**
6	$4.00 if the die is 1-6 $3.20 if the die is 7-10	$7.70 if the die is 1-6 $0.20 if the die is 7-10	**A** or **B**
7	$4.00 if the die is 1-7 $3.20 if the die is 8-10	$7.70 if the die is 1-7 $0.20 if the die is 8-10	**A** or **B**
8	$4.00 if the die is 1-8 $3.20 if the die is 9-10	$7.70 if the die is 1-8 $0.20 if the die is 9-10	**A** or **B**
9	$4.00 if the die is 1-9 $3.20 if the die is 10	$7.70 if the die is 1-9 $0.20 if the die is 10	**A** or **B**
10	$4.00 if the die is 1-10	$7.70 if the die is 1-10	**A** or **B**

Result of first random number generated (to determine Decision): ___________

Result of second random number generated (to determine Payoff): ___________

Payoff: ___________

APPENDIX Table 4: *AllLowProb* Condition (Experiment 2 Only)

<u>EVERYONE IN THE ROOM</u> WILL BE PAID FOR 1 OF THE 10 DECISIONS ON THIS SHEET.

Decision	Option A	Option B	Your Decision Circle One
1	$32.00 if the die is 1 - 5 $25.60 if the die is 6-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 5 $1.60 if the die is 6 - 50 $0 if the die is 51 - 400	**A** or **B**
2	$32.00 if the die is 1 - 10 $25.60 if the die is 11-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 10 $1.60 if the die is 11 - 50 $0 if the die is 51 - 400	**A** or **B**
3	$32.00 if the die is 1 - 15 $25.60 if the die is 16-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 15 $1.60 if the die is 16 - 50 $0 if the die is 51 - 400	**A** or **B**
4	$32.00 if the die is 1 - 20 $25.60 if the die is 21-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 20 $1.60 if the die is 21 - 50 $0 if the die is 51 - 400	**A** or **B**
5	$32.00 if the die is 1 - 25 $25.60 if the die is 26-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 25 $1.60 if the die is 26 - 50 $0 if the die is 51 - 400	**A** or **B**
6	$32.00 if the die is 1 - 30 $25.60 if the die is 31-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 30 $1.60 if the die is 31 - 50 $0 if the die is 51 - 400	**A** or **B**
7	$32.00 if the die is 1 - 35 $25.60 if the die is 36-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 35 $1.60 if the die is 36 - 50 $0 if the die is 51 - 400	**A** or **B**
8	$32.00 if the die is 1 - 40 $25.60 if the die is 41-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 40 $1.60 if the die is 41 - 50 $0 if the die is 51 - 400	**A** or **B**
9	$32.00 if the die is 1 - 45 $25.60 if the die is 46-50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 45 $1.60 if the die is 46 - 50 $0 if the die is 51 – 400	**A** or **B**
10	$32.00 if the die is 1 - 50 $0 if the die is 51 – 400	$61.60 if the die is 1 – 50 $0 if the die is 51 - 400	**A** or **B**

Result of first random number generated (to determine Decision): ___________

Result of second random number generated (to determine Payoff): ___________

Payoff: ___________

APPENDIX B: POST-EXPERIMENT SURVEY[1]

At this time, we would like all participants to complete this survey. All information provided in this survey will be treated as confidential and can not be used to identify individual participants. We will not ask for your name, address, or other personal information that can identify you. Please try to complete the entire survey, but you can choose not to answer certain questions if you don't want to, and you can end the survey at any time. All information you provide will be kept confidential.

1. Please enter the ID Number written on the index card at your desk.

2. In what year were you born?

3. What is your gender?
   Male
   Female

4. What category best describes your racial and ethnic background?
   White or Caucasian
   Black or African American
   Hispanic
   Asian Asian-American
   Multiracial or other

5. What is your marital status?
   Never married
   Married
   Divorced
   Separated
   Widowed

7. In what country were you raised?
   U.S.
   Other

8. In what country were your parents raised?
   Both in U.S.
   One in U.S.
   Neither in U.S.
   Unknown

9. What is the highest level of education you have completed?
   Less than high school degree
   High school degree
   Some college
   College degree
   Graduate degree
   Unknown

10. How would you describe your employment status?
    Retired
    Part-time employment
    Full-time employment
    Stay at home parent
    Not employed

[1] Every question also included an option for subjects to choose not to answer the question.

