Loss Attitudes in the U.S. Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE)*

Jonathan Chapman
NYUAD
jchapman@nyu.edu
www.jnchapman.com

Erik Snowberg
Caltech, UBC, CESifo, NBER
snowberg@caltech.edu
hss.caltech.edu/∼snowberg/

Stephanie Wang
University of Pittsburgh
swwang@pitt.edu
www.pitt.edu/∼swwang/

Colin Camerer
Caltech
camerer@hss.caltech.edu
hss.caltech.edu/∼camerer/

November 7, 2019

Abstract

To measure individual-level loss aversion in a representative sample of the U.S. population ($N = 2{,}000$), we introduce DOSE—Dynamically Optimized Sequential Experimentation. We find that around 50% of the U.S. population is loss tolerant. This is counter to earlier findings, which mostly come from lab/student samples, that a strong majority of participants are loss averse. Loss attitudes are correlated with cognitive ability: loss aversion is more prevalent in people with high cognitive ability, and loss tolerance is more common in those with low cognitive ability. We also use DOSE to document facts about risk and time preferences, and demonstrate that DOSE elicitations are more accurate, more stable across time, and faster to administer than standard methods.

JEL Classifications: C81, C9, D03, D81, D9

Keywords: Loss Aversion, Dynamic Experiments, DOSE, Risk Preferences, Time Preferences

*We thank Summer Clay, Alison Harris, and Alec Smith for generously sharing their data. Thanks to Roland Benabou, Doug Bernheim, John Beshears, Gary Charness, Mark Dean, Kate Johnson, Ian Krajbich, Andreas Krause, Pietro Ortoleva, Deb Ray, Antonio Rangel, Hans-Martin von Gaudecker, Peter Wakker, Nathaniel Wilcox, and the participants of seminars and conferences for their useful comments and suggestions. Judah Okwuobi and Michelle Filiba provided excellent research assistance. Camerer and Snowberg gratefully acknowledge the financial support of NSF Grant SMA-1329195. This paper subsumes the earlier, largely methodological, working paper, "Dynamically Optimized Sequential Experimentation (DOSE) for Estimating Economic Preference Parameters" (2010).

1 Introduction

An important hypothesis in behavioral economics is that people treat losses and gains differently, resulting in most being *loss averse*: even if they are risk neutral, they tend to shy away from positive expected value gambles with negative payoffs (losses). Loss aversion is used as an explanation for a number of important economic phenomena,[1] and is an essential ingredient in theories of reference-dependent preferences (Kahneman and Tversky, 1979; Kőszegi and Rabin, 2006; O'Donoghue and Sprenger, 2018).

Yet, most evidence of loss aversion comes from economics and psychology labs, usually with university student participants who often have different preferences than the general population (Snowberg and Yariv, 2018; Walasek et al., in progress). The hypothesis of differential responses to gains and losses would thus benefit from individual-level assessments in a representative sample. However, as we detail in Section 2, current methodologies make such an assessment difficult. To overcome these difficulties, we introduce DOSE—Dynamically Optimized Sequential Experimentation—which estimates preference parameters precisely and quickly by selecting a personalized sequence of simple choices for each participant.

Using DOSE, we find that around 50% of people in the U.S. are *loss tolerant*: even if they are risk neutral, they embrace gambles with negative expected values. Loss aversion is more prevalent in people with high cognitive ability, and loss tolerance is more common in those with low cognitive ability. Moreover, we find that risk aversion over gains and loss attitudes are equally stable across time (and more stable than previously appreciated), suggesting that both are equally important in understanding risk preferences.

It is important to emphasize that, although surprising, the prevalence of loss tolerance is *not* evidence against the hypothesis of gain-loss differences. Rather, it is evidence of substantial heterogeneity in the asymmetry, with potentially important consequences for

[1] Examples include the equity premium puzzle (Mehra and Prescott, 1985; Benartzi and Thaler, 1995), asymmetric consumer price elasticities (Hardie et al., 1993), reference-dependent labor supply (Dunn, 1996; Camerer et al., 1997; Goette et al., 2004), tax avoidance (Rees-Jones, 2017), opposition to free trade (Tovar, 2009), performance in athletic contests (Pope and Simonsohn, 2011; Allen et al., 2016), and more.

welfare economics. In particular, loss aversion can, in theory, reduce the propensity to use financial products that exploit common characteristics like overoptimism and skew-love (Kahneman and Lovallo, 1993; Åstebro et al., 2015). Loss tolerance, on the other hand, makes people more susceptible to exploitation of these characteristics. Moreover, our evidence suggests that loss tolerance is particularly prevalent in precisely the people who might benefit from additional reservations about problematic financial products: those with low income, education, and cognitive ability, and the aged (Kornotis and Kumar, 2010; Chang, 2016).

DOSE takes the challenges of eliciting loss aversion—the need for multiple choices and, usually, a parametric model—and designs around them.[2] DOSE uses the parametric structure and rapid computation of Bayesian updating to dynamically select a personalized sequence from a set of simple questions, as described in Section 2.2. That is, DOSE starts with a prior over parameters and/or models and, based on that prior, selects a question that will maximize information. After each question, DOSE then uses a participant's choice to dynamically update its priors about that participant, and selects the next question in the personalized sequence based on the new distribution over parameters and/or models.

We use DOSE to estimate loss aversion using an incentivized, representative survey of the U.S. population (N = 2,000), described in Section 2.4. This incentivized survey has several useful features. It is comprehensive, using a wide range of elicitations to measure different preferences. Moreover, it is repeated, meaning the same participants are asked the same questions twice, six months apart. These features allow us to establish a number of facts about loss aversion, as well as evaluate DOSE, in an important practical setting.

There is a much higher level of loss tolerance in the U.S. population than indicated by prior samples. The loss aversion parameter in Prospect Theory, λ, indicates loss aversion when $\lambda > 1$, and loss tolerance when $\lambda < 1$. In Section 3, we find that 52% of the U.S.

[2] Estimating an individual index of loss aversion without a parametric structure attributes all differences in the curvature of utility functions over gains and over losses to loss aversion. In principle, a non-parametric approach allows a classification of people into loss averse/neutral/tolerant, but, in practice, many cannot be classified. For example, using a non-parametric method, Abdellaoui et al. (2007) find that between 29% and 88% of participants cannot be classified, depending on which definition of loss aversion is used.

population is loss tolerant. This is higher than 13–30% (weighted average 22%, $N = 1{,}023$) in the eight studies we are aware of that investigate heterogeneity in loss aversion (all in lab samples).[3] Moreover, the median level of loss aversion in our study is 0.99, versus 1.5–2.5 in other samples. We show that this difference is not due to DOSE: among 369 student participants in an almost exact replication of our representative study, 10% are loss tolerant, with a median value of $\lambda = 1.84$. Moreover, those with greater education and cognitive ability, and lower age, are more likely to be loss averse in our representative sample.[4] These attributes describe the student samples usually used in studies of loss aversion. Indeed, in our representative sample, 23% of those under 35 with a college education ($N = 101$) were loss tolerant, with a median value of $\lambda = 1.75$. Altogether, this suggests that the prevalence of findings of loss aversion, rather than loss tolerance, may be the result of inadvertently selecting highly loss-averse samples, a topic we return to in our discussion in Section 5.

An important feature of DOSE is that it dynamically estimates, and adjusts for, an individual's level of choice consistency. This produces two more substantive results. First, although we find a correlation between higher cognitive ability and less risk aversion using DOSE, we do not find this relationship using an MPL-based measure of risk aversion. However, if we examine only those participants that DOSE tells us make consistent choices, we recover a similar relationship using the MPL measure, suggesting that choice inconsistency and resultant measurement error may lead to the mixed results on the relationship between cognitive ability and risk aversion (Dohmen et al., 2018). Second, we document that loss aversion is nearly as stable as risk aversion and time discounting, indicating that all three are similarly important in describing preferences. Moreover, prior studies seem to have underes-

[3]These studies are Schmidt and Traub (2002); Brooks and Zank (2005); Abdellaoui et al. (2007, 2008); Sokol-Hessner et al. (2009); Abdellaoui et al. (2011); Sprenger (2015); Goette et al. (2018). The figure for Sprenger (2015) is reported in Footnote 8 of Goette et al. (2018). We compare estimates of risk aversion in this study and prior studies in Appendix A.

[4]Most studies of the relationship between cognitive ability and risk preferences have focused on lotteries over gains (see Andersson et al., 2016b, Table E1 for a summary). The few studies with questions involving losses have found that lower cognitive ability is associated with fewer expected value-maximizing choices on those lotteries—consistent with our results—although differences in design and data reporting make it difficult to ascertain the degree of agreement.

timated the stability of risk and time preferences, suggesting that controlling measurement error is essential in estimating the true stability of preferences (Gillen et al., Forthcoming).

Our results from DOSE are robust to a number of factors, such as misspecification and removing participants most likely to not be paying attention, as shown in Section 4. Allowing for different specifications of the utility function still results in much lower estimates of loss aversion and much higher estimates of loss tolerance than prior studies on student/lab populations. Alternative reference points fit the data poorly. Removing participants that may be "rushing through" DOSE, or our entire study, has minimal effects on the distribution of DOSE-estimated parameters. Additional robustness checks are conducted in Appendix F.

The paper concludes with a discussion, in Section 5, of the potential reasons why prior studies have misunderstood the patterns of loss attitudes. We also discuss the potential for DOSE to be used more widely. We conclude with a brief discussion of how our results may inform broader discussions in welfare economics.

1.1 Related Literature

Our work is related to three broad literatures: loss aversion, optimal experimental design, and measuring economic preferences and their correlates in broad populations. We review these literatures here: relationships between specific factual findings in this paper and others are included when we discuss those specific findings, and in Appendix A.

There is a large literature interested in understanding loss aversion. Most studies focus on lab/student populations.[5] von Gaudecker et al. (2011) is the most similar to our work. As noted above, they focus on population distributions, and their estimates are sensitive to estimation choices. Depending on those choices, they report estimates (in their appendix)

[5]See Table 1 of Booij et al. (2010) and Table S4 of Sokol-Hessner et al. (2009) for estimates from lab studies. We are aware of four field studies that measure loss aversion in non-representative populations, but only report first moments. These studies feature samples of customers at a car dealer (Gächter et al., 2007), Vietnamese villagers (Tanaka et al., 2010), Mechanical Turk workers (Toubia et al., 2013), and U.S. mortgage holders (Atlas et al., 2017). Reported first moments of loss aversion are similar to lab studies. Laboratory experiments in psychology suggest that loss aversion may not extend across contexts (see the review in Gal and Rucker, 2018). Erev et al. (2008), for example, do not find evidence of loss aversion in decisions made with feedback.

of the median λ ranging from 0.12 to 4.47.[6] As we show in Section 4.1 and Appendix F.1, our results are relatively stable with respect to different specifications. It is worth noting, however, that their results are not inconsistent with ours: the shape of the loss aversion distribution we find (in Figure 3) is very similar to theirs. Moreover, some of their specifications produce results much closer to ours than to the prior literature.

While optimal experimental design is the subject of literature in both computer science and statistics, there is surprisingly little development of applications for economics.[7] Those few studies that exist focus on static, rather than dynamic, experiments (El-Gamal et al., 1993; El-Gamal and Palfrey, 1996). DOSE extends these studies by implementing a dynamic design, in which questions are selected sequentially based on a participant's answers. This allows for better identification of models and parameters at the individual level, in contrast to prior designs which could only discriminate between models or measure the distribution parameters at a population level. Taking advantage of recent advances in computing power, we are also able to account for a much larger range of parameters in designing an experiment.

Two recent papers that have examined dynamic experimental procedures draw on the working paper version of this manuscript (Wang et al., 2010).[8] Toubia et al. (2013) use DOSE to study risk and time preferences in the lab. Imai and Camerer (2018) use DOSE to evaluate time preferences on Mechanical Turk, but focus on model selection.[9]

---

[6] Their estimation strategy also does not allow them to use the S-shaped utility function suggested by Prospect Theory (Kahneman and Tversky, 1979).

[7] The idea of optimal experimental design appears to originate most clearly in Peirce (1879), who described an "economic" theory of experimentation and applied it to the study of gravity. The idea of dynamic designs begins with Wald (1950). Chaloner and Verdinelli (1995) provide a useful review of applications in statistics. Although little used in economics, optimal designs have been used in many applied fields including in neurophysiology (Lewi et al., 2009), psychophysics (Kujala and Lukka, 2006; Lesmes et al., 2006), marketing (Toubia et al., 2004; Abernethy et al., 2008), and medicine (Müller et al., 2007). See also Aigner (1979) for an early survey in economics, and Moffatt (2007) for a discussion of potential applications of optimal design to parameter estimation, including the elicitation of risk preferences.

[8] Cavagnaro et al. (2010, 2013a,b, 2016) independently develop an adaptive framework for model discrimination. Their implementations use many more questions than DOSE—for example, 80 in Cavagnaro et al. (2016), 101 in Cavagnaro et al. (2013a)—making it difficult to use with a representative sample.

[9] This paper uses a different information criterion—EC^2 rather than the Kullback-Leibler divergence we use—and a different model of inconsistencies in decision-making. The Kullback-Leibler criterion is particularly well suited to efficient parameter estimation (Ryan et al., 2016) but may not be as efficient in model selection. Thus, the main contribution of that paper is to illustrate how the novel criterion EC^2 is used, and apply it to distinguish different models of time preferences (in participants recruited from MTurk) more

Our paper also contributes to the recent literature studying the correlates of economic preferences in broad populations. Many of these studies focus on the role of cognitive ability in economic preferences, generally concluding that higher cognitive ability is associated with greater normative rationality (Frederick, 2005; Burks et al., 2009; Oechssler et al., 2009; Dohmen et al., 2010; Benjamin et al., 2013). We add to this literature in two ways. First, we examine the relationship between loss aversion and cognitive ability, and find that both low- and high-cognitive-ability people tend to depart from normative rationality, but in different ways. Second, we show that the mixed results on the relationship between risk aversion and cognitive ability (Dohmen et al., 2018) are likely due to measurement error and imprecision in the elicitation techniques used in prior studies.

2 The DOSE Procedure

This section introduces the DOSE procedure first from the point of view of participants, then from that of the experimenter. This emphasizes that, while DOSE appears simple to participants, it gives experimenters a number of new degrees of freedom. This leads into a discussion of our specific design and data-gathering choices.

Before describing DOSE, we note that a new technique is needed to measure loss aversion (and other preferences) in representative populations. In the three studies that elicit loss aversion in a representative population, two lose more than 70% of participants due to non-response and inconsistent choice (Booij and Van de Kuilen, 2009; Booij et al., 2010). A third only recovers population distributions of loss aversion, and is very sensitive to estimation choices (von Gaudecker et al., 2011). Depending on those choices, it produces population estimates that vary from a large majority being extremely loss averse, to almost everyone being loss tolerant. Individual-level estimates are necessary to study the correlates of heterogeneity of preferences (Harrison et al., 2002; Goette et al., 2018), or to calibrate personalized contracts (Andreoni et al., 2016). Moreover, standard techniques for measuring

rapidly and precisely than earlier research.

Figure 1: Examples of DOSE from a Participant's Point of View

In the next few questions, you will be asked to choose between two lotteries.

You will start this section with 10,000 points, which you may lose based on the lotteries you choose in this section. That is, some of the lotteries in this section may both **add** to or **subtract** from this initial 10,000 points.

For example, suppose you chose a lottery that had a 50% chance of adding 5,000 points, and a 50% chance of subtracting 5,000 points. In the case of winning, the 5,000 will be added to your additional 10,000. In the case of a loss, the 5,000 will be subtracted from your initial 10,000. Note that you will never have the possibility of losing more than 10,000, so at worst you will end this section with 0 points.

(a) DOSE Instructions

Which of the following options do you prefer?

- A lottery where you can either receive 7,000 points or lose 6,300 points, each with probability 50%;

OR

- Receiving 0 points for certain.

(b) Sample DOSE Choice

other economic preferences produce unstable estimates (Meier and Sprenger, 2015), and may unintentionally introduce reference points (Sprenger, 2015). DOSE produces stable (Section 3.4), fast (Section 4.3), and accurate (Appendix D) individual-level estimates of preference parameters that are robust to alternative specifications (Section 4.1). Using DOSE to study risk and time preferences produces similar—often less noisy—results to prior studies that use standard elicitation techniques (Section 3.2 and Appendix A).

2.1 The Participants' Point of View

DOSE can be designed to be quite simple from the participants' point of view, as shown in the examples taken from our implementation in Figure 1.[10] The participant is given a simple explanation of the upcoming choices, as in Figure 1a. He or she is then given a series of simple choices, similar to those in Figure 1b.

We chose lotteries and sure amounts in order to make the choices particularly simple. In particular, all lotteries contain 50% probabilities of different payoffs. When a lottery

[10]This figure includes screenshots from our actual implementation. For screenshots of all the questions used in this paper, see Appendix B.4. Full design documents and screenshots can be found at hss.caltech.edu/∼snowberg/wep.html.

contains a gain and a loss, then the sure amount is always zero. When the lottery does not contain a loss, one of the outcomes of the lottery is always zero points. This make expected value calculations simple, but does not allow for the identification of potential features of preferences such as probability weighting.

It is worth noting that the endowment of points in Figure 1a does not appear to affect the reference point of participants, as shown in Section 4.2. Like most methods that measure loss aversion, participants need to be endowed with points to prevent them from owing money. Whether this creates a reference point can easily be tested using DOSE: we can attempt to fit observed choices to a model with a reference point of 10,000 points ($10). Doing so yields a model that predicts only 53.5% of choices correctly—a 7% improvement over random guessing. In contrast, our preferred model, discussed in the next subsection, fits 88% of choices.

2.2 The Experimenter's Point of View

DOSE asks each participant a personalized set of questions. Questions are selected sequentially, using a participant's previous answers to identify the most informative question at that point in time. When selecting each successive question, DOSE accounts for the possibility that the participant may have made mistakes in his or her previous choices. Altogether, this leads to accurate parameter estimates after only a few questions.

The procedure starts with a prior over a set of parameter values, and then, according to the experimenter's chosen criterion, optimally selects questions to pinpoint a participant's preferences. After a participant answers the first question, DOSE updates beliefs using Bayes' law, optimally selects the next question, and so on. The process continues for as many questions as the experimenter wants.

DOSE can elicit more accurate parameter estimates than other common dynamic experimental designs because it allows for the possibility that participants make mistakes, as we illustrate by comparing DOSE with a simple partitioning method, in Figure 2.[11] Both methods start with a uniform prior, and offer participants a binary choice. In the first round, each participant faces the same question (Q_1 or q_1). Beliefs are then updated depending on the answer they provide, and the next question is picked optimally given the new beliefs. The key difference between the two procedures is that a partitioning method successively eliminates ranges of parameter values after each question. DOSE, in contrast, allows for the possibility that any choice may have been a mistake, and hence places a positive probability on all parameter values regardless of previous answers.

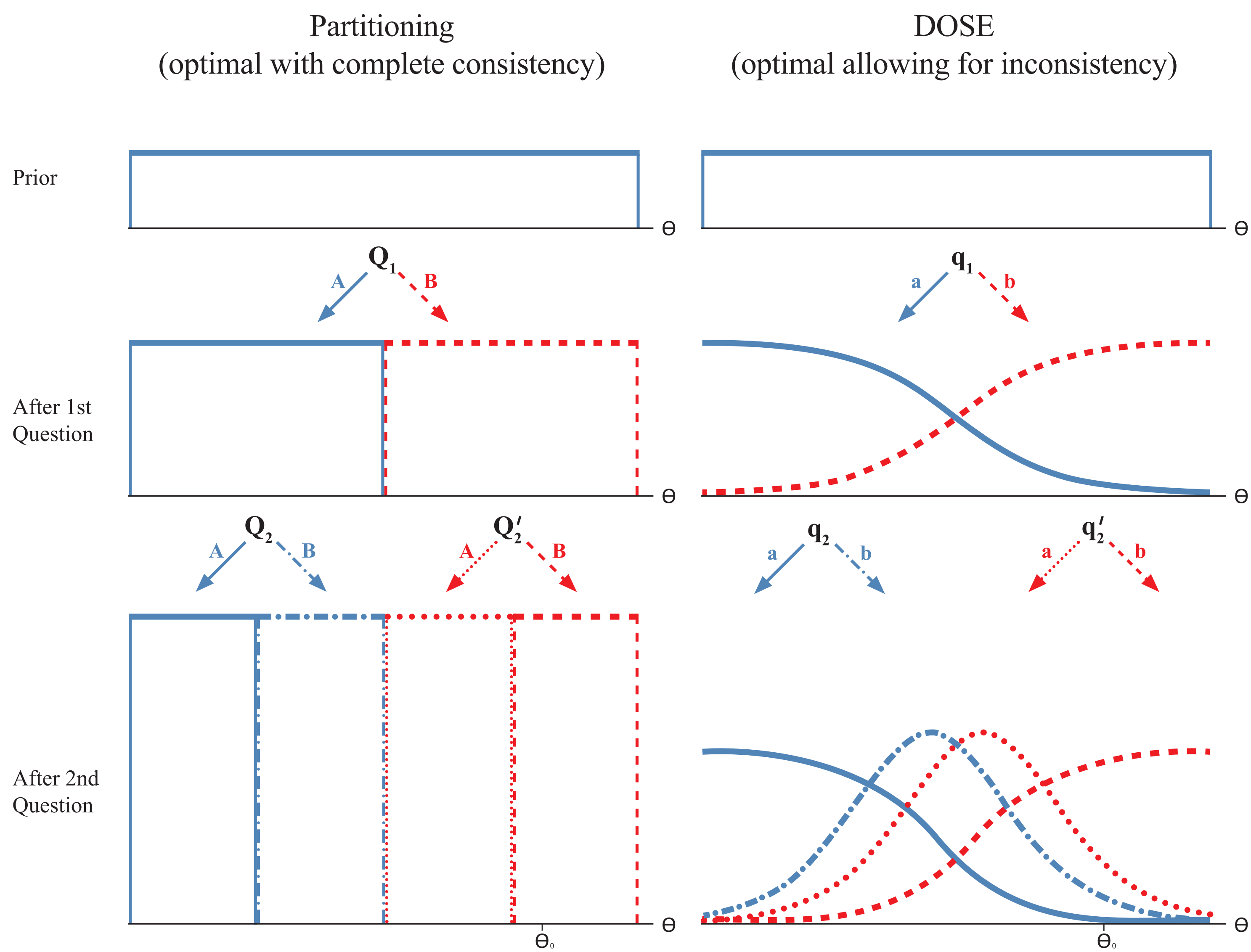
In a partitioning method, a single incorrect choice can cause considerable inaccuracy. Consider a participant with true parameter θ_0, displayed in the bottom panel of Figure 2. This participant should choose B in both questions of a partition method. If, however, he or she incorrectly chooses A in the first question, his or her estimated parameter value is constrained to be less than the median value—regardless of the number of rounds of questions. Errors early in the procedure thus lead to considerable measurement error. Any error makes it particularly hard to identify parameter values at the extremes of the distribution.

DOSE, in contrast, can elicit accurate parameter values even after a participant makes a mistake. Even after an initial incorrect choice of a, the posterior distribution places a positive probability on the true parameter value θ_0. As a result, with enough correct answers in future rounds, an accurate parameter estimate will still be obtained. Further, the procedure keeps track of the extent of inconsistent choice, which, as we demonstrate empirically in Section 3.3, provides a valuable measure of participant behavior.

The questions DOSE uses, the precise way in which it selects questions, accounts for possible mistakes, and models the resultant choices, can easily be adapted to meet the needs of a particular research question. Our selections for these various facets were made partly

[11] Partitioning techniques include the iterative MPL (see, for example, Andersen et al., 2006; von Gaudecker et al., 2011) and the staircase method (Falk et al., 2018). The iterative MPL presents participants with an initial MPL, and then offers them a refined set of options in another MPL. For example, if the choice on the first MPL implied a participant's certainty equivalent for a lottery lay between $X and $Y, the next MPL would have options in [$X,$Y].

Figure 2: DOSE improves estimate accuracy by allowing for choice inconsistency.



on the basis of theory, and partly on the basis of simulations. The latter is described in the next subsection. Here, we describe the theoretical structure necessary to understand our results. A more technical discussion of the DOSE procedure can be found in Appendix B.1.

We model risk and loss aversion using a Prospect Theory utility function with power utility (Kahneman and Tversky, 1979). In this specification, participants value payments relative to a reference point, which we assume is zero. The standard S-shaped utility function in Prospect Theory implies that, for common parameter values, participants are risk averse over positive payments (gains), and risk loving over negative payments (losses). A kink in the utility function at zero represents loss aversion. Formally:

$$v(x, \rho_i, \lambda_i) = \begin{cases} x^{\rho_i} & \text{for } x \geq 0 \\ -\lambda_i(-x)^{\rho_i} & \text{for } x < 0, \end{cases} \tag{1}$$

in which λ_i parameterizes loss aversion, ρ_i parameterizes risk aversion, and $x \in \mathbb{R}$ is a monetary outcome relative to the reference point. If $\lambda_i > 1$, then the participant is loss averse. If $\lambda_i < 1$, then the participant is loss tolerant. An individual with $\rho_i < 1$ demonstrates risk aversion over gains and risk love over losses. So that higher numbers indicate greater risk aversion, we use the *coefficient of relative risk aversion*: $1 - \rho_i$, in tables and figures.

We model mistakes, or, more precisely, the mapping between utility and choices, using the logit function, which has been widely used in both economics and psychophysics due to its connection with the random utility model.[12] For any choice between options o_1 and o_2 with $V(o_1) > V(o_2)$:

$$\text{Prob}[o_1] = \frac{1}{1 + e^{-\mu_i(V(o_1)-V(o_2))}}. \tag{2}$$

The logit function depends both on the utility difference between options o_1 and o_2 and the choice *consistency* parameter $\mu_i \in \mathbb{R}^+$. The probability of making a mistake—that is,

[12]Specifically, choice probabilities will be logit if the errors in the random utility model have an Extreme Value Type I distribution. See McFadden (2001) for a broader discussion of the history of the logit specification and its properties. DOSE can easily be implemented with multi-answer questions using a multinomial logit or alternative probabilistic choice function.

not choosing the value-maxmizing option—is $1 - \text{Prob}[o_1]$. This is decreasing in the value difference between o_1 and o_2. This decrease is more rapid when μ_i is larger, so higher values of this parameter represent greater consistency in choices.

We use this specification both to choose questions, and to analyze the resultant data. Thus, our parameter estimates for each person are just the mean of the posterior distribution after observing all choices.[13] However, the model used to analyze participants' choices can differ from that used to select questions. Indeed, we engage in such ex-post re-specification in Section 4 to examine the robustness of our results to other potential explanations.

2.3 Simulations

Before fielding our study, we conducted a number of exercises to help us choose features of our design. These exercises involved some laboratory piloting of DOSE, and extensive modeling using parameter distributions found in lab trials. These exercises are extensively documented in Appendices B, C, and D. Here, we summarize those that are most relevant to our implementation decisions.

First, dynamic methods are not generally incentive compatible, and we chose to not modify DOSE in order to make it so. In particular, participants may intentionally mislead the researcher in earlier questions in order to obtain higher payoffs in later questions. There are a number of ways to make DOSE incentive compatible, but they are complex and may confuse participants. In order to ascertain whether such confusion might be worthwhile, we piloted DOSE with and without incentive-compatible mechanisms. The parameter distributions estimated in these pilots were statistically indistinguishable, as described in Appendix B.2. This data is consistent with participants responding to DOSE *as if* it is incentive compatible, perhaps because they are not explicitly told that earlier choices may affect later choice sets, and without extensive knowledge of the dynamic process it is extremely difficult to devise a

[13]Each participant starts with a uniform prior over all parameters. The support of the prior distribution covers individual estimates obtained in lab data: $\lambda \in [0, 4.6]$, $\rho \in [0.2, 1.7]$ and $\mu \in [0, 8]$; see Appendix C. Questions are chosen to maximize the Kullback-Leibler divergence, see Appendix B for a technical treatment.

robust strategy to increase payoffs (Ray, 2015).

Second, we chose to present each participant with 10 questions. This decision was based on two exercises. First, a parameter recovery exercise, documented in Appendix D, suggested that 10 questions were sufficient to almost double the precision of parameter estimates versus a Multiple Price List (MPL), a standard way of eliciting preferences. As shown in Section 3.3 and Appendix D.2 (the latter of which compares our simulations with outcomes from the representative sample), these simulations appear to have underestimated the precision gains of DOSE versus an MPL. Second, 10 questions selected by DOSE provide about twice the information as an equal number of randomly selected questions.

The simulations in Appendices C and D also show that our particular implementation of DOSE is robust to specification of the prior, and various misspecifications of the parametric structure of risk and loss aversion. We examine many of the same potential misspecifications using our actual data in Section 4.1.

2.4 Our Data

We now turn to the practical details of implementing DOSE in two waves of a large, representative, incentivized survey of the U.S. population. The two waves asked the same questions of the same participants, about six months apart. The first wave of the survey collected responses from 2,000 U.S. adults and was conducted online by YouGov between March 27 and April 3, 2015. A second wave recontacted the same population and received 1,465 responses between September 21 and November 23, 2015. We use data from the first wave for most analyses. Results are similar in the second wave data, as shown in Appendix F.6.

Participants in the survey were drawn from a panel of participants maintained by YouGov. YouGov continually recruits new people to the panel, especially from hard-to-reach and low-socioeconomic-status groups. To generate a representative sample, it randomly draws people from various Census Bureau products, and matches them on observables to members of their panel. Differential response rates lead to the over- and under-representation of certain populations, so YouGov provides sample weights to recover estimates that would be obtained from a fully representative sample. We use these weights throughout the paper.[14]

The behavioral measures in this paper were all incentivized: at the end of the survey, two survey modules were selected for payment at random.[15] All outcomes were expressed in YouGov points, an internal YouGov currency used to pay panel members, which can be converted to U.S. dollars using the approximate rate of $0.001 per point.[16] To enhance the credibility of these incentives, we took advantage of YouGov's relationship with its panel, and restricted the sample to those who had already been paid (in cash or prizes) for their participation in surveys. The average payment to participants (including the show-up fee) was $9 (9,000 points), which is approximately three times the average for YouGov surveys. The median completion time of the first wave was 40 minutes. For comparability, we convert points to dollars, using the exchange rate, in our analyses.

3 Economic Preferences in a Representative Sample

The U.S. population is more loss tolerant than participants in student samples. Consistent with this finding, higher-cognitive-ability participants are more loss averse. This contrasts with the prior literature that suggests that those with higher cognitive ability are more "rational," or, as it applies here, more likely to make expected value maximizing choices.

[14]The attrition rate of ≈25% is lower than most online surveys. This is due, in part, to YouGov's panel management, and, in part, to the large incentives we offered. According to Pew Research, YouGov's sampling and weighting procedure yields better representative samples than traditional probability sampling methods with non-uniform response rates, including Pew's own probability sample (Pew Research Center, 2016, YouGov is Sample I).

[15]We chose to pay two randomly selected questions to increase the stakes while making fewer participants upset about their payoffs. Paying for two questions instead of one may theoretically induce some wealth effects, but these are known to be negligible, especially in an experiment such as ours (Charness et al., 2016). Paying for randomly selected questions is incentive compatible under Expected Utility, but not necessarily under more general risk preferences, where it is known that no such mechanism may exist (Karni and Safra, 1987; Azrieli et al., 2018). An old, and still growing, literature suggests this theoretical concern may not be empirically important (Beattie and Loomes, 1997; Cubitt et al., 1998; Hey and Lee, 2005; Kurata et al., 2009), but there are some exceptions (Freeman et al., 2015).

[16]The conversion from points to awards can only be done at specific point values, which leads to a slightly convex payoff schedule. This is of little concern here as these cash-out amounts are further apart than the maximum payoff from the survey.

DOSE estimates of risk aversion, discounting, and choice consistency are in-line with the perspective in the prior literature: higher cognitive ability participants are less risk averse and more patient. The literature, however, has found mixed results about the relationship between risk aversion and cognitive ability. We use DOSE estimates of choice consistency (μ) to show that these results may be explained by inconsistent choice: if we examine those with above-median choice consistency we recover a correlation between MPL-based measures of risk aversion and cognitive ability that is obscured when examining all participants.

Finally, the cross-time consistency of loss aversion, which has not been measured before, is similar to that of risk aversion. Moreover, the cross-time stability of risk aversion and discounting are higher using DOSE than using MPLs, and higher than measured in prior studies. Thus, loss attitudes are at least as stable a descriptor of preferences as these other, more common, preferences, and the stability of those other preferences is likely understated.

3.1 Loss Aversion in the U.S. Population

We find far more loss tolerant participants in the general population than in a student sample, as displayed in Figure 3. While the distribution of parameter estimates from DOSE is largely the same across both waves of our incentivized survey, the one from a student sample is markedly different. This is also true for risk aversion: lab-based populations are less risk averse than the general population, in-line with prior research (see Snowberg and Yariv, 2018, and references therein).[17]

The median estimate of the loss aversion parameter, $\lambda = 0.99$, in the U.S. population is much lower than the "standard" estimate of 2 (Fehr-Duda and Epper, 2012). However, in-line with prior studies, DOSE in a student sample ($N = 369$) produces a median estimate of $\lambda = 1.84$, with 90% of students being classified as loss averse, as also shown in Figure 3.

[17] The median CRRA coefficient ($1 - \rho$) in the general population is 0.28 vs. 0.05–0.09 in the student and lab samples. The median monthly discount factor here (0.90) is in the lowest quartile of the results of three recent laboratory studies using the Convex Time Budget method of Andreoni and Sprenger (2012)—see Appendix Table D1 in Imai and Camerer (2018). See Appendix A for a more thorough review. The distribution of both the discounting and choice consistency measures are displayed in Appendix Figure F.4.

These students were recruited from the University of Pittsburgh Experimental Laboratory (PEEL) mailing list on January 9, 2019 with all students who expressed an interest in the study receiving a unique link to participate. The study was extremely similar to the one used with YouGov's panel. There were two differences. First, we removed some elicitations that are not used in this paper from the survey.[18] Second, although questions were presented using the same point values as in our representative sample, students were told ahead of time that points would be converted into cash value on a Visa gift card at an exchange rate of 1,000 points =$1 within two weeks of completing the survey.[19]

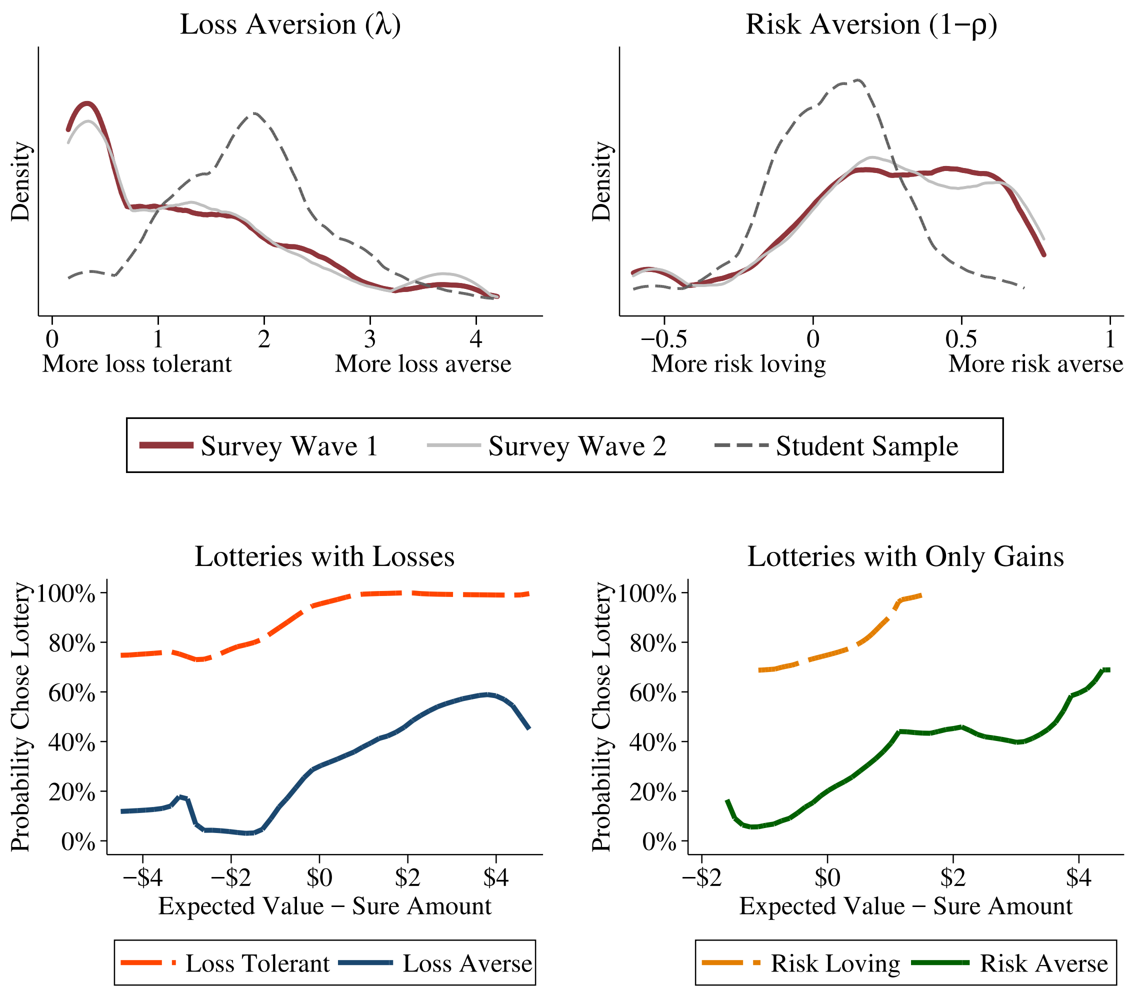
The proportion of loss-tolerant participants in the U.S. population—52%—is also higher than in the eight studies, referenced in the Introduction, that have investigated heterogeneity in loss aversion—all in lab samples. In those experiments, between 13% and 30% (weighted average: 22%, $N = 1{,}023$) of participants that could be classified are loss tolerant. However, as noted above, the methodologies used in several of these studies could not classify many participants. Our representative sample also produces different results than the three extant field studies. Two of those studies obtain estimates for less than 30% of their participants (Booij and Van de Kuilen, 2009; Booij et al., 2010). A third produces estimates between 0.12 and 4.47 of the median level of loss aversion, depending on the specification (von Gaudecker et al., 2011). As we show in Section 4.1 and Appendix F.1, our results are much more stable under different estimation specifications.

The DOSE estimates reflect clear patterns in choice, as shown in the bottom panel of Figure 3. The x-axis is the difference between the expected value of a lottery and the

[18] This resulted in a median time to complete the survey of 22 minutes. The run time was shorter in part because there were fewer tasks, and in part because the students went through each task an average of 25% faster, possibly due to greater familiarity with the lab-based elicitations on the survey. The average payment was $\approx$ \$15.50 due to the removal of some low-average-value tasks. We were not aware that the removal of some tasks would change the expected value so drastically, so we advertised the study as having an average payoff of \$10, about the same as the YouGov study. Design documents and screenshots for this experiment can be found at hss.caltech.edu/~snowberg/wep.html.

[19] We also collected data from lab studies that elicited risk and loss aversion using DOSE. While the details of these implementations vary somewhat from the one used here, these experiments had a total sample size of $N = 439$, with a median loss aversion parameter of 1.99. Further, 90% of students are classified as loss averse. For greater detail, see Appendix E.

Figure 3: Distribution of Economic Preferences within the U.S. population



Notes: The top panel displays the kernel density of each parameter, plotted using Epanechnikov kernel with bandwidth chosen by rule-of-thumb estimator. The bottom panel displays the Nadaraya-Watson (local mean smoothing) estimator (bandwidth 0.6) with Epanechnikov kernel and without sample weights.

sure amount in a given choice. Loss-tolerant participants ($\lambda < 1$) are clearly more likely to choose lotteries with losses than those who are loss averse ($\lambda > 1$), with loss-tolerant participants choosing lotteries nearly 100% of the time when the expected difference is zero. Note, however, that the flat parts of both lines in the left-hand side of the bottom-left panel are due to the fact that DOSE only exposes those who have already revealed loss tolerance through prior choices of lotteries with large negative expected values. Similar patterns exist for those who are risk averse versus those who are risk loving: the latter are more likely to choose gambles with gains at every expected value difference. For all four groups of participants, the probability of choosing the lottery increases with the difference between the expected value of the lottery and the sure amount.

3.2 Economic Preferences and Cognitive Ability

Our comprehensive survey allows us to document new facts about the correlates of loss aversion in the U.S. population. An examination of the simple correlations between economic preferences, socioeconomic characteristics, and cognitive ability reveals that cognitive ability is the most important correlate of loss aversion and the other DOSE-estimated parameters. High-cognitive-ability participants are more loss averse, while those of lower cognitive ability are more loss tolerant, on average. Higher-cognitive-ability participants are more patient and more consistent, in-line with previous studies. Higher-cognitive-ability participants are also less risk averse. In the following subsection, we demonstrate that the mixed evidence on the relationship between risk aversion and cognitive ability in previous studies may be explained by inconsistent choice (Andersson et al., 2016b; Dohmen et al., 2018).

This section uses additional data on time preferences and cognitive ability. Discounting was also elicited using DOSE: it is modeled with a monthly discount factor and the power utility function in (3).[20] Utility from the perspective of the survey date is given by $u(x_t, \rho_i, \delta_i) = \delta_i^t x_t^{\rho_i}$, where δ_i is a discount factor and ρ_i captures the curvature of the utility function from (3), t is the time from the survey date in months, and x_t is a payment at time t. Participants were given 10 binary choices between a lower amount of points at an earlier date or a higher amount at a later date (up to 90 days in the future).

Cognitive ability was measured using a set of nine questions. Six questions were from the International Cognitive Ability Resource (ICAR, Condon and Revelle, 2014): three were similar to Raven's Matrices, and the other three involved rotating a shape in space. We also

[20]The specification used for question selection in the time preference module also allowed for present bias. In practice, however, we found very little evidence of present bias either in the DOSE module or the time-discounting MPLs—possibly due to the fact that payment was, in general, not instantly convertible into consumption. As such, the specification used to obtain estimates did not include a present bias parameter.

Table 1: DOSE preference parameters are correlated with individual characteristics.

	Loss Aversion (λ)	Risk Aversion ($1-\rho$)	Patience (δ)	Choice Consistency (μ)
Cognitive Ability	0.21***	−0.21***	0.18***	0.16***
	(.030)	(.028)	(.029)	(.026)
Income	0.15***	−0.14***	0.11***	0.06*
	(.032)	(.034)	(.034)	(.033)
Education	0.13***	−0.10***	0.17***	0.11***
	(.032)	(.033)	(.037)	(.031)
Male	0.07**	−0.10***	−0.02	0.00
	(.033)	(.032)	(.035)	(.033)
Age	−0.11***	0.01	0.18***	0.06*
	(.033)	(.032)	(.036)	(.036)
Stock Investor	0.06**	−0.11***	0.10***	−0.01
	(.031)	(.029)	(.031)	(.032)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors, in parenthesis, come from a standardized regression. Each cell corresponds to a single regression.