Full-time student

11. How many people participating in this experiment today do you consider to be your friend?
    How often do you recycle?
    Nearly all the time (every day)
    Frequently (a few times a week)
    Occasionally (a few times a month)
    Never

12. Are you a U.S. citizen?
    Yes
    No

13. How often do you buy environmentally of socially labeled products (for example, fair trade products, low energy light bulbs, or recycled products)?
    Nearly all the time when I shop
    Occasionally when I shop
    Never

14. During the past two years have you been a member, contributed time, or contributed money to a social organization (for example, soup kitchens or Big Brother-Big Sister).
    Yes
    No

15. If you are a member of a political party, to which party do you belong?
    Democratic
    Republican
    Libertarian
    Green
    Other
    I am not a member of a political party

16. Which political party best represents your interests?
    Democratic
    Republican
    Libertarian
    Green
    Other

17. How often do you wear a seatbelt when driving or riding in a car?
    Always, or almost always
    Most of the time
    Some of the time
    Never, or almost never

18. If you drive a car, how often do you drive over the speed limit?
    Always, or almost always
    Most of the time
    Some of the time
    Never, or almost never
    Not applicable; I don't drive a car

19. How often have you gambled or purchased lottery tickets in the last year?
    Never
    Once or twice
    Between three and twelve times

More than 12 times

20. What best describes your religious affiliation?
    None
    Catholic
    Protestant
    Jewish
    Muslim
    Other religion

21. Do you smoke cigarettes every day, some days, or not at all?
    Every day
    Some days
    Not at all

22. A drink of alcohol is 1 can or bottle of beer, 1 glass of wine, 1 can or bottle of wine cooler, 1 cocktail, or 1 shot of liquor. During the past week, how many days did you have at least one drink of any alcoholic beverage?

23. On the days when you drank, about how many drinks did you drink on average?

24. Are your parents living now? (Parents refers to biological or legally adoptive parents, not step-parents.)
    Only mother is living
    Only father is living
    Both mother and father are living
    Both mother and father are deceased

25. Thinking first about your father: If he is living, how old is he now? Or if he is deceased, at what age did he die?

26. Now, thinking about your mother: If she is living, how old is she now? Or if she is deceased, at what age did she die?

27. Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee you an annual income for life that is equal to your current total family income. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would double your total lifetime income and a 50-50 chance that it would cut it by a third. Which job would you take -- the first job or the second job?
    First job
    Second job
    Do not know

28. Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it in half. Would you take the first job or the second job?
    First job
    Second job
    Do not know

29. Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by twenty percent. Would you take the first job or the second job?
    First job
    Second job
    Do not know

30. Suppose that a distant relative left you a share in a private business worth one million dollars. You are immediately faced with a choice -- whether to cash out now and take the one million dollars, or to wait until the company goes public in one month, which would give you a 50-50 chance of doubling your money to two million and a 50-50 chance of losing one-third of it, leaving you 667 thousand dollars. Would you cash out immediately or wait until after the company goes public?
    Cash out
    Wait
    Do not know

31. Suppose that waiting a month, until after the company goes public, would result in a 50-50 chance that the money would be doubled to two million dollars and a 50-50 chance that it would be reduced by half, to 500 thousand dollars. Would you cash out immediately and take the one million dollars, or wait until the company goes public?
    Cash out
    Wait
    Do not know

32. Suppose that waiting a month, until after the company goes public, would result in a 50-50 chance that the money would be doubled to two million dollars and a 50-50 chance that it would be reduced by twenty percent, to 800 thousand dollars. Would you cash out immediately and take the one million dollars, or wait until after the company goes public?
    Cash out
    Wait
    Do not know

33. How do you finance your college education? Check all that apply.
    Scholarship(s)
    Student Loans
    Income from part-time or full-time jobs
    Parent/Guardian contributions
    Other

34. About how much do you weigh in pounds?

35. For your height, please enter two numbers, feet in the first box and inches in the second box. For example, a 5 foot 7 inch tall person would put a 5 for the first height question and a 7 for the second. About how tall are you? - FEET

36. About how tall are you?- INCHES

37. During the past two years have you been a member, contributed time, or contributed money to an environmental organization (for example, a campus environmental group of the Nature Conservancy).
    Yes
    No

38. What is your zip code?

39. Which category best describes your college residency status?
    In-state
    Out-of-state

40. Which category best describes your college social status?
    Freshman
    Sophomore
    Junior
    Senior

41. What is your primary academic interest area/major area?
    Sciences
    Social Sciences
    Arts and Humanities
    Business

    Closing Statement: Thank you for completing the survey. Please remain seated momentarily and someone will come to your desk to pay you for your participation in the experiment.