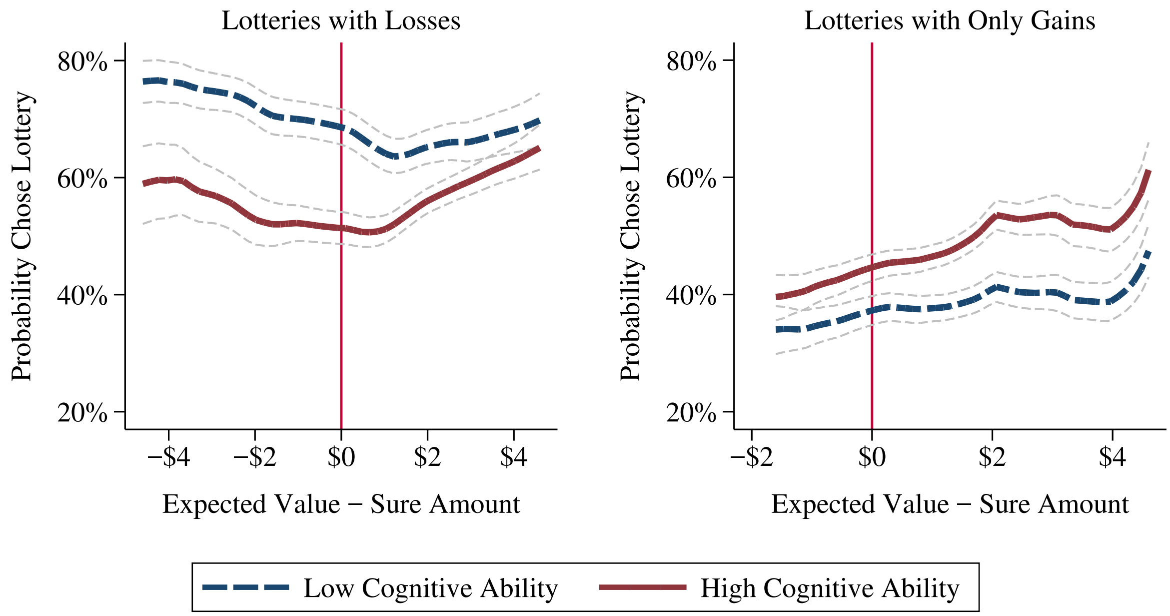
administered the Cognitive Reflection Test (CRT; Frederick, 2005): three arithmetically straightforward questions with an instinctive, but incorrect, answer. Our cognitive ability score was the sum of correct answers to these nine questions.

The relationships in Table 1 are consistent with the analyses in the prior subsection showing that the general population is more risk averse and less loss averse than lab/student populations. In particular, more educated, higher income, and more cognitively able individuals tend to be more loss averse and less risk averse; and lab populations have higher cognitive ability than the general population (Snowberg and Yariv, 2018). Men and younger people tend to be more loss averse, and those that own stock are less loss averse.[21]

The strong correlations between cognitive ability and economic preferences are robust to controlling for the other individual characteristics in Table 1—see Appendix Table F.4.

[21] Appendix Table F.3 presents additional correlations with Church Attendance, Ethnicity, and Home Ownership, and shows that the correlations with the two components of cognitive ability (CRT and IQ) are similar to the correlations in Table 1.

Figure 4: Low-cognitive-ability participants chose more lotteries with losses.



Notes: Figure displays the Nadaraya-Watson (local mean smoothing) estimator (bandwidth 1) with Epanechnikov kernel. Grey dotted lines represent 95% confidence intervals, constructed with 10,000 clustered bootstrap replications. High and low cognitive ability refer to the top and bottom terciles, respectively.

In fact, differences in cognitive ability appear to explain most of the relationship between education and economic preferences.

The choices participants make differ by cognitive ability, as shown in Figure 4.[22] Across the range of expected value differences, high- and low-cognitive-ability participants exhibit different patterns of choice. In the first panel, which focuses only on lotteries with a loss, low-cognitive-ability participants are significantly more likely to choose the lottery than high-cognitive-ability participants. The u-shape of the curve for both ability terciles is driven by the fact that DOSE only presents very negative expected value difference choices to those who have already expressed significant loss tolerance. In contrast, in the second panel, which focuses on lotteries that only contain a zero payoff and a gain, low-cognitive-ability participants are significantly less likely to choose the lottery. In sum: the patterns of correlation between cognitive ability and risk and loss aversion in Table 1 are clearly driven by underlying choices. Low-cognitive-ability participants are especially willing to

[22]Appendix Figure F.4 presents the results in Figure 3 by cognitive ability tercile.

accept lotteries with losses, even when those result in an expected value loss. However, low-cognitive-ability participants are also less willing to choose a lottery over gains.

Very few participants consistently make expected value (EV) maximizing choices, regardless of cognitive ability. Fewer than 2% of participants made 10 EV-maximizing choices, and fewer than 5% made more than 8 such choices. Further, in contrast to some previous studies (for example, Burks et al., 2009; Benjamin et al., 2013), we find the proportion of choices that maximize expected value is only slightly higher for high-cognitive-ability participants: those in the highest tercile of cognitive ability made 57% EV-maximizing choices compared to 52% for participants in the lowest tercile.

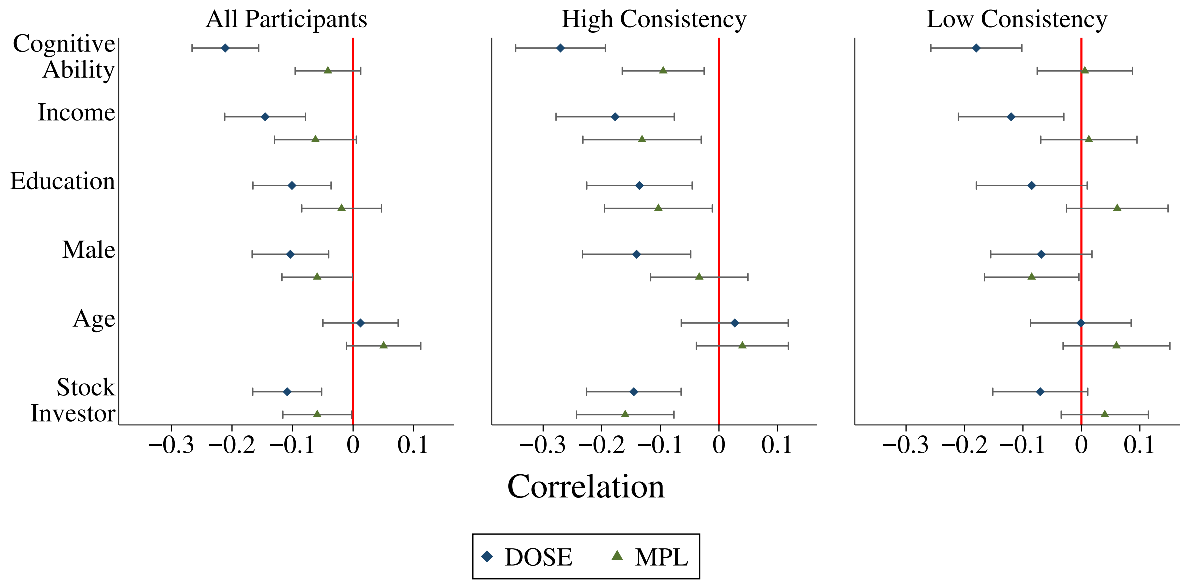
3.3 Choice Consistency and Estimate Accuracy

Accounting for inconsistent choice can explain the mixed evidence about the relationship between risk aversion and cognitive ability in the literature (Dohmen et al., 2018). As discussed in Section 2.3, and shown in Appendix D.3, multiple price lists (MPLs) and other techniques measure risk aversion with considerable error when participants make inconsistent choices. Error attenuates, and potentially biases, any estimated relationship between these measures and other factors. In this subsection, we show that inconsistent choice is related to attenuation in our survey. An MPL measure of risk aversion on our incentivized survey is weakly associated with cognitive ability, in contrast to the DOSE measure.[23] However when we focus only on participants that make (more) consistent choices, the MPL and DOSE measures exhibit similar correlations.

The MPL-based risk aversion measure is less correlated with other characteristics than the DOSE measure, as shown in the first panel of Figure 5. For example, the correlation with cognitive ability is -0.04 (s.e. = .028), compared to -0.21 (.028) for DOSE. This pattern is

[23] Two MPLs asked participants to choose between a fixed 50/50 lottery and a series of ascending sure amounts. The row in which the participant first chose the sure amount identified a range of possible certainty equivalents for the lottery—we use the midpoint of this range. There were two MPLs of this type: the first had a 50/50 lottery over 0 and 10,000 points, the second, a 50/50 lottery over 2,000 and 8,000 points. See Appendix B.3 for more details.

Figure 5: DOSE measure of risk aversion is more highly correlated with individual characteristics before accounting for choice consistency.



Notes: Figure displays correlations between the DOSE and MPL measures of risk aversion and individual characteristics. The left-hand panel includes all participants, the middle contains those with above median choice consistency, and the right-hand panel contains those with below median choice consistency. The survey contained two MPL measures of risk preference. Correlations are estimated by stacking the two and clustering standard errors by participant.

consistent with the simulation results in Appendix D.3, which show greater error in MPLs than in DOSE.

Inconsistent choice is related to the attenuation of correlations. Once we use the DOSE consistency measure to exclude inconsistent participants, there is a strong negative relationship between cognitive ability and the MPL risk aversion measure—see the middle panel of Figure 5, which contains only those with above median choice consistency parameters μ.[24] The magnitudes of correlations are consistently higher for both risk aversion measures; however the contrast is particularly striking for the MPL measure, where a number of relationships—including with cognitive ability—are now statistically significant. This is despite the fact that standard errors are increased by using only half the sample. As shown in the right-hand panel of the figure, DOSE estimates exhibit similar correlations even for

[24] As we discuss in Section 4.3, DOSE estimates may also contain less measurement error because the binary choice questions are easier to understand than MPLs.

very inconsistent participants, while correlations with MPL estimates are almost zero.

The patterns in Figure 5 also demonstrate that the choice consistency parameter can identify individuals who make more mistakes even outside of the DOSE module, and thus help researchers address survey noise. This information is difficult to obtain through other easily available measures. For example, as we demonstrate in Appendix F.5, the correlations in Figure 5 cannot be recovered by truncating the sample based on response time rather than consistency. In fact, the consistency measure helps distinguish whether fast responses reflect a lack of attention: restricting the sample to high-consistency participants recovers correlations even among the subgroup of participants with particularly fast response times.

The value of the choice consistency measure is particularly striking when we consider that our DOSE design did not focus primarily on eliciting this measure. Simple design tweaks—such as allowing the procedure to ask questions multiple times—could allow the variable to be measured more accurately, and hence provide even more information to researchers.

3.4 Within-person Stability of Loss Aversion

Our estimates of loss aversion (λ) are relatively stable, indicating the importance of this parameter to risk attitudes. The correlation of DOSE estimates of loss aversion across survey waves was 0.40 (s.e. = .04). This is comparable to the over-time stability of DOSE estimates of risk aversion (ρ) of 0.44 (.04), and of discounting (δ)—0.47 (.04). Moreover, the stability of loss aversion and loss tolerance are comparable, indicating that this classification is not due to mismeasurement. In particular, Stango and Zinman (2019) document that behavioral attributes are much less stable when they are not in the "expected" direction. This is true in our data for risk aversion: of those who DOSE classifies as risk averse on the first survey, 88% are also classified as risk averse on the second, whereas for those classified as risk loving on the first survey, only 43% are classified as risk loving on the second. However, loss tolerance and loss aversion have similar levels of stability: of those classified as loss averse on the first study, 68% are also classified as loss tolerant on the second, whereas for loss aversion the figure is 71%. This suggests that, in our surveys, risk love, but not loss tolerance, is partly an artifact of mismeasurement.

Moreover, the within-person stability of DOSE estimates of risk and time preference is higher than the within-person stability of estimates from other techniques. The inter-temporal correlation between choices in the two risk MPLs were 0.29 and 0.26 (.04 for both), and for choices in a risky project measure (Gneezy and Potters, 1997) the correlation was 0.33 (.04). The stability in the two time preference MPLs was 0.28 and 0.20 (.06 for both).[25] These findings are consistent with higher measurement error in other techniques.

The over-time correlations of DOSE estimates compare favorably with methods in prior studies. The only study we are aware of that measures stability of risk attitudes in the loss domain is Levin et al. (2007), who report over-time correlations for 62 participants of 0.29 for a risk measure over gains, and 0.20 for a measure of differential risk-taking between the gain and loss domain.[26] Two further studies use incentivized methods to investigate the stability of risk aversion (over gains) over lengthy periods, both finding lower over-time correlations than the DOSE estimates. Gillen et al. (Forthcoming) find an inter-temporal correlation of 0.32 for both of two risk MPLs and 0.36 and 0.47 for two risky project questions. Lönnqvist et al. (2015) report a within-subject correlation of 0.21 for an MPL measure across a year.[27] There is a similar pattern when comparing the stability of the DOSE-measured time preferences to the prior literature, although differences in methodology and samples make it harder to compare (see discussion in Appendix A).

[25]The stability of the consistency parameter (μ) was 0.24 (S.E.=.04); lower than the other DOSE measures but similar to the MPLs discussed above. Part of the explanation for this relatively low correlation is that our DOSE implementation was designed to update more on other parameters: the relatively small number of questions made it harder to identify inconsistent choices.

[26]Levin et al. (2007) report correlations for 62 pairs of parents and children. The figures above are from the adults, for comparability. For the children, over-time correlations are 0.38 and 0.30 respectively.

[27]Andersen et al. (2008b) elicit risk aversion over time, however, they do not report over-time correlations. See Chuang and Schechter (2015) for a review of studies measuring the stability of risk and time preferences, including those using hypothetical questions.

4 Robustness

At a basic level, our main results are obviously robust. The choices we observe, for example in Figure 3, demonstrate widespread loss tolerance in the U.S. population without making any parametric assumptions. A loss-averse participant should never accept a lottery with negative expected value in our implementation, as they could always choose a certain amount of $0 instead. Yet, many participants do. Moreover, to explain the difference between our results and the extant literature, one would also need to explain the difference in results between our representative and student samples. This is difficult to do as the elicitations in these two samples are nearly identical.

Even with these caveats, our data provide the opportunity to reduce concerns about the robustness of our results while learning more about behavior in the representative sample. In this section, we present four such analyses. The first two subsections estimate different utility specifications using choice data. The first subsection focuses on differences in risk aversion across the gain and loss domain, and the second on reference points. It is worth noting that our pre-specified model, which we have used up until this point, fits the data best out of all the different specifications we have tried, matching 88% of choices.[28] The second two subsections utilize response times and random placement of different modules in the survey to examine the role of fatigue and inattention in our results. In short, these factors, to the extent they vary within our study, are unrelated to the result.

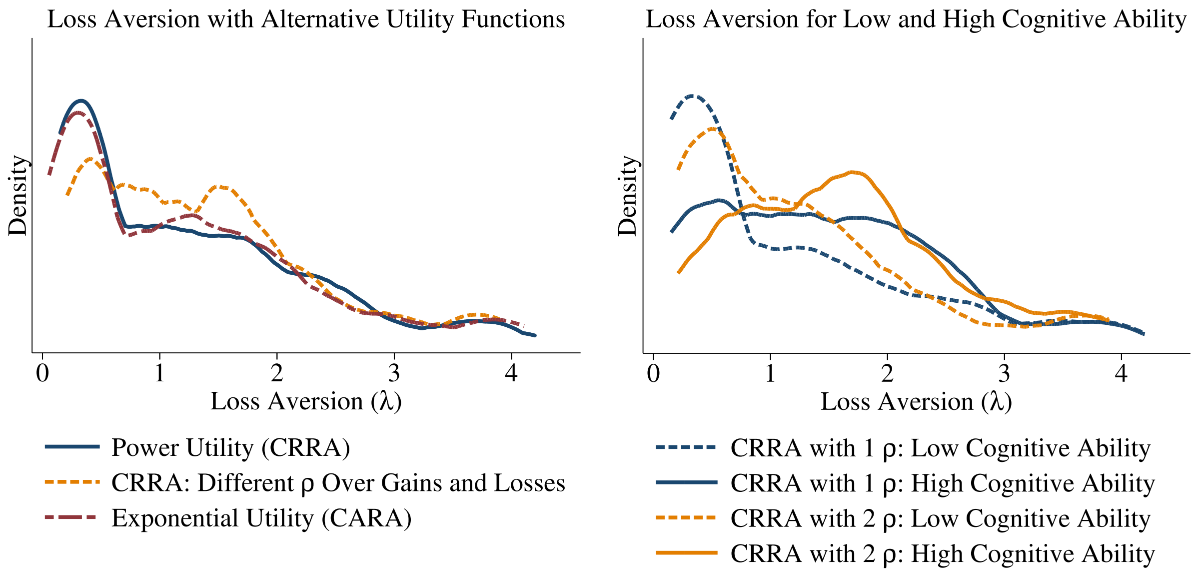
4.1 Misspecification

DOSE is largely robust to misspecification of the parametric form used to select questions, as the choice data can be used to estimate the parameters of different functions.[29] This re-estimation can be done in response to new information, or simply as a robustness check. In this subsection, and the one that follows, we focus on the latter application.

[28]This model was specified in our design documents found at hss.caltech.edu/∼snowberg/wep.html.

[29]In simulations, the re-estimation exercises we employ in this section work well even if the question-selection model is misspecified. See Appendix D.4.

Figure 6: Results on loss aversion are robust to different parametric specifications.



Notes: Plotted using Epanechnikov kernel, with bandwidth chosen by rule-of-thumb estimator. Figure displays the estimated loss aversion parameter after re-estimating the model parameters with alternative functional forms for the utility function—see Appendix F.1 for full details. "Power utility (CRRA)" does not allow for differential curvature over gains and losses—see (3), whereas "CRRA: Different ρ over gains and losses" allows for this possibility.

We can easily estimate different models using our choice data, and do so in Figure 6. This figure shows that our conclusions regarding loss aversion are similar using different utility functions. In addition to the CRRA utility function used in (3), we add a CARA utility function, and a CRRA utility function with different curvature parameters over losses and gains (Tversky and Kahneman, 1992). The latter specification produces the biggest difference. However, the finding that a large portion of the population is loss tolerant is unchanged: the proportion of participants with estimated $\lambda < 1$ ranges from 44% to 52% across the three models. The right-hand panel of Figure 6 breaks down both the one- and two-parameter CRRA utility parameters by cognitive ability. Again, there is the same pattern as in the prior section: those with higher cognitive ability are more loss averse, and those with lower cognitive ability are more loss tolerant.

The CRRA model we used in prior sections fits the data best: it predicts 88% of choices correctly, compared to less than 85% for the other two functions. Moreover, we estimate that most (68%) participants are risk averse over gains and risk loving over losses, in-line with prior experiments and Prospect Theory (Kahneman and Tversky, 1979). The average difference in the curvature between the domains is close to zero, offering support for specification (3), although there is considerable individual heterogeneity—see Appendix F.1.

A particularly pernicious form of misspecification—although one that does not seem to affect our results—is of the error process. In particular, Apesteguia and Ballester (2018) show that the logit error structure may lead to considerable biases. While this issue will not affect the loss aversion (λ) estimate directly—as λ enters utility linearly—it may affect the risk aversion parameter ρ, and this may then lead to some bias in λ. Thus, we re-estimate ρ using a random parameter model (RPM) that Apesteguia and Ballester show to be immune to the issue they describe, and then estimate λ using the RPM estimates of ρ. There was little difference between our standard estimates of ρ and those estimated using the RPM—the correlation between the two sets of estimates is 0.98. Thus, using the RPM estimates of ρ had little effect on the distribution of loss aversion: 49% of participants were classified as loss tolerant, with a median loss aversion parameter of 1.02. For more details about robustness to this and other error processes, see Appendix F.1.

4.2 Reference Points

Our preferred model fits the choice data quite well, especially compared to other common reference point models, as shown in Table 2. The first row shows the statistics for our preferred model. The second row features the most obvious alternative: participants incorporating the endowment of \$10, as shown in Figure 1a, into their values of the various options. If participants did this, every payoff—even those that were negative (with respect to 0)—would appear to be a gain with respect to the amount they began the survey with. As can be seen from the table, this model provides a worse fit: instead of a 77% improvement over chance (random guessing), a \$10 endowment provides an 7% improvement over chance. This model only fits 1% of participants better than our preferred model. We also examine model fit

Table 2: Our preferred model fits better than other standard reference-dependent models.

Reference Point	All Participants: % Improvement over Chance	All Participants: % Participants Improved Fit	Loss-Tolerant Subgroup: % Improvement over Chance	Loss-Tolerant Subgroup: % Participants Improved Fit
Preferred—$0	77%	–	82%	–
Endowment	7%	1%	−4%	1%
EV of Lottery	22%	8%	12%	8%
Sure Option	39%	13%	22%	9%
Stochastic	13%	6%	13%	9%
Choice	29%	8%	33%	11%

Notes: Loss-Tolerant Subgroup is the group of participants our preferred model classifies as loss tolerant. % Improvement over chance is equal to 2∗(the percent of choices fit − 50%). % Participants Improved Fit is the percent of participants for whom the model in that row fits better than our preferred model.

among those participants who our preferred model classified as loss tolerant, reasoning that if there were an error in our model, it would likely be among these participants. The model fit is even worse among these participants: a 4% decline in fit over chance.

The next two rows feature models with fixed reference points: either the expected utility (EV) of the lottery or the sure amount in each question.[30] Either of these reference points could captures the "first focus" concept of Kőszegi and Rabin (2006). These models fare a bit better. However, when the sure option is the reference point, 56% of choices—those containing lotteries with losses—have a reference point of $0, the same as the reference point in our preferred model. That is, by adopting the same reference point as our preferred model for half the choices, this model is roughly halfway between random guessing and our preferred model. It is also worth noting that both of these models fit the choices for loss-

[30] The concept of the expected utility as the reference point has been used in the disappointment models of Loomes and Sugden (1986) and Bell (1985). With our question structure, the sure amount represents the "MaxMin", "MinMax", and "X at Max P" (the outcome with the highest probability) reference points analyzed in Baillon et al. (2019).

tolerant participants slightly worse than those for all participants—whereas our preferred model does slightly better when focusing on loss-tolerant participants.

The final two rows contain stochastic reference point models, as in Kőszegi and Rabin (2006, 2007). First we model a stochastic reference point—that is allowing the lottery reference point to vary probabilistically according to the distribution of prizes in the lottery. The fit here is worse than always using the expected utility of the lottery. Next, we implement Kőszegi and Rabin's (2007) "Choice-Acclimating Personal Equilibrium", in which the decision determines both the reference point and the outcome. That is, before a participant chooses, he or she evaluates the lottery with the EV of the lottery as the reference point, and evaluates the sure amount with the sure amount as the reference point.[31]

Each model considered in Table 2 provides a superior fit to our preferred model for (largely overlapping) sets of up to ∼10% of the population. This suggests an intriguing possibility: perhaps different people use different reference points. Unfortunately, our data are not terribly informative about this possibility, as the choices were not designed to discriminate, at the individual level, between different models of reference point formation. Baillon et al. (2019) explore the possibility of person-specific reference points. To the extent that results are comparable, ours agree with theirs: one of the two best models in their exercise is a "status quo" reference point ($0 in our implementation), as in our preferred model.

This and the prior subsection demonstrate that the DOSE estimates reflect participants' choices. However, these estimates cannot speak to the extent to which those choices accurately reflect individual preferences or, specifically, whether participants paid attention to or understood the DOSE module. Naturally, we lack direct evidence on individuals' understanding or fatigue. However, the next two subsections marshal indirect evidence.

[31] In many cases, scholars fit stochastic reference point models using linear utility. Doing so here makes the model fit slightly worse in all cases.

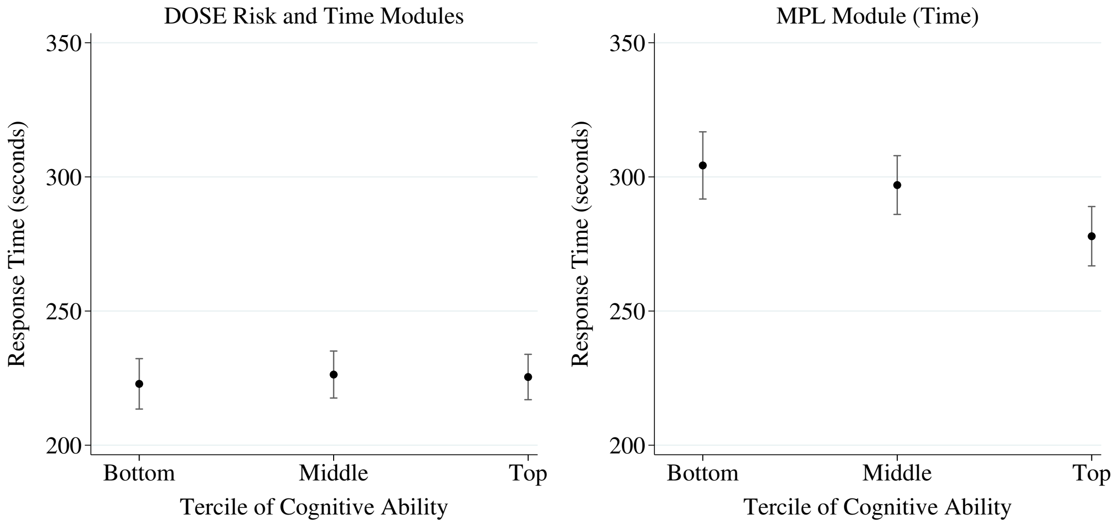
4.3 Complexity

The questions we used in DOSE appear to be easy to understand. First, participants across the cognitive ability spectrum complete DOSE questions in similar amounts of time, while lower-cognitive-ability participants take longer to complete more complicated methods. Second, survey placement, and thus participant fatigue, has little impact on our results.

Earlier results provide some evidence that participants did, in fact, find DOSE relatively simple. The higher correlation of DOSE elicitations across time (compared to MPL elicitations) suggests that it was easier to maintain similar response patterns more than six months later. Moreover, the fact that correlations between DOSE estimates of risk aversion, discounting, and sociodemographic characteristics largely mirror the existing literature, and are stronger than the correlations with MPL measures, also suggests that they more reliably extract preferences. However, the fact that student and general population results differ could be interpreted as evidence of possible confusion in the representative sample.

The fact that DOSE took equally long for all terciles of cognitive ability to complete, as shown in Figure 7, suggests that our results are not driven by question complexity. Question complexity has been shown to affect response patterns (Friedman et al., 2014; Charness et al., 2018). Thus, one may be concerned that those with lower cognitive ability found the DOSE questions more complex, and thus made different choices than they might have otherwise. A way to assess the complexity of a question is the amount of time participants take to answer: if participants struggle to understand a question they will usually take longer to answer it (or much shorter if they give up). However, the length of time it took to complete DOSE (including instructions) shows no evidence that lower-cognitive-ability people found DOSE more complex. This is not because response times do not vary: the (arguably more complex) first MPL module (with instructions) in our study took participants longer, on average, and took low-cognitive-ability participants longer still. Moreover, the variance of time taken on DOSE was relatively constant across terciles, as indicated by the confidence intervals in the figure. Together, these facts suggests that DOSE was equally easy for participants all along

Figure 7: Low-cognitive-ability participants take longer on MPL questions, but not on DOSE.



<u>Notes:</u> DOSE module includes 20 questions addressing both risk and time preferences. MPL module includes two MPLs assessing time preferences, which was the first MPL module on the survey. participants with response times greater than 15 minutes for either module are excluded. Both panels include time taken for questions and explanatory text.

the cognitive ability spectrum.

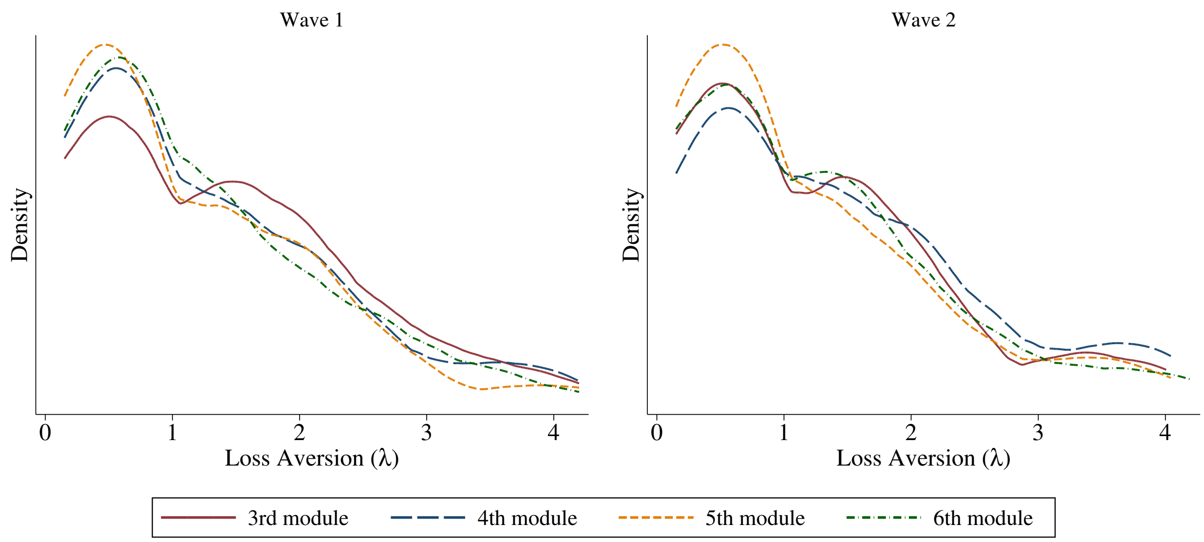
A final manifestation of complexity-induced inattention might be choosing the same option in each question: either the lottery (always listed first), or the sure amount (always listed second). However, there is little evidence of this pattern of inattention in our results: fewer than 6% participants chose the same option in all ten question rounds. While we cannot rule out that some participants rapidly clicked through the DOSE module, such behavior does not appear to affect our results.

It is worth noting that the simplicity of DOSE for participants is helpful to experimenters. DOSE elicited four useful behavioral parameters in less time than it took to elicit a single preference parameter using an MPL.

4.4 Fatigue

Our results also appear to be independent of fatigue. In particular, the further into our survey a participant goes, the more likely he or she will change their response patterns due to fatigue. The position of the DOSE module was randomized across participants: it appeared

Figure 8: Distribution of loss aversion is similar regardless of survey position.



Notes: Plotted using Epanechnikov kernel, with bandwidth chosen by rule-of-thumb estimator.

in four different positions, as the third, fourth, fifth, or sixth task.[32] The distributions of loss aversion for those participants who completed DOSE in a given position, for both wave 1 and wave 2, are plotted in Figure 8. As can be seen, the distributions are similar regardless of position. Moreover, there is no particular pattern of people appearing either more or less loss averse as the modules get later in the survey. In wave 1, the percent of participants classified as loss tolerant is, respectively: 45%, 51%, 58%, and 53%. For wave 2, the percentages are: 49%, 46%, 57%, and 48%. The other DOSE-estimated parameters also show little difference by position, as can be seen in Appendix Figure F.5.

The fact that there is relatively little difference between the parameter distributions is perhaps unsurprising when we consider that the median time it took to reach the DOSE module when it was in the third position was 8 minutes, which may not be enough time for true fatigue to set in. However, there may have been some participants who sped through the entire survey. The distribution of economic preferences is similar when we remove the fastest 20%, 40%, 60%, and 80% of participants as shown in Appendix Figure F.7. The distributions overlap almost entirely—and the median loss aversion parameter consistently remains very

[32]The first position was always an MPL to elicit discounting, and the second an MPL to elicit risk aversion, either through a WTA or WTP for a lottery. All four modules in positions 3–6 were randomized.

similar in the whole sample (1.03 or below): fast response or inattention cannot be said to be the explanation for loss tolerance. Moreover, correlations with other characteristics are also similar when removing the fastest participants—see Appendix Table F.11.

5 Discussion

In this paper, we study loss aversion and other economic preferences in a representative sample of the U.S. population using a new method: DOSE. Our results are summarized in Table 3. A few are worth highlighting. First, we find that around 50% of the U.S. population is loss tolerant over small stakes, differing from prior studies that have found a strong majority of loss-averse participants, usually in lab/student samples. Second, those with greater cognitive ability, education, and income are more likely to be loss averse, and those with lower cognitive ability are more likely to be loss tolerant. This, along with the fact that DOSE in a student sample produces similar results to prior studies, suggests that differences in samples are likely the source of the difference between our results and prior studies. Third, using DOSE's choice consistency parameter, we show that those with high consistency exhibit a correlation on MPL-based measures between higher cognitive ability and less risk aversion. This suggests that the mixed results about this relationship in the literature (Dohmen et al., 2018) may be due to measurement error (Gillen et al., Forthcoming). Fourth, across a range of evaluations, DOSE produces better measures: more accurate, more stable, faster, and so on.

Our findings about loss aversion diverge significantly from conventional wisdom, raising the possibility that the literature may have been influenced by factors beyond the inadvertent sample selection mentioned above. Hints can be found in Fehr-Duda and Epper (2012, p. 576), who observe, "Since the publication of Tversky and Kahneman (1992), any estimates of loss aversion that deviate significantly from the value of two have been eyed with great suspicion, notwithstanding the fact that the original estimate was based on 25

Table 3: Comparison of DOSE with Other Elicitation Methods

		DOSE	MPL	Risky Project / Lottery Menu (†)
Substantive	Loss Attitudes			
	Percent Loss Tolerant	52%	n.a.	n.a.
	Median Loss Aversion Parameter	0.99	n.a.	n.a.
	Correlations w/Cognitive Ability			
	Loss Aversion	0.40	n.a.	n.a.
	Risk Aversion	-0.21	≈ 0	-0.07
	Patience	0.18	0.17	n.a.
	Correlations w/ Demographics			
	Loss Aversion	✓	n.a.	n.a.
	Risk Aversion	✓	≈ 0	✓
	Patience	✓	✓	n.a.
Methodological	Representative Survey			
	Speed	115 secs	259 secs	33 secs
	Stability: Loss Aversion	0.40	n.a.	n.a.
	Stability: Risk Aversion	0.44	0.28	0.33
	Stability: Patience	0.47	0.24	n.a.
	Parameter Recovery Analysis			
	Inaccuracy: Loss Aversion	15%	36%	n.a.†
	Correlation: Loss Aversion	0.91	0.65	n.a.†
	Inaccuracy: Risk Aversion	15%	37%	35%†
	Correlation: Risk Aversion	0.79	0.45	0.28†

Notes: "Inaccuracy" is the average absolute percentage difference between the estimated and true parameter values. "Correlation" is the correlation between the estimated and true parameter values. "Stability" is the correlation across survey waves. ✓ indicates that a pattern of statistically significant correlations were identified. Observations denoted with a † are from the Lottery Menu elicitation—which was included in our simulations, but not our survey—rather than the Risky Project—which was included in our survey, but not our simulations. Parameter Recovery Analysis tests how well a procedure recovers known parameters from simulated data. Details can be found in Appendix D.

subjects, hypothetical decisions over relatively large stakes, and that no standard errors were reported." Relatedly, the one study that examines loss aversion in a representative sample, von Gaudecker et al. (2011), reports a median estimate of $\lambda = 2.38$, although specifications in the appendix have medians ranging from 0.12 to 4.47. More directly, Walasek et al. (in progress) analyze 19 studies of loss aversion in lab/student populations, and find evidence of publication bias, and Yechiam (2018, p. 1) asserts in a review of the loss aversion literature that, "[T]he findings of some of these studies have been systematically misrepresented to reflect loss aversion, though they did not find it."

Our DOSE implementation provides estimates of risk aversion and time preferences that also appear to dominate MPL-based elicitations. DOSE risk aversion measures show stronger correlations with cognitive ability and other characteristics than MPL-based measures, as shown in Table 3, which summarizes many of the results in this paper and its appendices. Further, DOSE-based measures of risk and time preferences show greater stability than MPL-based measures. These facts indicate, in-line with our simulations, that DOSE produces estimates with lower measurement error in these domains as well. Thus, DOSE is an answer to Meier and Sprenger's (2015, p. 286) challenge to develop, "A more precise experimental technique for eliciting time preferences...to make further study of stability."

The accuracy, speed, and simplicity of DOSE also potentially expands the range of research settings in which incentivized preference elicitation is viable. The procedure may be particularly valuable in field experiments, where it is difficult to provide participants with detailed instructions. Similarly, it may also be easier to implement than more complex or time-consuming designs when conducting experiments with low-literacy participants, children, patients with medical disorders, or even animals.[33] DOSE performs better with low-cognitive-ability, low-education, and low-income participants, suggesting that it could be particularly useful in development environments, where current elicitation methods can be plagued by inconsistent choice (Jacobson and Petrie, 2009; Charness and Viceisza, 2012)

[33]Economists have often shied away from using incentivized measures in large samples because of high measurement error and the prohibitive cost of implementing multiple elicitations (Schildberg-Hörisch, 2018).

and new techniques are particularly needed (Berry et al., 2015). DOSE can also be used to discriminate between models on an individual level in real time—an application developed for time preferences in Imai and Camerer (2018), based on an earlier working paper version of this manuscript (Wang et al., 2010).

While we have provided evidence that many concerns about our results are unlikely to be of much import, there are two potential explanations that our data provide no information on: differences in "zero avoidance" (Ert and Erev, 2013) and/or mental accounting (Thaler, 1985) between our representative and student sample. The questions we have used throughout always involve a prize with a (possible) zero payoff—either in the lottery or the sure amount. Avoiding those zero payoffs would produce a choice profile that is consistent with extremely high risk aversion, and extreme loss tolerance, whereas embracing them would be consistent with high loss aversion, and extreme risk-seeking. Thus, we cannot rule out a hypothesis that students are zero seeking, while half the representative population is zero avoidant. While this is a possible rationalization of our results, it represents an extreme departure from the literature. Mental accounting cannot be ruled out because of the slight differences in the way payments are made between the representative and student samples. While new implementations of DOSE can be designed to test these possibilities, they are beyond the scope of the manuscript. Moreover, as we discuss in Section 4.2, the most fruitful application of DOSE at this point is likely to isolate the possibility of person-specific reference points.[34]

We close with some speculative thoughts about how one might isolate the mechanisms behind the association between cognitive ability and loss attitudes, and the implications of that finding for behavioral welfare economics.

Differential attention may lead to the correlation between cognitive ability and loss attitudes. Attention is a scarce mental resource, and lower-ability people might well allocate it differently than higher-ability people do (Mullainathan and Shafir, 2013; Ong et al., 2019). Recent evidence indicates aversion or tolerance to loss is associated with differential attention

[34] We thank an anonymous reviewer for pointing out this possible application.

paid to losses and gains (Yechiam and Hochman, 2013; Bhatia and Golman, 2015; Clay et al., 2017). Exogenously manipulating attention to losses—by visually presenting losses for an extra time—increases estimated loss aversion (Pachur et al., 2018). This hypothesized cause of the ability-behavior correlation could be tested by measuring whether those with lower cognitive ability pay relatively less attention to losses, and how exogenous manipulation of attention differentially affects loss aversion in the high- and low-cognitive-ability people.

Alternatively, low cognitive ability may be indicative of individuals that experience different patterns of losses and gains in their everyday life. This may then, in turn, lead to different adaptive responses to the prospects of losses and gains.[35] To explain the cognitive ability-loss attitude correlation we observe requires that lower-ability people would experience losses more frequently, and become more adapted to them, than higher-ability people.[36] This hypothesis could be tested by gathering more detailed data on whether everyday financial experiences of low- and high-ability people include relatively more losses—that is, do lower-ability people, for example, get more parking tickets, bank overdraft fines, and unexpected bills?

Our finding that loss aversion is more common among the cognitively sophisticated suggests that the conflation of rationality and cognitive ability may be premature, with potentially important consequences for welfare economics. Welfare analysis is difficult in the presence of (apparently) contradictory choices: one must either find a way to write-off one or the other choice as a "mistake," or more deeply understand the reasons for the apparent inconsistency. The former approach has been standard, and various forms of data were employed to argue that certain choices were mistaken.[37] This often leads to choices, and the preferences of those making those choices, being ignored because they seem "irrational."

[35] The last part of this hypothesis is the well-known phenomenon of "adaptive coding" first suggested by Barlow (1961) for sensory encoding in vision. Adaptive coding occurs everywhere in the sensory system, including in the valuation of risk and goods, as a way of allocation attention efficiently (Stewart, 2009; Walasek and Stewart, 2015; Polania et al., 2019).

[36] Experiencing frequent losses could also generate more loss-aversion (Kirby, forthcoming).

[37] More recently, behavioral welfare economics has taken various approaches to make more principled judgements about apparently contradictory choices. For a useful survey see Bernheim and Taubinsky (2019).

The fact that those with lower cognitive ability are more likely to make seemingly irrational choices furthers the notion that seeming-irrationality is a mistake (Benjamin et al., 2013), but also suggests that the preferences of those with lower cognitive ability may be poorly accounted for in welfare analyses.

One reason that the preferences of low-cognitive-ability individuals are often overlooked may be that they are hard to measure: we present a tool—DOSE—that could help overcome this constraint. Those with lower cognitive ability are both harder to reach (they are not well represented amongst undergraduates) and more likely to answer inconsistently (Dave et al., 2010), making it difficult to identify accurate parameter estimates. DOSE offers a solution to both issues: it is relatively easy to implement in non-laboratory environments, and provides accurate estimates even in the presence of inconsistent choice. Further, it is straightforward to adapt the procedure to tackle alternative research questions, including identifying additional parameters for welfare analyses. Slight modifications of the DOSE implementation used here, for example, could distinguish whether inconsistent choice appears random, or whether there are "consistent inconsistencies" in the choices of participants, suggesting a need for improved behavioral models.

References

Abdellaoui, Mohammed, Han Bleichrodt, and Corina Paraschiv, "Loss Aversion under Prospect Theory: A Parameter-Free Measurement," *Management Science*, 2007, *53* (10), 1659–1674.

— , — , **and Olivier l'Haridon**, "A Tractable Method to Measure Utility and Loss Aversion under Prospect Theory," *Journal of Risk and Uncertainty*, 2008, *36* (3), 245.

— **, Olivier L'Haridon, and Corina Paraschiv**, "Experienced vs. Described Uncertainty: Do We Need Two Prospect Theory Specifications?," *Management Science*, 2011, *57* (10), 1879–1895.

Abernethy, Jacob, Theodoros Evgeniou, Olivier Toubia, and Jean-Philippe Vert, "Eliciting Consumer Preferences Using Robust Adaptive Choice Questionnaires," *IEEE Transactions on Knowledge and Data Engineering*, 2008, *20* (2), 145–155.

Aigner, Dennis J., "A Brief Introduction to the Methodology of Optimal Experimental Design," *Journal of Econometrics*, 1979, *11* (1), 7–26.

Allen, Eric J. Patricia M. Dechow, Devin G. Pope, and George Wu, "Reference-Dependent Preferences: Evidence from Marathon Runners," *Management Science*, 2016, *63* (6), 1657–1672.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström, "Elicitation Using Multiple Price List Formats," *Experimental Economics*, 2006, *9* (4), 383–405.

— , — , — , **and** — , "Eliciting Risk and Time Preferences," *Econometrica*, 2008, *76* (3), 583–618.

— , — , — , **and** — , "Lost in State Space: Are Preferences Stable?," *International Economic Review*, 2008, *49* (3), 1091–1112.

Andersson, Ola, Håkan J. Holm, Jean-Robert Tyran, and Erik Wengström, "Deciding For Others Reduces Loss Aversion," *Management Science*, 2016, *62* (1), 29–36.

— , — , — , **and** — , "Risk Aversion Relates to Cognitive Ability: Preferences or Noise?," *Journal of the European Economic Association*, 2016, *14* (5), 1129–1154.

Andreoni, James and Charles Sprenger, "Estimating Time Preferences from Convex Budgets," *The American Economic Review*, 2012, *102* (7), 3333–3356.

— **, Michael Callen, Yasir Khan, Karrar Jaffar, and Charles Sprenger**, "Using Preference Estimates to Customize Incentives: An Application to Polio Vaccination Drives in Pakistan," 2016. NBER Working Paper Series # 22019.

Apesteguia, Jose and Miguel A. Ballester, "Monotone Stochastic Choice Models: The Case of Risk and Time Preferences," *Journal of Political Economy*, 2018, *126* (1), 74–106.

Åstebro, Thomas, José Mata, and Luís Santos-Pinto, "Skewness Seeking: Risk Loving, Optimism or Overweighting of Small Probabilities?," *Theory and Decision*, 2015, *78* (2), 189–208.

Atlas, Stephen A., Eric J. Johnson, and John W. Payne, "Time Preferences and Mortgage Choice," *Journal of Marketing Research*, 2017, *54* (3), 415–429.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy, "Incentives in Experiments: A Theoretical Analysis," *Journal of Political Economy*, 2018, *126* (4), 1472–1503.

Baillon, Aurélien, Han Bleichrodt, and Vitalie Spinu, "Searching for the Reference Point," 2019. Erasmus School of Economics, *mimeo.*

Barlow, Horace B., "Possible Principles Underlying the Transformation of Sensory Messages," in Walter A. Rosenblith, ed., *Sensory Communication*, Oxford, UK: Wiley, 1961, pp. 217–234.

Barsky, Robert B. F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *The Quarterly Journal of Economics*, 1997, *112* (2), 537–579.

Beattie, Jane and Graham Loomes, "The Impact of Incentives upon Risky Choice Experiments," *Journal of Risk and Uncertainty*, March 1997, *14* (2), 155–168.

Bell, David E, "Disappointment in Decision Making under Uncertainty," *Operations Research*, 1985, *33* (1), 1–27.

Benartzi, Shlomo and Richard H. Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle," *The Quarterly Journal of Economics*, 1995, *110* (1), 73–92.

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro, "Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences," *Journal of the European Economic Association*, 2013, *11* (6), 1231–1255.

Bernheim, B.D̃ouglas and Dmitry Taubinsky, "Behavioral Public Economics," in B.D̃ouglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics*, Vol. 1, New York: Elsevier, 2019.

Berry, James, Greg Fischer, and Raymond P. Guiteras, "Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana," 2015. University of Delaware, *mimeo.*

Bhatia, Sudeep and Russell Golman, "Attention and Reference Dependence," 2015. University of Pennsylvania, *mimeo.*

Booij, Adam S. and Gijs Van de Kuilen, "A Parameter-Free Analysis of the Utility of Money for the General Population under Prospect Theory," *Journal of Economic Psychology*, 2009, *30* (4), 651–666.

— , **Bernard M.S. Van Praag, and Gijs Van De Kuilen**, "A Parametric Analysis of Prospect Theory's Functionals for the General Population," *Theory and Decision*, 2010, *68* (1-2), 115–148.

Brooks, Peter and Horst Zank, "Loss Averse Behavior," *Journal of Risk and Uncertainty*, 2005, *31* (3), 301–325.

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini, "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment," *Proceedings of the National Academy of Sciences*, 2009, *106* (19), 7745–7750.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler, "Labor Supply of New York City Cabdrivers: One Day at a Time," *The Quarterly Journal of Economics*, 1997, *112* (2), 407–441.

Cavagnaro, Daniel R., Gabriel J. Aranovich, Samuel M. McClure, Mark A. Pitt, and Jay I. Myung, "On the Functional Form of Temporal Discounting: An Optimized Adaptive Test," *Journal of Risk and Uncertainty*, 2016, *52* (3), 233–254.

— , **Jay I. Myung, Mark A. Pitt, and Janne V. Kujala**, "Adaptive Design Optimization: A Mutual Information-Based Approach to Model Discrimination in Cognitive Science," *Neural Computation*, 2010, *22* (4), 887–905.

— , **Mark A. Pitt, Richard Gonzalez, and Jay I. Myung**, "Discriminating Among Probability Weighting Functions using Adaptive Design Optimization," *Journal of Risk and Uncertainty*, 2013, *47* (3), 255–289.

— , **Richard Gonzalez, Jay I. Myung, and Mark A. Pitt**, "Optimal Decision Stimuli for Risky Choice Experiments: An Adaptive Approach," *Management Science*, 2013, *59* (2), 358–375.

Chaloner, Kathryn and Isabella Verdinelli, "Bayesian Experimental Design: A Review," *Statistical Science*, 1995, *10* (3), 273–304.

Chang, Alvin, "Four Ways the Lottery Preys on the Poor," January 2016. https://www.vox.com/identities/2016/1/13/10763268/lottery-poor-prey.

Chapman, Jonathan, Mark Dean, Pietro Ortoleva, Erik Snowberg, and Colin F. Camerer, "Willingness to Pay and Willingness to Accept are Probably Less Correlated Than You Think," 2017. NBER Working Paper #23954.

— , — , — , — , **and** — , "Econographics," 2018. NBER Working Paper #24931.

Charness, Gary and Angelino Viceisza, "Comprehension and Risk Elicitation in the Field: Evidence from Rural Senegal," 2012. University of California Santa Barbara, *mimeo.*

— , **Catherine Eckel, Uri Gneezy, and Agne Kajackaite**, "Complexity in Risk Elicitation may Affect the Conclusions: A Demonstration using Gender Differences," *Journal of Risk and Uncertainty*, 2018, *56* (1), 1–17.

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

Chuang, Yating and Laura Schechter, "Stability of Experimental and Survey Measures of Risk, Time, and Social Preferences: A Review and Some New Results," *Journal of Development Economics*, 2015, *117*, 151–170.

Clay, Summer N. Allison M. Harris, and Catherine L. Reed, "Loss Aversion Affects Inhibitory Processes for Reward and Loss as Indicated by Inhibition of Return," *Journal of Vision*, 2016, *16* (12), 95.

_ , **John A. Clithero, Alison M. Harris, and Catherine L. Reed**, "Loss Aversion Reflects Information Accumulation, Not Bias: A Drift-Diffusion Model Study," *Frontiers in Psychology*, 2017, *8*, 1708.

Condon, David M. and William Revelle, "The International Cognitive Ability Resource: Development and Initial Validation of a Public-Domain Measure," *Intelligence*, 2014, *43*, 52–64.

Cubitt, Robin, Chris Starmer, and Robert Sugden, "On the Validity of the Random Lottery Incentive System," *Experimental Economics*, 1998, *1* (2), 115–131.

Dave, Chetan, Catherine C. Eckel, Cathleen A. Johnson, and Christian Rojas, "Eliciting Risk Preferences: When is Simple Better?," *Journal of Risk and Uncertainty*, 2010, *41* (3), 219–243.

Dean, Mark and Anja Sautmann, "Credit Constraints and the Measurement of Time Preferences," 2014. Columbia University, *mimeo.*

DellaVigna, Stefano and M. Daniele Paserman, "Job Search and Impatience," *Journal of Labor Economics*, 2005, *23* (3), 527–588.

Dohmen, Thomas, Armin Falk, Armin Huffman, and Uwe Sunde, "Are Risk Aversion and Impatience Related to Cognitive Ability?," *The American Economic Review*, 2010, *100* (3), 1238–1260.

_ , _ , **David Huffman, and Uwe Sunde**, "On the Relationship between Cognitive Ability and Risk Preference," *Journal of Economic Perspectives*, 2018, *32* (2), 115–34.

Dunn, Lucia F., "Loss Aversion and Adaptation in the Labor Market: Empirical Indifference Functions and Labor Supply," *The Review of Economics and Statistics*, 1996, *78* (3), 441–450.

Eckel, Catherine C. and Philip J. Grossman, "Sex Differences and Statistical Stereotyping in Attitudes toward Financial Risk," *Evolution and Human Behavior*, 2002, *23* (4), 281–295.

_ **and** _ , "Men, Women and Risk Aversion: Experimental Evidence," in Charles R. Plott and Vernon L. Smith, eds., *Handbook of Experimental Economics Results*, Vol. 1, North-Holland, 2008, pp. 1061–1073.

El-Gamal, Mahmoud A. and Thomas R. Palfrey, "Economical Experiments: Bayesian Efficient Experimental Design," *International Journal of Game Theory*, 1996, *25* (4), 495–517.

El-Gamal, Mahmoud A., Richard D. McKelvey, and Thomas R. Palfrey, "A Bayesian Sequential Experimental Study of Learning in Games," *Journal of the American Statistical Association*, 1993, *88* (422), 428–435.

Erev, Ido, Eyal Ert, and Eldad Yechiam, "Loss Aversion, Diminishing Sensitivity, and the Effect of Experience on Repeated Decisions," *Journal of Behavioral Decision Making*, 2008, *21* (5), 575–597.

Ert, Eyal and Ido Erev, "On the Descriptive Value of Loss Aversion in Decisions under Risk: Six Clarifications," *Judgment and Decision Making,*, 2013, *8* (3), 214–235.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," *The Quarterly Journal of Economics*, 2018, p. https://doi.org/10.1093/qje/qjy013.

Fehr-Duda, Helga and Thomas Epper, "Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences," *Annual Review of Economics*, 2012, *4* (1), 567–593.

Frederick, Shane, "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 2005, *19* (4), 25–42.

Freeman, David, Yoram Halevy, and Terri Kneeland, "Eliciting Risk Preferences using Choice Lists," 2015. University of British Columbia, *mimeo.*

Friedman, Daniel, R Mark Isaac, Duncan James, and Shyam Sunder, *Risky Curves: On the Empirical Failure of Expected Utility*, Routledge, 2014.

Frydman, Cary, Colin Camerer, Peter Bossaerts, and Antonio Rangel, "MAOA-L Carriers are Better at Making Optimal Financial Decisions under Risk," *Proceedings of the Royal Society of London B: Biological Sciences*, 2011, *278* (1714), 2053–2059.

Gächter, Simon, Eric J. Johnson, and Andreas Herrmann, "Individual-level Loss Aversion in Riskless and Risky Choices," 2007. *mimeo.*

Gal, David and Derek D Rucker, "The Loss of Loss Aversion: Will it Loom Larger than its Gain?," *Journal of Consumer Psychology*, 2018, *28* (3), 497–516.

Gillen, Ben, Erik Snowberg, and Leeat Yariv, "Experimenting with Measurement Error: Techniques and Applications from the Caltech Cohort Study," *Journal of Political Economy*, Forthcoming.

Gneezy, Uri and Jan Potters, "An Experiment on Risk Taking and Evaluation Periods," *The Quarterly Journal of Economics*, 1997, *112* (2), 631–645.

Goette, Lorenz, David Huffman, and Ernst. Fehr, "Loss Aversion and Labor Supply," *Journal of the European Economic Association*, 2004, *2* (2-3), 216–228.

_ , Thomas Graeber, Alexandre Kellogg, and Charles Sprenger, "Heterogeneity of Loss Aversion and Expectations-Based Reference Points," *Working Paper*, 2018.

Hardie, Bruce G.S. Eric J. Johnson, and Peter S. Fader, "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice," *Marketing Science*, 1993, *12* (4), 378–394.

Harrison, Glenn W. Morten I. Lau, and E. Elisabet Rutström, "Estimating Risk Attitudes in Denmark: A Field Experiment," *The Scandinavian Journal of Economics*, 2007, *109* (2), 341–368.

_ , _ , and Melonie B. Williams, "Estimating Individual Discount Rates in Denmark: A Field Experiment," *The American Economic Review*, 2002, *92* (5), 1606–1617.

Hey, John and Jinkwon Lee, "Do Subjects Separate (or Are They Sophisticated)?," *Experimental Economics*, September 2005, *8* (3), 233–265.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *The American Economic Review*, 2002, *92* (5), 1644–1655.

Imai, Taisuke and Colin F. Camerer, "Estimating Time Preferences from Budget Set Choices Using Optimal Adaptive Design," 2018. Ludwig Maximilian University of Munich, *mimeo.*

Jacobson, Sarah and Ragan Petrie, "Learning from Mistakes: What Do Inconsistent Choices Over Risk Tell Us?," *Journal of Risk and Uncertainty*, 2009, *38* (2), 143–158.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

_ and Dan Lovallo, "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking," *Management Science*, 1993, *39* (1), 17–31.

Karni, Edi and Zvi Safra, "'Preference Reversal' and the Observability of Preferences by Experimental Methods," *Econometrica*, 1987, *55* (3), 675–685.

Kirby, Kris N., "One-year Temporal Stability of Delay-Discount Rates," *Psychonomic Bulletin & Review*, 2009, *16* (3), 457–462.

Kirby, Kris N. Ricardo Godoy, Victoria Reyes-Garcıa, Elizabeth Byron, Lilian Apaza, William Leonard, Eddy Perez, Vincent Vadez, and David Wilkie, "Correlates of Delay-Discount Rates: Evidence from Tsimane'Amerindians of the Bolivian Rain Forest," *Journal of Economic Psychology*, 2002, *23* (3), 291–316.

Kirby, Nielsen, "Dynamic Risk Preferences under Paper and Realized Outcomes," *Journal of Economic Behavior & Organization*, forthcoming.

Köbberling, Veronika and Peter Wakker, "An Index of Loss Aversion," *Journal of Economic Theory*, 2005, *122*, 119–131.

Kornotis, George M. and Alok Kumar, "Cognitive Abilities and Financial Decisions," in H. Kent Baker and John Nofsinger, eds., *Behavioral Finance*, Hoboken, NJ: John Wiley & Sons, Inc., 2010.

Kőszegi, Botond and Matthew Rabin, "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

_ **and** _ , "Reference-Dependent Risk Attitudes," *The American Economic Review*, 2007, *97* (4), 1047–1073.

Krajbich, Ian, Colin Camerer, and Antonio Rangel, "Exploring the Scope of Neurometrically Informed Mechanism Design," *Games and Economic Behavior*, 2017, *101*, 49–62.

Kujala, Janne V. and Tuomas J. Lukka, "Bayesian Adaptive Estimation: The Next Dimension," *Journal of Mathematical Psychology*, 2006, *50* (4), 369–389.

Kurata, Hiroshi, Hiroshi Izawa, and Makoto Okamura, "Non-Expected Utility Maximizers Behave as if Expected Utility Maximizers: An Experimental Test," *Journal of Socio-Economics*, 2009, *38* (4), 622–629.

Lesmes, Luis Andres, Seong-Taek Jeon, Zhong-Lin Lu, and Barbara Anne Dosher, "Bayesian Adaptive Estimation of Threshold Versus Contrast External Noise Functions: The Quick TvC Method," *Vision Research*, 2006, *46* (19), 3160–3176.

Levin, Irwin P., Stephanie S. Hart, Joshua A. Weller, and Lyndsay A. Harshman, "Stability of Choices in a Risky Decision-Making task: A 3-year Longitudinal Study with Children and Adults," *Journal of Behavioral Decision Making*, 2007, *20* (3), 241–252.

Lewi, Jeremy, Robert Butera, and Liam Paninski, "Sequential Optimal Design of Neurophysiology Experiments," *Neural Computation*, 2009, *21* (3), 619–687.

Lönnqvist, Jan-Erik, Markku Verkasalo, Gari Walkowitz, and Philipp C. Wichardt, "Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison," *Journal of Economic Behavior & Organization*, 2015, *119*, 254–266.

Loomes, Graham and Robert Sugden, "Disappointment and Dynamic Consistency in Choice under Uncertainty," *The Review of Economic Studies*, 1986, *53* (2), 271–282.

McFadden, Daniel, "Economic Choices," *The American Economic Review*, 2001, *91* (3), 351–378.

Mehra, Rajnish and Edward C. Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145–161.

Meier, Stephan and Charles D. Sprenger, "Temporal Stability of Time Preferences," *Review of Economics and Statistics*, 2015, *97* (2), 273–286.

Moffatt, P., "Optimal Experimental Designs in Models of Decision and Choice," in M. Boumans, ed., *Measurement in Economics: A Handbook*, Elsevier, 2007.

Mullainathan, Sendhil and Eldar Shafir, *Scarcity: Why Having Too Little Means so Much*, New York: Times Books, Henry Holt and Company, LLC, 2013.

Müller, Peter, Don A. Berry, Andy P. Grieve, Michael Smith, and Michael Krams, "Simulation-Based Sequential Bayesian Design," *Journal of Statistical Planning and Inference*, 2007, *137* (10), 3140–3150.

O'Donoghue, Ted and Charles Sprenger, "Reference-Dependent Preferences," in Bernheim B. Douglas, DellaVigna Stefano, and Laibson David, eds., *Handbook of Behavioral Economics: Foundations and Applications*, Vol. 1, Elsevier, 2018.

Oechssler, Jörg, Andreas Roider, and Patrick W. Schmitz, "Cognitive Abilities and Behavioral Biases," *Journal of Economic Behavior & Organization*, 2009, *72*, 147–152.

Ong, Qiyan, Walter Theseira, and Irene Y.H. Ng, "Reducing Debt Improves Psychological Functioning and Changes Decision-making in the Poor," *Proceedings of the National Academy of Sciences*, 2019.

Pachur, Thorsten, Michael Schulte-Mecklenbeck, Ryan O Murphy, and Ralph Hertwig, "Prospect Theory Reflects Selective Allocation of Attention," *Journal of Experimental Psychology: General*, 2018, *147* (2), 147–169.

Paninski, Liam, "Asymptotic Theory of Information-Theoretic Experimental Design," *Neural Computation*, 2005, *17* (7), 1480–1507.

Peirce, Charles S., "Note on the Theory of the Economy of Research," 1879. United States Coast Survey for the fiscal year ending June 1876, US Government Printing Office 1879, reprinted in *Operations Research*, 1967, vol. XV, 643–648.

Pew Research Center, *Evaluating Online Nonprobability Surveys*, www.pewresearch.org, May 2016.

Polania, Rafael, Michael Woodford, and Christian C. Ruff, "Efficient Coding of Subjective Value," *Nature Neuroscience*, 2019, *22* (1), 134–142.

Pope, Devin and Uri Simonsohn, "Round Numbers as Goals: Evidence from Baseball, SAT Takers, and the Lab," *Psychological Science*, 2011, *22* (1), 71–79.

Ray, Debajyoti, "Efficient Methods for Empirical Tests of Behavioral Economics Theories in Laboratory and Field Experiments." PhD dissertation, California Institute of Technology 2015.

Rees-Jones, Alex, "Quantifying Loss-Averse Tax Manipulation," *The Review of Economic Studies*, 2017, *85* (2), 1251–1278.

Ryan, Elizabeth G. Christopher C. Drovandi, James M. McGree, and Anthony N. Pettitt, "A Review of Modern Computational Algorithms for Bayesian Optimal Design," *International Statistical Review*, 2016, *84* (1), 128–154.

Schildberg-Hörisch, Hannah, "Are Risk Preferences Stable?," *Journal of Economic Perspectives*, 2018, *32* (2), 135–54.

Schmidt, Ulrich and Stefan Traub, "An Experimental Test of Loss Aversion," *Journal of Risk and Uncertainty*, 2002, *25* (3), 233–249.

Shamosh, Noah A and Jeremy R Gray, "Delay Discounting and Intelligence: A Meta-Analysis," *Intelligence*, 2008, *36* (4), 289–305.

Snowberg, Erik and Leeat Yariv, "Testing the Waters: Behavior across Subject Pools," 2018. NBER Working Paper #24781.

Sokol-Hessner, Peter, Ming Hsu, Nina G. Curley, Mauricio R. Delgado, Colin F Camerer, and Elizabeth A. Phelps, "Thinking Like a Trader Selectively Reduces Individuals' Loss Aversion," *Proceedings of the National Academy of Sciences*, 2009, *106* (13), 5035–5040.

Sprenger, Charles, "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Journal of Political Economy*, 2015, *123* (6), 1456–1499.

Stango, Victor and Jonathan Zinman, "We are all Behavioral, More or Less: Measuring and using Consumer-level Behavioral Sufficient Statistics," 2019. Dartmouth College, *mimeo.*

Stewart, Neil, "Decision by Sampling: The Role of the Decision Environment in Risky Choice," *The Quarterly Journal of Experimental Psychology*, 2009, *62* (6), 1041–1062.

Tanaka, Tomomi, Colin F Camerer, and Quang Nguyen, "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *The American Economic Review*, 2010, *100* (1), 557–71.

Thaler, Richard, "Mental Accounting and Consumer Choice," *Marketing Science*, 1985, *4* (3), 199–214.

Toubia, Olivier, Eric Johnson, Theodoros Evgeniou, and Philippe Delquié, "Dynamic Experiments for Estimating Preferences: An Adaptive Method of Eliciting Time and Risk Parameters," *Management Science*, 2013, *59* (3), 613–640.

_ , John R. Hauser, and Duncan I. Simester, "Polyhedral Methods for Adaptive Choice-Based Conjoint Analysis," *Journal of Marketing Research*, 2004, *41* (1), 116–131.

Tovar, Patricia, "The Effects of Loss Aversion on Trade Policy: Theory and Evidence," *Journal of International Economics*, 2009, *78* (1), 154–167.

Tversky, Amos and Daniel Kahneman, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–323.

Urminsky, Oleg and Gal Zauberman, "The Psychology of Intertemporal Preferences," in Gideon Keren and George Wu (Editor), eds., *Handbook of Judgment and Decision Making*, Vol. 2, Wiley-Blackwell, 2016, pp. 141–181.

von Gaudecker, Hans-Martin, Arthur Van Soest, and Erik Wengström, "Heterogeneity in Risky Choice Behavior in a Broad Population," *The American Economic Review*, 2011, *101* (2), 664–694.

Wakker, Peter P., *Prospect Theory: For Risk and Ambiguity*, Cambridge University Press, 2010.

Walasek, Lukasz and Neil Stewart, "How to Make Loss Aversion Disappear and Reverse: Tests of the Decision by Sampling Origin of Loss Aversion.," *Journal of Experimental Psychology: General*, 2015, *144* (1), 7–11.

_ , Timothy L. Mullett, and Stewart Neil., "A Meta-Analysis of Loss Aversion in Risky Contexts," in progress. *mimeo.*

Wald, Abraham, *Statistical Decision Functions*, John Wiley and Sons, New York; Chapman and Hall, London., 1950.

Wang, Stephanie W. Michelle Filiba, and Colin F. Camerer, "Dynamically Optimized Sequential Experimentation (DOSE) for Estimating Economic Preference Parameters," 2010. *mimeo.*

Wölbert, Eva and Arno Riedl, "Measuring Time and Risk Preferences: Reliability, Stability, Domain Specificity," 2013. *CESifo Working Paper Series No. 4339.*

Yechiam, Eldad, "Acceptable Losses: The Debatable Origins of Loss Aversion," *Psychological Research*, 2018, *https://doi.org/10.1007/s00426-018-1013-8.*

_ and Guy Hochman, "Losses as Modulators of Attention: Review and Analysis of the Unique Effects of Losses over Gains," *Psychological Bulletin*, 2013, *139* (2), 497–518.

Online Appendix—Not Intended for Publication

A Risk and Time Preferences and the Literature

Our findings regarding risk aversion and discounting are broadly similar to those of previous studies in representative populations; the few differences appear to be explained by our elicitation method. Risk aversion and discounting are widespread amongst our survey participants, but the median level of risk aversion is lower than found in the previous literature—a difference explained by our use of binary choice questions rather than MPLs. As discussed in Section 3.3, our elicitation method also explains the one major difference with previous studies of the correlates of these two preferences: by accounting for variation in choice consistency, we identify a strong negative relationship between risk aversion and cognitive ability.

Differences in elicitation method appear to explain the lower level of risk aversion estimated by DOSE than found in previous studies of representative samples. The mean and median Coefficient of Relative Risk Aversion (1-ρ) are 0.25 and 0.28 respectively, compared to previous findings ranging from approximately 0.4 (Dohmen et al., 2010) to 0.7 (Harrison et al., 2007; Andersen et al., 2008a). This pattern is consistent with laboratory studies finding lower levels of risk aversion using binary choice questions: the median coefficient using DOSE is 0.05 in the lab and 0.09 in the student survey—similar to the value of 0.12 found by Sokol-Hessner et al. (2009) using binary questions, but much lower than the range of 0.3–0.5 found by Holt and Laury (2002). Moreover, the median coefficients on the MPLs on our survey (0.4 and 2.1) are more in line with previous studies.

The patterns of correlation between risk aversion and discounting and sociodemographic characteristics we find (see Table 1), largely match the literature. The relationship between cognitive ability and patience, unlike risk aversion, is well-established in both economics and psychology (for example, Shamosh and Gray, 2008; Burks et al., 2009; Dohmen et al., 2010;

Benjamin et al., 2013). Patient individuals have also been found to have higher income, greater savings, and more education (see DellaVigna and Paserman, 2005; Falk et al., 2018; Urminsky and Zauberman, 2016). In the laboratory most studies have documented that women are more risk averse than men (Eckel and Grossman, 2008), as have Falk et al. (2018) in a representative sample. There is also some evidence of a negative relationship between risk aversion and income, although results have been mixed (see, for instance Dohmen et al., 2010; Barsky et al., 1997). The most important difference from the literature is the relationship between risk aversion and cognitive ability, where we find a strong negative correlation, whereas results in the previous literature are mixed between somewhat smaller negative correlations, no correlations and, occasionally positive correlations (for a summary see Andersson et al., 2016b, Figure 1). As discussed in Section 3.3, it appears this difference is explained by the fact DOSE accounts for inconsistent choice.

Measurement error may also explain the lack of consistent patterns emerging from the few other studies that have examined the correlates of loss aversion in representative samples.[1] Only one of those studies examined the association with cognitive ability, finding no evidence of a relationship (Andersson et al., 2016a). The findings for both education and income have been very mixed, with correlations sometimes negative, sometimes zero and—for income—sometimes positive. The most consistent pattern emerging from other studies, but not reflected in our results, is that women have been found to be more loss averse—whereas we find no relationship with sex after controlling for other sociodemographic variables.

As noted in Section 3.4, the over-time correlation of the DOSE time preference measure is larger than estimates in most previous studies, but direct comparisons are complicated by differences in the sample used. In the most comparable study, Meier and Sprenger (2015) report correlations of 0.36 for present bias and 0.25 for discounting parameters among 250 low- to middle-income Americans. In another field study, Kirby et al. (2002) reports correlations of 0.09–0.23 over a six month period among Bolivian Amerindians. The only

[1] As discussed in Section 1, those studies include Booij and Van de Kuilen (2009); Booij et al. (2010); von Gaudecker et al. (2011) and Andersson et al. (2016a)

study (Kirby, 2009) that finds a higher correlation than DOSE (between 0.63 and 0.71) took place in a more controlled (laboratory) environment than our survey. The variety of the samples makes comparisons difficult—it is not clear, for instance, how to compare our representative online survey to the in-person, low-income sample in Meier and Sprenger (2015). However, the results are, at least, consistent with the DOSE estimates being more stable over time due to reduced measurement error.[2]

B DOSE Procedure and Survey Implementation

This Appendix presents details of the DOSE implementation we used to estimate loss aversion in the U.S. population. We start by explaining the technical features of DOSE, focusing on the choices experimenters can make to tune the procedure for other applications. We then describe the elicitations we used in our online survey.

B.1 Technical Details of DOSE Procedure

DOSE can be customized for particular research questions. The main objects of choice for a researcher are the parametric specification(s), the prior distribution over parameters or models, the set of choices to present to participants, how parameters map to choices—that is, the structure of possible mistakes—and the information criterion used to select the next question based on current beliefs. This subsection details the choices we made to elicit risk aversion, loss aversion, and discounting.

[2] Chuang and Schechter (2015) provide a detailed review of previous studies of stability of risk or time preferences. They document two additional studies that reported correlations from incentivized measures over short periods of time. Dean and Sautmann (2014) find correlations of up to 0.67 over a one week period in Mali. Wölbert and Riedl (2013) report correlations of between 0.36 and 0.68 for 20 risk MPLs, and between 0.61 and 0.68 for three measures of discount rates over a 5–10 week period.

B.1.1 Utility Function and Priors over Parameters

We elicit risk and loss aversion using a Prospect Theory utility function with power utility (Kahneman and Tversky, 1979). This utility function assumes that participants value payments relative to a reference point, which we assume is zero. The standard S-shaped utility function in Prospect Theory implies that, for common parameter values, participants are risk averse over positive payments (gains), and risk loving over negative payments (losses). A kink in the utility function at zero represents loss aversion. Formally:

$$v(x, \rho_i, \lambda_i) = \begin{cases} x^{\rho_i} & \text{for } x \geq 0 \\ -\lambda_i(-x)^{\rho_i} & \text{for } x < 0, \end{cases} \tag{3}$$

in which λ_i parameterizes loss aversion, ρ_i parameterizes risk aversion, and $x \in \mathbb{R}$ is a monetary outcome relative to the reference point. If $\lambda_i > 1$, then the participant is loss averse. If $\lambda_i < 1$, then the participant is loss tolerant. An individual with $\rho_i < 1$ demonstrates risk aversion over gains and risk love over losses.

The specification in (3) focuses on accurately estimating loss aversion with as few questions as possible. It does not allow for other common features of Prospect Theory: probability weighting or differential curvature of the utility function over losses and gains. The lotteries in our questions are further designed to minimize probability distortions, as all have 50/50 probabilities of two outcomes. Moreover, as most studies have found limited difference in curvature across the two domains (see Booij et al., 2010, Table 1), we impose the same utility curvature for both gains and losses.[3] This improves the accuracy of the estimates of λ_i as questions are not selected to separately identify the curvature in the loss and gain domain. However, it is quite simple to allow for different curvatures even after conducting

[3]Assuming the same curvature across gains and losses also avoids an issue with power utility: different curvatures mean that estimates of loss aversion depend on scaling (Köbberling and Wakker, 2005). Moreover, there will always be an amount x s.t. $U(x) > U(-x)$ (Wakker, 2010). Our results are similar using the exponential (CARA) utility function Köbberling and Wakker (2005) suggest to avoid these issues (see Section 4.1).

the experiment and, as shown in Section 4.1, the assumption of equal curvature is supported by the data our results are robust to estimating this alternative specification, and that .

Time discounting is modeled with a standard monthly discount factor and the power utility function in (3). Utility from the perspective of the survey date is given by $u(x_t, \rho_i, \delta_i) = \delta_i^t x_t^{\rho_i}$, where δ_i is a discount factor and ρ_i captures the curvature of the utility function from (3), t is the time from the survey date in months, and x_t is a payment at time t.

While researchers must choose a parametric specification and prior distribution for data collection, an alternative prior or specification can be used to calculate parameter values ex post. The experimenter's initial choices are used only to generate the personalized question sequence: once participant choices are recorded, any other prior distribution or parametric specification can be used to derive new estimates from the data.

B.1.2 Mistakes and Choice Consistency

An important feature of DOSE is that, when selecting the personalized sequence of questions, it takes into account the possibility that participants make mistakes. The process by which mistakes are made must be parametrically modeled for DOSE to account for it. We model the mapping between utility and choices using the logit function, which has been widely used in both economics and psychophysics due to its connection with the random utility model.[4] For any choice between options o_1 and o_2 with $V(o_1) > V(o_2)$:

$$\text{Prob}[o_1] = \frac{1}{1 + e^{-\mu_i (V(o_1) - V(o_2))}}. \tag{4}$$

The logit function depends both on the utility difference between options o_1 and o_2 and the choice *consistency* parameter $\mu_i \in \mathbb{R}^+$. The probability of making a mistake—that is, not choosing the value-maximizing option—is $1 - \text{Prob}[o_1]$. This is decreasing in the value

[4] Specifically, choice probabilities will be logit if the errors in the random utility model have an Extreme Value Type I distribution. See McFadden (2001) for a broader discussion of the history of the logit specification and its properties. DOSE can easily be implemented with multi-answer question using a multinomial logit or alternative probabilistic choice function.

difference between o_1 and o_2. This decrease is more rapid when μ_i is larger, so higher values of this parameter represent greater consistency in choices.

The set of questions can be designed to reduce the likelihood of mistakes. Our options were constructed to make expected value comparisons as simple as possible. All of our questions include only two options, with only one of these being a non-degenerate lottery. Each lottery has only two possible prizes. One of three payoffs (the sure payoff, and the two lottery payoffs) is always zero. Choices are thus between either a lottery with a zero payoff and some gain, versus some (possibly negative) sure amount; or between a lottery with a gain and loss, versus a sure amount of zero. In the former case, the expected value can be found by dividing by two. In the latter, one can ascertain if the expected value of the lottery is greater or less than zero by comparing the size of the positive and negative payoffs.

We believe these questions are also unlikely to produce inadvertent reference points, as MPLs have been shown to do (Sprenger, 2015; Chapman et al., 2017). However, this is also a testable prediction: we can try to fit specifications with alternative reference points, such as incorporating the endowment of points given at the beginning of the DOSE module. As shown in Section 4.2, this model produces a much worse fit: it only predicts 54% of choices correctly, whereas our main specification predicts 88% correctly.

B.1.3 Information Criterion

In our implementation, DOSE selects each question to maximize the expected Kullback-Leibler (KL) divergence between the prior and possible posteriors associated with each answer. That is, the question that is picked at each point is the one with the highest expected information gain given the initial prior and previous answers. The KL criterion has been used widely in the optimal design literature in statistics due to its conceptual simplicity and grounding in information theory (see Ryan et al., 2016, for a discussion and examples). Further, the information maximization approach leads to consistent and efficient parameter estimates under weak modeling conditions (Paninski, 2005). However, DOSE is easily modified to incorporate alternative information criteria—for example, Imai and Camerer (2018) use DOSE with the EC^2 criterion to discriminate between models of time preferences.

Formally, consider a finite set of possible parameter vectors θ_k for $k = 1, ..., K$, where each $\theta_k = (\rho_k, \lambda_k, \delta_k, \mu_k)$ is a combination of possible values of the parameters of interest.[5] Each θ_k has an associated probability p_k of being the correct parameters. In the first question, these probabilities are the priors chosen by the experimenter; they are then updated in each round according to the participant's answers. The expected Kullback-Leibler divergence between the prior and the posterior when asking question Q_j is:

$$KL(Q_j) = \sum_{k \leq K} \sum_{a \in A} \log \left(\frac{l_k(a; Q_j)}{\sum_{j \in \mathcal{K}} p_j l_j(a; Q_j)} \right) p_k l_k(a; Q_j) \tag{5}$$

where $a \in A$ are the possible answers to the question, and $l_k(a; Q_j)$ is the likelihood of answer a given θ_k—in our implementation this is determined by the logit funciton in (4). DOSE selects the question that maximizes $KL(Q)$, the participant answers it, model posteriors are updated, the question Q_j that now maximizes $KL(Q)$ is selected, and so on.[6]

B.2 Potential for Manipulation

The concern that DOSE could—like most dynamic designs—be strategically manipulated by participants is unlikely to be important in practice. The adaptive nature of the DOSE question selection procedure means that individuals could have an incentive to misrepresent their true preferences in early questions to obtain more generous offers in future questions. For example, participants could misleadingly say they prefer a lottery to a sure amount

[5] We assume a finite space of parameters for computational ease. The survey questions were selected using the Kullback-Leibler information criterion suggested by El-Gamal and Palfrey (1996). Their variant maximizes the distance between posteriors (information) obtained under different models, whereas Equation (5) maximizes the information about parameters. In addition, in the survey procedure, after each question round the joint posterior was used to construct marginal distributions for each of the parameters. These updated marginal distributions were then used to construct the new probability distribution used for question selection in the following question round under the assumption that the distributions were independent.

[6] In the survey, we restrict the procedure to only consider questions that had not yet been asked of that participant. In order to improve the estimate of μ, the procedure would eventually ask the same question multiple times.

in the first question in order to increase the magnitude of the sure amounts offered in the future. However, such behavior is unlikely as the incentives for manipulation are small, and it requires that participants understand both the adaptive nature of the procedure, which is not explained to them, and how to manipulate the question sequence. In the laboratory there is little evidence that subjects attempt manipulation, and such behavior is even less likely in our representative survey where participants are unlikely to have any experience of economic experiments.

DOSE can easily be adapted to address concerns about manipulation, however laboratory evidence obtained during piloting suggests that doing so is unlikely to lead to different results. We used a 10-question DOSE procedure to implement the risk aversion measure developed by Holt and Laury (2002), and tested whether using an incentive compatible payment method affected the estimates obtained.[7] Specifically, 23 participants were paid according to a randomly selected question (as in our survey), and 22 were paid using an incentive compatible method suggested by Ian Krajbich. Under the latter, DOSE parameter estimates were used to project a person's choice on an unanswered question. This question was then paid. The distribution of estimates was similar across the payment methods, with no statistically significant difference in means (p-value$= 0.40$).

We also do not observe substantial differences between the results of our student survey and those of two laboratory studies that implemented incentive compatible DOSE procedures.[8]. The first of these studies, Krajbich et al. (2017), implemented the procedure described in the previous paragraph. The second, carried out by Alec Smith at UCLA, implemented an alternative incentive compatible procedure suggested by Kate Johnson.[9] In this approach the actual question chosen for payment was randomly selected from *all* possible questions after the personalized question sequence is completed. If that question had already

[7]In this measure, participants face a series of choices between two lotteries, with the probability of the high payoff varying between questions. Holt and Laury (2002) present these choices as a Multiple Price List, whereas DOSE chooses the questions based on previous answers. See Wang et al. (2010) for further details of the implementation.

[8]For further discussion of these previous laboratory studies, see Appendix E

[9]This experiment was not published: we are grateful to Alec Smith for providing the data.

been answered, the answer determined the payment. If not, the participant answered it, and this answer determined the payoff. The level of loss aversion in the lab was similar to our survey: across the two lab experiments, 8% of participants were loss tolerant (compared to 10% in our survey) and the median λ was 1.95 (compared to 1.84). It does not appear that incentive compatible implementation led to changed behavior in these student subject pools.

Strategic behavior is even more unlikely in our survey than in these laboratory samples, given the short question sequence and that participants are unlikely to have previous experience of economic experiments. We draw from a general population sample who—unlike the subject pools used in laboratory experiments—are unlikely to have received formal training in economics or to have participated in previous incentivized studies. Participants also had little opportunity to learn during the survey itself, as each DOSE module contained only ten questions. Further, behavior does not obviously change between survey waves, as we would expect if participants identified the potential for manipulation during the first wave and then acted strategically when taking the survey a second time. The distribution of preferences is similar across survey waves (Figure F.5), and our preference parameters are stable over time (Section 3.4). There is thus no reason to believe that participants attempted to manipulate the question selection procedure.

Ray (2015) examines potential manipulation in more detail when using an adaptive method similar to ours, and also concludes that participants did not engage in strategic behavior. In particular, based on a post-experiment survey, he reports that most participants did not even try to manipulate the algorithm, while the few that did attempt manipulation were unable to find an effective method. Participants that were informed that the algorithm was adaptive did not receive higher average earnings, and he also finds little behavioral evidence of manipulation (as identified by either choice patterns or response times changing significantly between early and later question rounds). A likely explanation for the lack of behavioral change is simply that the benefits of doing so are not large: with a 10 question adaptive sequence the excess earnings of a risk neutral clairvoyant agent are only 8% higher than a myopic one who maximizes earnings in each choice.

Future researchers interested in incentive (in)compatibility in dynamic processes might also examine adapting the DOSE procedure to directly assess the extent of strategic manipulation among participants. In particular, the possibility of manipulation could be built into DOSE as a separate theory of behavior with associated prior beliefs. The DOSE questions would then be selected in order to identify whether there is strategizing or not (as well as the other parameters of interest).

There are, however, important drawbacks to incentive compatible designs. They are complicated to explain and, if they rely on computer choices, may not be credible to participants. Further, there may be a trade-off between the incentive for truthful response and the strength of incentives—the two incentive compatible payment procedures above, for instance, mean that each question has a lower probability of influencing the final payoff.

B.3 DOSE in a Representative Survey

We now turn to the practical details of implementing DOSE in two waves of a large, representative, incentivized survey of the U.S. population. The survey includes two DOSE modules—one relating to risk and one to time preferences—as well as other behavioral elicitations, and cognitive and sociodemographic questions.

B.3.1 DOSE Modules

Both of the DOSE modules were comprised of ten questions selected using the procedure described in Appendix B.1, slightly modified due to the practicalities associated with implementing the survey online. The design of YouGov's online platform precluded using DOSE to choose questions in real time and so, instead, simulated responses were used to map out all possible sets of binary choices in advance. That tree was then used to route participants through the survey. Mapping such a tree with a refined prior was infeasible given both computational constraints and the limitations of YouGov's interface (mapping such a tree over 20 questions would involve over 500,000 routes through the survey). As such, questions were selected using a coarser prior and then final individual-level estimates were obtained by performing the Bayesian updating procedure with a joint 100-point discretized uniform prior.[10]

We now describe the particular design choices—the priors, utility specification, and question set—that we made for each of the two DOSE modules.

Risk Preferences: The first DOSE module elicited risk and loss aversion. Participants were given 10,000 points and offered a sequence of ten binary choices between a 50:50 lottery and a sure amount. Two types of lottery were used. The first had a 50% chance of 0 points, and a 50% chance of winning a (varying) positive amount of points (of up to 10,000). The second had a 50% chance of winning an amount up to 10,000 points, and a 50% chance of a loss of up to 10,000 points. In the latter case, the sure amount was always 0 points.[11]

The question sequence for the risk preference module was selected using the specification in Equation (1) and a prior constructed using the estimates for laboratory participants obtained by Sokol-Hessner et al. (2009) and Frydman et al. (2011): 0.2–1.7 for ρ, and 0–4.6 for λ. To account for the survey environment we restricted the question selection procedure in two ways. First, to focus the procedure on obtaining a precise estimate of ρ before moving onto estimates of λ, the first four questions in the module were restricted to be lotteries over gains. Second, to make it harder for participants to identify the adaptive nature (and hence attempt to manipulate) the procedure, the maximum prize was restricted to be no more than 7,000 points in each even numbered round.

[10] The prior used for question selection included 12 mass points for ρ and δ, 20 for λ, and 4 for μ. To utilize the information about the curvature of the utility function from the risk-loss module, for the time preference module participants were assigned to one of ten prior distributions over ρ, based on their estimated ρ from the risk-loss module.

[11] The set of potential questions used for the risk module allowed for gains ranging between 1,000 and 10,000 points in 500 point increments, and sure amounts and losses varying ranging from 500 points to 10,000 points in 100 point increments. Questions were excluded if one choice was first order stochastically dominated for all values of the prior distribution. Questions were also selected as if the prize amounts were 3 times the actual amounts offered in the lottery to improve discrimination of the risk and loss aversion parameters.

Time Preferences: The second DOSE module elicited discount factors and refined the estimates of the curvature of the utility function elicited in the risk preference module. Participants were offered a sequence of ten binary choices between a lower amount of points at an earlier date (either the day of the survey, or in the future) or a higher amount at a later date (up to 90 days in the future). The maximum payment in each question was 10,000 points.[12]

The time discounting questions were selected accounting for both discounting and present bias.[13] However, there was little evidence of present bias in the survey—possibly due to the fact that points were, in general, not instantly convertible into consumption—and so time preferences were re-estimated allowing for discounting only. As with the risk module, some restrictions were placed on the question selection procedure. The first five questions were restricted to the choice between payment on two dates in the future. In addition, when considering two options in the future (that is, $t_1 > 0$ and $t_2 > 0$), individuals were assumed to choose as if they have a fixed value of the present bias parameter (β=0.64, based on the estimates from Tanaka et al. (2010)).

B.3.2 Additional Measures

We also utilize survey measures of cognitive ability and more standard elicitations of risk and time preferences.

Cognitive Ability: Cognitive ability was measured using a set of nine questions: six from the International Cognitive Ability Resource (ICAR, Condon and Revelle, 2014) and three from the Cognitive Reflection Test (CRT) developed by Frederick (2005).

The IQ test consisted of three questions from the ICAR matrix module and three from the 3-D rotational module. In the matrix questions, participants were presented with a 3x3

[12]Possible payment amounts in the time module ranged from 1,000 to 10,000 points in 1,000 point increments, and possible payment days were of 0, 1, 3, 5, 7, 9, 10, 12, 16, 21, 28, 35, 42, 49, 56, 60, 70, 80, 90 days after the survey. The question set included all combinations of these payment amounts and dates in which the early payment was less than the later payment.

[13]For both discounting (δ) and present bias (β) we used a prior ranging between 0.2 and 1.0.

matrix with 8 geometric designs. They then had to choose the correct design to complete the pattern from a list of 6 possible options. In the 3-D rotational questions, participants were shown a picture of three sides of a cube. They then had to choose which of six options (each also showing three sides of a cube) was compatible with a rotated version of the original.

The CRT includes three arithmetically straightforward questions with an instinctive, but incorrect, answer. The test thus measures the tendency for an individual to reflect upon a question rather than answer instinctively.

Risk Aversion MPLs: Two MPLs asked participants to choose between a fixed 50/50 lottery and a series of ascending sure amounts. The row in which the participant first chose the sure amount identified a range of possible certainty equivalents for the lottery—we use the midpoint of this range. There were two MPLs of this type: the first had a 50/50 lottery over 0 and 10,000 points, the second, a 50/50 lottery over 2,000 and 8,000 points.[14]

Time Preference MPLs: In addition to the DOSE module, the survey included two MPLs to elicit time preferences. The first time MPL elicited the amount of points that the participant valued the same as 6,000 points 45 days later. The second MPL elicited the amount of points in 45 days that the participant valued the same as 6,000 points in 90 days. This measure is used primarily in Section 3.2.

B.4 Screenshots

This subsection contains screenshots of all the types of questions analyzed in this paper. Full design documents and screenshots can be found at hss.caltech.edu/∼snowberg/wep.html.

[14]These MPLs were designed to elicit the willingness to pay (WTP) for a lottery; participants were endowed with a fixed set of points and asked how much they would exchange for the lottery. An additional two MPLs in the survey elicited the willingness to accept (WTA) for the same lotteries; that is, participants were endowed with the lottery and asked the amount of points that they would need to be willing to exchange the lottery. As shown in Appendix F.2, the results are similar measuring risk aversion using the WTA MPLs. See Chapman et al. (2017) for more details of the design of these measures, as well as extensive discussion of the relationship between WTA, WTP, and other risk elicitations in the survey.

Figure B.1: DOSE Risk/Loss Aversion Instruction Screen

Section 4 of 12

In the next few questions, you will be asked to choose between two lotteries.

You will start this section with 10,000 points, which you may lose based on the lotteries you choose in this section. That is, some of the lotteries in this section may both **add** to or **subtract** from this initial 10,000 points.

For example, suppose you chose a lottery that had a 50% chance of adding 5,000 points, and a 50% chance of subtracting 5,000 points. In the case of winning, the 5,000 will be added to your additional 10,000. In the case of a loss, the 5,000 will be subtracted from your initial 10,000. Note that you will never have the possibility of losing more than 10,000, so at worst you will end this section with 0 points.

Figure B.2: DOSE Risk/Loss Aversion Example Question: Gains Only

Which of the following options do you prefer?

- A lottery where you can either receive 10,000 points or receive 0 points, each with probability 50%;

OR

- Receiving 5,200 points for certain.

Figure B.3: DOSE Risk/Loss Aversion Example Question: Both Gains and Losses

Which of the following options do you prefer?

- A lottery where you can either receive 7,000 points or lose 6,300 points, each with probability 50%;

OR

- Receiving 0 points for certain.

Figure B.4: DOSE Discounting Instruction Screen

YouGov

The next few questions ask you to choose between amounts of points at different times, many of which are **in the future**. If one of these questions is selected for payment, the number of points displayed will be credited to your account on the day shown.

For your reference, today is April 17.

Figure B.5: DOSE Discounting Example Question

YouGov

Which of the following options do you prefer?

- 10,000 points put in your account 90 days from now (July 16)
- 9,750 points put in your account today

Figure B.6: Risky Project Question

YouGov

You are endowed with 2,000 points that you can choose to keep or invest in a risky project. Points that are not invested in the risky project are yours to keep. The risky project has a 40% (that is a 4 out of 10) chance of success.

- If the project is successful, you will receive 3 times the amount you chose to invest.
- If the project is unsuccessful, you will lose the amount invested.

Please choose how many points you want to invest in the risky project. Note that you can pick any number between 0 and 2,000, including 0 or 2,000.

Figure B.7: Multiple Price List to Measure Risk Aversion

YouGov

For this question, you **have been given 8,000 points**. You will be offered the opportunity to exchange some of these points for a lottery ticket. This lottery ticket has a **50% chance** of paying you **8,000 points**, and a **50% chance** of paying **2,000 points**.

For example, if you choose to pay 3,000 points for a lottery ticket, and this question is chosen for payment, you will:

- Pay 3,000 points for the lottery ticket
- Keep 5,000 points for yourself
- Earn whatever proceeds you get from the lottery ticket (if any)

For each row in the table below, which option would you prefer?

✓ Keep 8,000 points	or	Buy the lottery ticket for 8,000 points and keep the remaining 0 points
Keep 8,000 points	or	Buy the lottery ticket for 6,000 points and keep the remaining 2,000 points
Keep 8,000 points	or	Buy the lottery ticket for 5,500 points and keep the remaining 2,500 points
Keep 8,000 points	or	Buy the lottery ticket for 5,250 points and keep the remaining 2,750 points
Keep 8,000 points	or	Buy the lottery ticket for 5,000 points and keep the remaining 3,000 points
Keep 8,000 points	or	Buy the lottery ticket for 4,750 points and keep the remaining 3,250 points
Keep 8,000 points	or	Buy the lottery ticket for 4,500 points and keep the remaining 3,500 points
Keep 8,000 points	or	Buy the lottery ticket for 4,250 points and keep the remaining 3,750 points
Keep 8,000 points	or	Buy the lottery ticket for 4,000 points and keep the remaining 4,000 points
Keep 8,000 points	or	Buy the lottery ticket for 3,750 points and keep the remaining 4,250 points
Keep 8,000 points	or	Buy the lottery ticket for 3,500 points and keep the remaining 4,500 points
Keep 8,000 points	or	Buy the lottery ticket for 3,250 points and keep the remaining 4,750 points
Keep 8,000 points	or	Buy the lottery ticket for 3,000 points and keep the remaining 5,000 points
Keep 8,000 points	or	Buy the lottery ticket for 2,500 points and keep the remaining 5,500 points
Keep 8,000 points	or	✓ Buy the lottery ticket for 2,000 points and keep the remaining 6,000 points

Reset

Autofill

Review the instructions

Figure B.8: Risk Aversion MPL used in Robustness Checks (Appendix F.2)

YouGov

For this question, you are given a lottery ticket that has a **50% chance** of paying you **10,000 points,** and a **50% chance** of paying you **0 points**.

You have two options for this lottery ticket:

1. Keep it or
2. Sell it for a certain amount of points (for example, 2,000 points)

For each row in the table below, which option would you prefer?

The lottery ticket	or	Sell it for 0 points
The lottery ticket	or	Sell it for 1,000 points
The lottery ticket	or	Sell it for 2,000 points
The lottery ticket	or	Sell it for 2,500 points
The lottery ticket	or	Sell it for 3,000 points
The lottery ticket	or	Sell it for 3,250 points
The lottery ticket	or	Sell it for 3,500 points
The lottery ticket	or	Sell it for 3,750 points
The lottery ticket	or	Sell it for 4,000 points
The lottery ticket	or	Sell it for 4,250 points
The lottery ticket	or	Sell it for 4,500 points
The lottery ticket	or	Sell it for 4,750 points
The lottery ticket	or	Sell it for 5,000 points
The lottery ticket	or	Sell it for 5,250 points
The lottery ticket	or	Sell it for 5,500 points
The lottery ticket	or	Sell it for 6,000 points
The lottery ticket	or	Sell it for 8,000 points
The lottery ticket	or	Sell it for 10,000 points

Reset

Autofill

Review the instructions

Figure B.9: Multiple Price List to Estimate Discount Rate (δ)

YouGov

For each row in the table below, which option would you prefer?

6,000 points in 45 days (June 1)	or	0 points today
6,000 points in 45 days (June 1)	or	1,000 points today
6,000 points in 45 days (June 1)	or	2,000 points today
6,000 points in 45 days (June 1)	or	3,000 points today
6,000 points in 45 days (June 1)	or	3,500 points today
6,000 points in 45 days (June 1)	or	4,000 points today
6,000 points in 45 days (June 1)	or	4,500 points today
6,000 points in 45 days (June 1)	or	5,000 points today
6,000 points in 45 days (June 1)	or	5,500 points today
6,000 points in 45 days (June 1)	or	5,600 points today
6,000 points in 45 days (June 1)	or	5,700 points today
6,000 points in 45 days (June 1)	or	5,800 points today
6,000 points in 45 days (June 1)	or	5,900 points today
6,000 points in 45 days (June 1)	or	5,950 points today
6,000 points in 45 days (June 1)	or	5,975 points today
6,000 points in 45 days (June 1)	or	6,000 points today
6,000 points in 45 days (June 1)	or	6,100 points today

Reset

Autofill

Review the instructions

Figure B.10: Multiple Price List Instruction Screen

YouGov

This survey often uses a special type of question. We want to help you answer these questions **quickly** and **accurately**.

This special type of question has many similar choices, as in the example below. The options on the left are always the same, while those on the right change — getting better and better.

If a question like this is picked for payment, **one row** will be selected, and you will be paid according to the choice **you made in that row**. It is important that your answers in each row **are accurate** so you will get the payment **you want**.

You will see a screen that looks like this.

☑ 5,000 points	or	☐ 0 points
☐ 5,000 points	or	☐ 1,000 points
☐ 5,000 points	or	☐ 2,000 points
☐ 5,000 points	or	☐ 3,000 points
☐ 5,000 points	or	☐ 4,500 points
☐ 5,000 points	or	☐ 5,500 points
☐ 5,000 points	or	☐ 6,000 points
☐ 5,000 points	or	☐ 7,000 points
☐ 5,000 points	or	☐ 8,000 points
☐ 5,000 points	or	☐ 9,000 points
☐ 5,000 points	or	☑ 10,000 points

>

Figure B.11: Multiple Price List Practice Screen

YouGov

To answer these types of questions **quickly** and **accurately** we suggest you:

1. Start by looking at the **top row**, and think carefully about each row in turn.
2. For each row where you **prefer the option on the left** over the option on the right, check the box on the left hand side.
3. When you find the **first question where you prefer the option on the right** over the option on the left, check the box on the right.
4. Notice that the option on the right is always better as you go down the list. This means that after you choose one option on the right, you should choose the option on the right for all rows below. Your answers should therefore "cross over" from left to right **only once**.
5. Once you have filled in the "cross over" point you may hit the Autofill button to fill in the rest of the chart faster. Alternatively, you may check every box manually.

All rows must have a box checked for you to continue to the next page

If you need to start over at any point, hit the **Reset** button to clear out all of the checkmarks.

Example question: For each row in the table below, which option would you prefer?

5,000 points	or	0 points
5,000 points	or	1,000 points
5,000 points	or	2,000 points
5,000 points	or	3,000 points
5,000 points	or	4,500 points
5,000 points	or	5,500 points
5,000 points	or	6,000 points
5,000 points	or	7,000 points
5,000 points	or	8,000 points
5,000 points	or	9,000 points
5,000 points	or	10,000 points

Reset Autofill

Review the instructions

Figure B.12: Multiple Price List Practice Screen with Error Message

YouGov

To answer these types of questions **quickly** and **accurately** we suggest you:

1. Start by looking at the **top row**, and think carefully about each row in turn.
2. For each row where you **prefer the option on the left** over the option on the right, check the box on the left hand side.
3. When you find the **first question where you prefer the option on the right** over the option on the left, check the box on the right.
4. Notice that the option on the right is always better as you go down the list. This means that after you choose one option on the right, you should choose the option on the right for all rows below. Your answers should therefore "cross over" from left to right **only once**.
5. Once you have filled in the "cross over" point you may hit the Autofill button to fill in the rest of the chart faster. Alternatively, you may check every box manually.

All rows must have a box checked for you to continue to the next page

If you need to start over at any point, hit the **Reset** button to clear out all of the checkmarks.

Example question: For each row in the table below, which option would you prefer?

You have indicated in row 2 that you prefer 1,000 points to 5,000 points. But 1,000 points is less than 5,000 points, which means you would get more by selecting 5,000 points. Please correct this.

In all the other questions on this survey, there is no right or wrong answer. However, you should make sure that you select the option that you prefer on **each line**.

☑ 5,000 points	or	☐ 0 points
☐ 5,000 points	or	☑ 1,000 points
☐ 5,000 points	or	☑ 2,000 points
☐ 5,000 points	or	☑ 3,000 points
☐ 5,000 points	or	☑ 4,500 points
☐ 5,000 points	or	☑ 5,500 points
☐ 5,000 points	or	☑ 6,000 points
☐ 5,000 points	or	☑ 7,000 points
☐ 5,000 points	or	☑ 8,000 points
☐ 5,000 points	or	☑ 9,000 points
☐ 5,000 points	or	☑ 10,000 points

Reset Autofill

Review the instructions

C Simulations using Laboratory Choices

In this appendix we demonstrate the benefits of DOSE's personalized question sequence using data from 120 subjects in two prior laboratory experiments.[15] In each of these experiments, participants were asked the same set of 140 binary choices over gains and losses—with each choice having the same structure as the questions used in our survey. A 20-question DOSE procedure obtains parameter estimates that are close to (within 15% of) parameter estimates after 140 questions. Further, a joint uniform prior over the parameters is close to optimal for question selection. The Bayesian procedure also requires many fewer questions to elicit individual-specific estimates of risk and loss aversion than Maximum Likelihood Estimation.

C.1 Estimates of Risk and Loss Aversion

In our simulation, we optimally order the questions for each participant using DOSE and compare the parameter estimates to those that would be obtained under a random question ordering. After DOSE selects a question, we provide it with the answer the participant gave in the experiment. The procedure then updates the probability distribution over parameters, selects the next question, and so on. This allows us to compare, question by question, the *inaccuracy*—the absolute distance from the true parameter value as a percentage of the true value—of DOSE's estimates with those elicited by a random question ordering. As we do not have access to true parameter values, we substitute the values one would obtain using the choices in all 140 questions.

A 20-question DOSE sequence provides a similar amount of information as about 50 randomly ordered questions, as shown in the top panel of Figure C.1.[16] The DOSE estimates of both risk and loss aversion are consistently closer to the final parameter estimate, indicating—under the assumption that the final estimate closely approximates an individual's true parameters—that the procedure provides accurate estimates considerably faster

[15] 90 participants come from Frydman et al. (2011) and 30 from Sokol-Hessner et al. (2009).

[16] For loss aversion, 45 randomly-ordered questions are needed to be as close to the final estimate as 20 DOSE questions. For risk aversion, 55 questions are required.

than selecting questions at random.[17] After 20 questions, the DOSE estimates are almost twice as close to the final estimate as those under a random question ordering (12% vs. 21–22%).

As shown in the bottom panel of Figure C.1, the DOSE estimates are also more highly correlated with the final estimates, an important feature when seeking to identify correlations between preferences and other population characteristics. After a 20 question DOSE sequence, the correlations are higher than obtained under the random ordering for both loss aversion (0.94 versus 0.87) and risk aversion (0.89 versus 0.70).

These simulations also show that using the uniform prior is close to optimal for question selection. To do so, we compare the performance of DOSE question selection using a uniform prior to that using an *optimal prior* constructed from the distribution of the estimates after 140 questions. To focus on the question selection impacts of the prior, we estimate the parameter values using a uniform prior in both cases. As shown in Figure C.1, both the accuracy and the correlations are similar whether using the optimal or uniform prior.

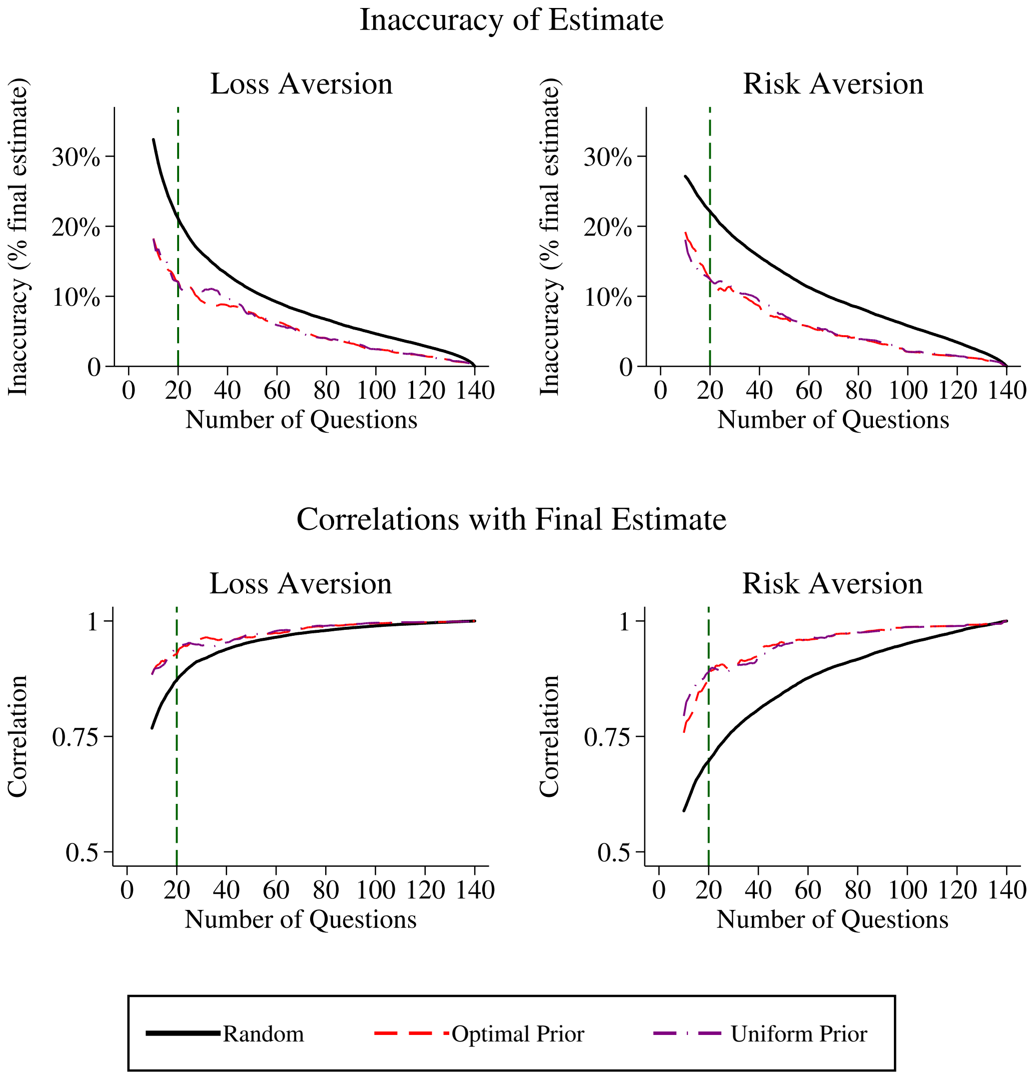
DOSE improves the accuracy of individual-specific estimates for the full distribution of parameter values, as shown in Figures C.2 and C.3. These figures display the progression towards the final parameter estimates after asking 10, 20, 50 and 100 DOSE questions. After just 10 questions, the estimates for both risk and loss aversion are clustered around the 45 degree line, reflecting a high degree of correlation with the final estimates. There is no evidence for any of the parameters that the procedure converges faster for particular parameter values: indicating that accuracy improvements from DOSE are not an artefact of the particular distribution of preference parameters we observe.

C.2 Estimates of Choice Consistency

DOSE provides relatively accurate estimates of the choice consistency parameter as well as risk and loss aversion, as shown in Figure C.4. Compared to the random ordering, the DOSE

[17] Supporting this assumption, the parameter recovery exercise in Appendix D.3 finds that DOSE achieves similar levels of accuracy in a simulation where we know the true parameter values.

Figure C.1: Optimal question selection rapidly leads to accurate estimates.

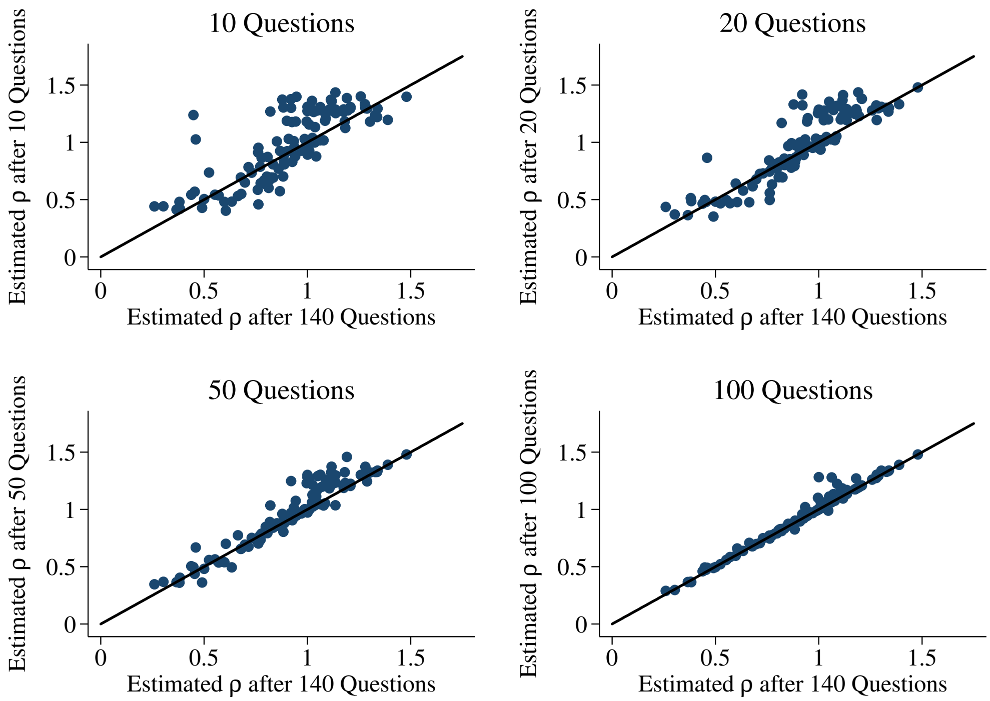


Notes: Based on data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). "Optimal Prior" and "Uniform Prior" refer to DOSE question selection using corresponding priors. "Random" orders questions randomly, averaging over 100 different random orderings. In the top panel, each line shows the inaccuracy of Bayesian estimates obtained after each question (starting at question 10). The bottom panel displays the correlation between the Bayesian estimates obtained after each question and the final estimate. All parameter values are estimated using a uniform initial prior.

estimates are closer to and more highly correlated with the final parameter estimate, and are more highly correlated with the final estimate throughout the question sequence. Again, these benefits are similar regardless of whether we use the uniform prior or the "optimal prior" (discussed in the previous subsection).

The choice consistency estimates takes longer to converge to the final estimate than

Figure C.2: Correlations between final estimates of the risk aversion parameter and the estimates after selected rounds.



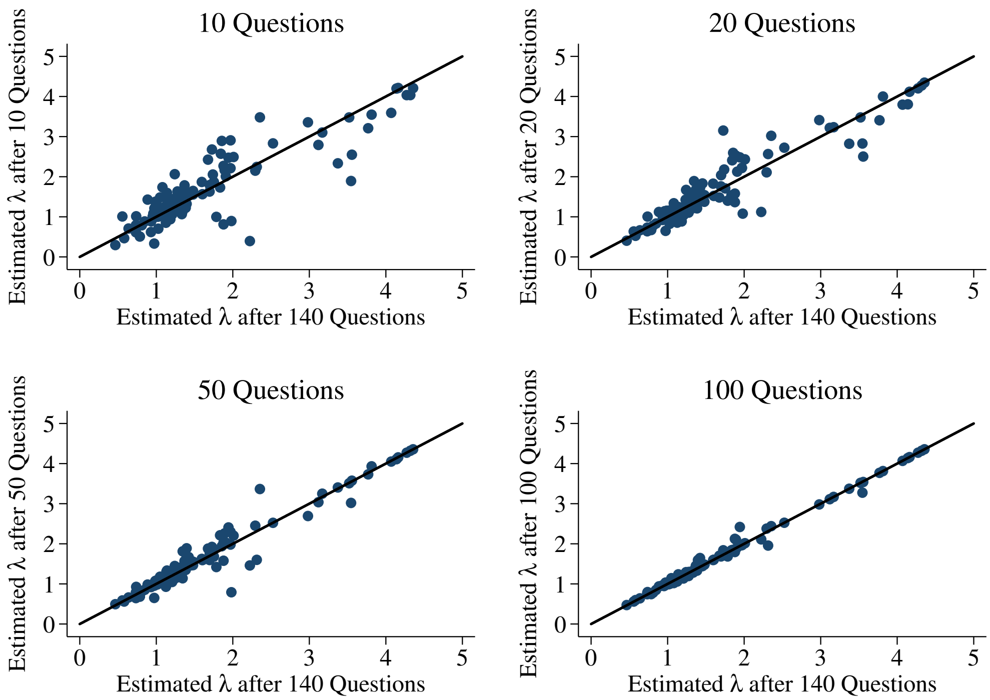
Notes: The figure is based on authors' analysis of data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). Each panel plots the DOSE estimate (using a uniform prior) of the exponent from the utility function (3) against the Bayesian estimate after 140 questions.

the other parameters, reflecting the fact that several similar questions must be asked to pin down this parameter. The estimated inaccuracy, for example, after 20 questions is 63%, much higher than the 21–22% for risk and loss aversion (but much lower than the inaccuracy (94%) achieved with the random ordering). Similarly, Figure C.5 shows that after 50 questions the individual-specific estimates are not as closely clustered around the 45 degree line. Again, however, there is no evidence that the accuracy improvements from DOSE are limited to particular values of the choice consistency parameter.

C.3 Maximum Likelihood Estimation

We also attempted to obtain individual parameter estimates using Maximum Likelihood Estimation (MLE), however we were frequently unable to estimate parameters for several

Figure C.3: Correlations between final estimates of the loss aversion parameter and the estimates after selected rounds.

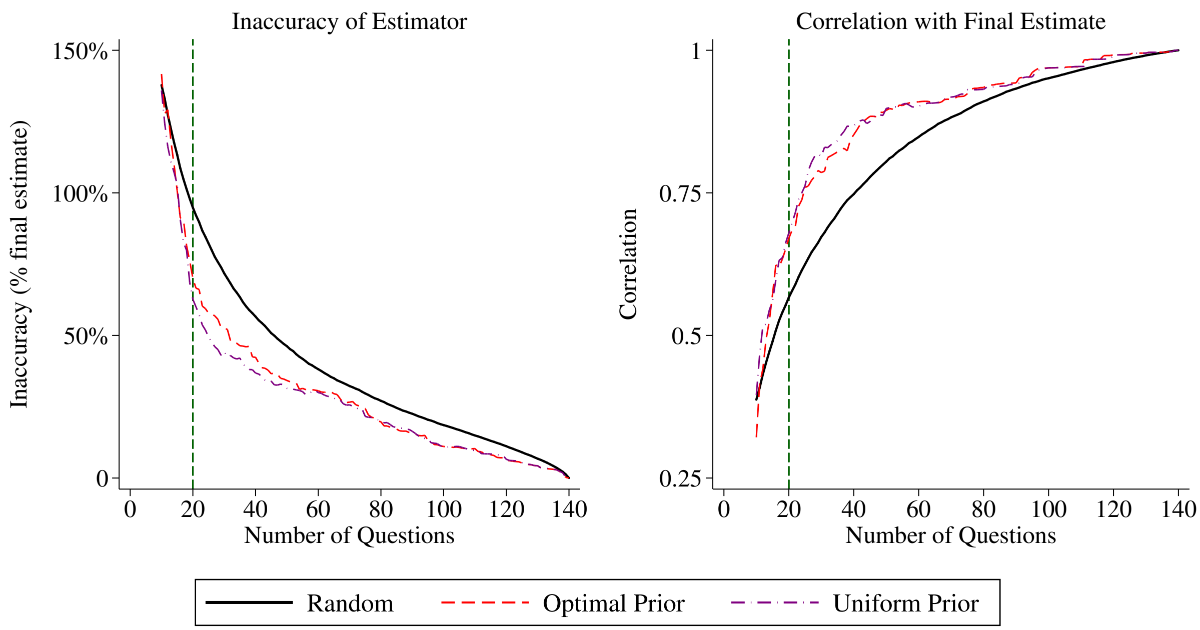


Notes: The figure is based on authors' analysis of data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). Each panel plots the DOSE estimate (using a uniform prior) of the loss aversion parameter from the utility function (3) against the Bayesian estimate after 140 questions.

participants.[18] As shown in Figure C.6, when using fewer than 40 questions (using the original order reported in the original datasets), we could not estimate parameter values for one quarter of the sample, and we could not obtain estimates for all participants even when using the full set of 140 questions. This failure is particularly striking given that, for this purpose, we do not exclude any unrealistic values (such as negative parameters) and that, in a final attempt to obtain an estimate, we initiated the search algorithm with the final Bayesian estimate of each individual's parameters. As such these numbers are an over-estimate of the proportion of participants for whom meaningful estimates could be recovered in reality; Frydman et al. (2011) in their initial study obtained estimates for only 64 of 83

[18]The MLE procedure was implemented using STATA's modified Newton-Raphson algorithm. Similar results were obtained using alternative algorithms. For each participant estimation was attempted three times (each with up to 16,000 iterations), allowing for alternative initial conditions, different stepping procedures in non-concave regions and relaxing convergence requirements on the gradient vector.

Figure C.4: DOSE elicits accurate estimates for the choice consistency parameter faster than the random ordering.



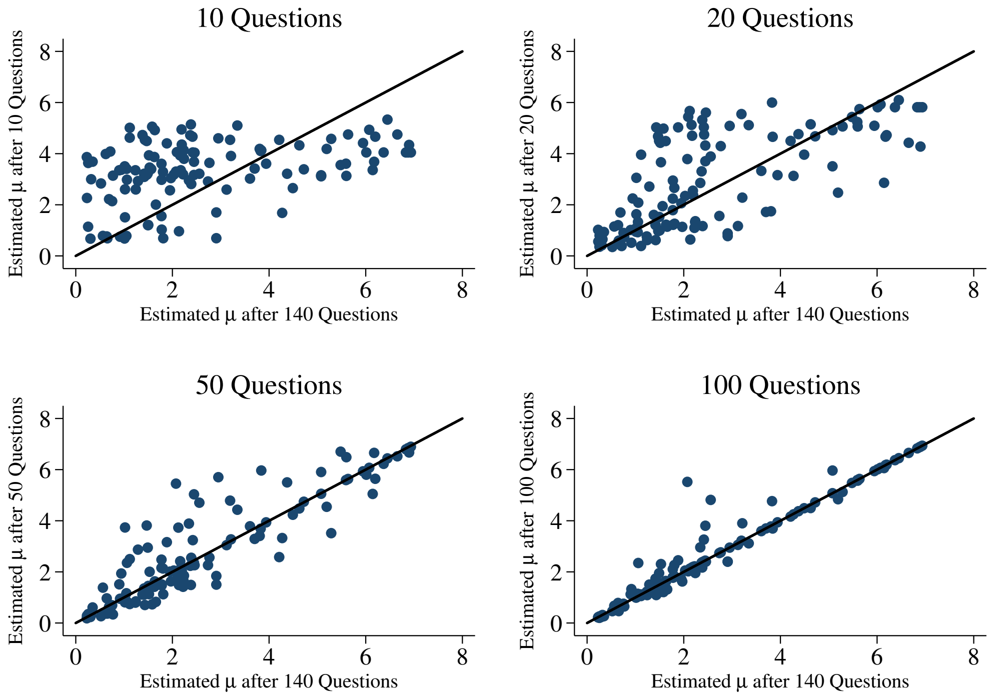
Notes: Based on data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). Left (right) hand panel shows the inaccuracy (correlation with final estimate) of Bayesian estimates obtained after each question, under different orders. "Optimal Prior" and "Uniform Prior" refer to DOSE question selection using corresponding priors. "Random" orders questions randomly, averaging over 100 different random orderings. All parameter values are estimated using a uniform initial prior.

participants (7 were excluded for other reasons), whereas we report estimates for 82 out of the 90 participants.

Further, the estimates that were obtained by MLE with a small number of questions appear much more inaccurate than those from the Bayesian procedure, as shown by the line plots in Figure C.6. After 140 rounds, the estimates from the different procedures are, as expected, very similar: the correlation between the final MLE and final Bayesian estimates was 0.85 for risk aversion, and 0.95 for loss aversion, while the median distance between the two estimates was less than 2% (of the Bayesian estimate) for both parameters. However, the Bayesian estimates are much closer to these final values after many fewer questions.[19] In addition, the Bayesian estimates are generally more accurate than the MLE

[19]To ensure comparability between the two sets of estimates, when calculating the distance from the final estimate we constrain the MLE estimates to the bounds of the prior used for the Bayesian estimates.

Figure C.5: Correlations between final estimates of the consistency parameter and the estimates after selected rounds.



Notes: The figure is based on authors' analysis of data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). Each panel plots the DOSE estimate (using a uniform prior) of the choice consistency parameter in (4) against the Bayesian estimate after 140 questions.

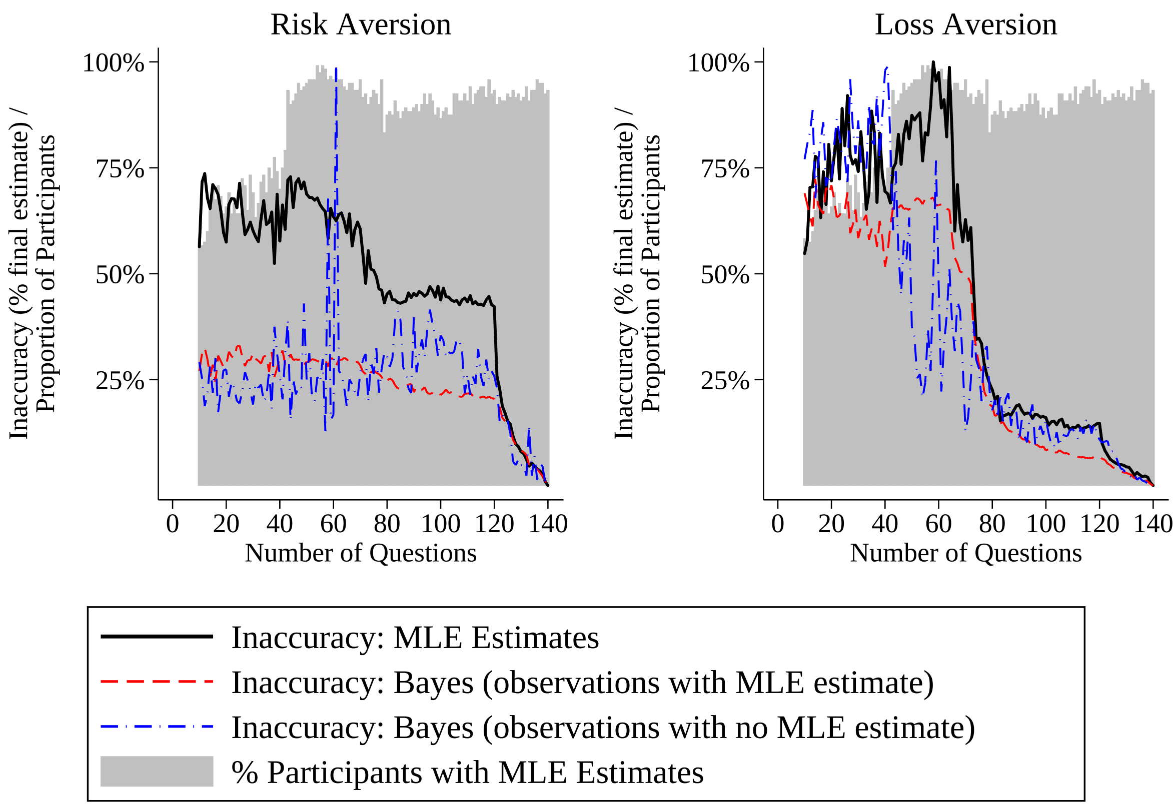
estimates that do exist even where no MLE estimate can be obtained at all.[20] Not only can the Bayesian procedure obtain an estimate in those circumstances, those estimates contain valuable information.

D Parameter Recovery Exercise

When participants make mistakes DOSE produces estimates that are about twice as accurate as traditional risk and loss aversion elicitation mechanisms. We demonstrate this with a parameter recovery exercise (or Monte Carlo simulation). This is conducted with an entirely simulated dataset that allows us to both know and control the true parameters governing

[20]The "jerky" nature of the line relating to the inaccuracy when no MLE estimate is available is explained by the fact that—particularly after question 40—few participants do not have MLE estimates, with the precise number varying from round to round. The large spike at round 61, for example, is explained by all but two participants having MLE estimates available.

Figure C.6: With a small number of questions the Bayesian procedure provides more accurate estimates than Maximum Likelihood Estimation.



Notes: The figure is based on authors' analysis of data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). The bars refer to the proportion of participants for whom a parameter estimate could be obtained using Maximum Likelihood Estimation. The lines plot the distance from the estimate obtained after 140 questions after each question round using i) Maximum Likelihood Estimation, ii) Bayesian estimation (where MLE estimates were available), and iii) Bayesian estimation (where MLE estimates were not available).

(simulated) participant behavior. We first provide a detailed explanation of the procedure used to simulate DOSE and two other elicitation methods: the double MPL, and Lottery Menu methods. To understand whether our assumptions about the level of noise in the survey are reasonable, we then compare simulated choices to real survey data. The simulation appears to underestimate the level of noise in the survey MPL.

The DOSE estimates are more accurate than other elicitation methods even when the utility function is misspecified. Accurate estimates can still be obtained by using the correct utility function after the fact and, even without re-estimating the results, the DOSE estimates are highly correlated with the true parameter values. Further, DOSE still performs well when estimating a utility function with differential curvature across gains and losses, despite the absence of questions just with losses in our dataset.

D.1 Parameter Recovery Procedure

Simulation Dataset A dataset of 10,000 simulated individuals was generated as follows. First, we estimated the 140 question DOSE procedure on the 120 participants from Sokol-Hessner et al. (2009) and Frydman et al. (2011). We then aggregated the 120 individual posterior distributions to form a joint probability distribution over the three parameters ρ, λ and μ. The 10,000 participants were then drawn from the resulting distribution.

DOSE Simulation We simulate a 20 question DOSE procedure for each individual, with each binary choice made probabilistically according to the logit probability (4). The possible question space included 760 questions, allowing for gains in \$0.25 increments up to \$10, and losses in \$0.5 increments up to \$10.

Lottery Menu In the lottery menu procedure, developed by Eckel and Grossman (2002), participants are offered a choice between multiple lotteries over gains. We calculate the expected measurement error for the menu of six 50:50 lotteries presented in Table D.1. This implementation is based on the menu used by Dave et al. (2010), adjusted so that the largest prize is \$10 (for comparability with the other elicitation procedures). The first lottery is a safe option (it has zero variance), while the subsequent lotteries increase in both expected value and variance.

The choice of lottery implies a range of possible CRRA coefficients, as shown in the penultimate column of Table D.1. For lotteries 2-5 we estimate the estimated CRRA coefficient $\hat{\rho}$ as the midpoint of this range. Since the midpoint is undefined for lotteries 1 and 6, for these lotteries we use the end-point of the range. To ensure comparability with the DOSE estimates, we then truncate the estimated parameters to the range defined by the Sokol-Hessner-Frydman distribution.

Table D.1: Choices in Simulation of Lottery Menu Procedure

	Low Prize	High Prize	CRRA Range	Estimated ρ
Lottery 1	4.00	4.00	$\rho < -2.46$	0.20
Lottery 2	3.43	5.14	$-2.46 < \rho < -0.16$	0.20
Lottery 3	2.86	6.29	$-0.16 < \rho < 0.29$	0.20
Lottery 4	2.29	7.43	$0.29 < \rho < 0.50$	0.40
Lottery 5	1.71	8.57	$0.50 < \rho < 1.00$	0.75
Lottery 6	0.29	10.00	$1 < \rho$	1.00

Notes: Lottery menu choices taken from Dave et al. (2010), adjusted so that maximum prize is $10. "CRRA range" is the implied range of CRRA coeffcients implied by the choice of each lottery. "Estimated ρ" is the estimated value of the CRRA coefficient associated with the choice of each lottery used in the calculation of expected inaccuracy.

The procedure for the simulation was as follows. Consider a menu over a set of lotteries $l_1, l_2, ..., l_{\mathcal{L}}$. We define a probability distribution over the set of lotteries by assuming that individuals make a series of binary choices in which they compare the set of lotteries in order. That is, they first compare lottery 1 with lottery 2, making a choice according to the logit probability. They then compare the winner of that choice with lottery 3, and then the winner of the latter choice with lottery 4. The procedure is repeated until lottery $\mathcal{L}$.

We define a probability distribution over the full lottery menu for each participant i as follows. For two lotteries l, k let $q^i_{l,k}$ be the probability that i chooses l when faced with a binary choice between l and k. This probability is defined by the logit function (4), and as such depends on the participant's value of ρ and μ; for simplicity we do not display these parameters or the i index in the following. Define the probability that lottery l is chosen after L choices as p^L_l. Then $p^1_1 = q_{1,2}$ and for all other l, L:

$$p^L_l = \sum_{k=1}^{l-1} q_{l,k} p^{l-1}_k \times \prod_{m=l+1}^{L} q_{l,m}$$

The probability distribution over the choice from the set of lotteries is then $\{p^{\mathcal{L}}_1, p^{\mathcal{L}}_2, ..., p^{\mathcal{L}}_{\mathcal{L}}\}$. Defining $\hat{\rho}_l$ as the estimated CRRA coefficient associated with a choice of lottery l, the

Table D.2: Hypothetical MPL 1 used to estimate ρ

Left Hand Choice	Right Hand Choice	CRRA Range	Estimated ρ
50% of \$0, 50% of \$10	\$0	n.a.	n.a.
50% of \$0, 50% of \$10	\$1	$\rho < 0.30$	0.23
⋮	⋮	$0.30 < \rho < 0.43$	0.37
⋮	⋮	$0.43 < \rho < 0.58$	0.50
⋮	⋮	$0.58 < \rho < 0.76$	0.66
⋮	⋮	$0.76 < \rho < 1.00$	1.57
⋮	⋮	$1.00 < \rho < 1.36$	1.16
⋮	⋮	$1.36 < \rho < 1.94$	1.61
⋮	⋮	$1.94 < \rho < 3.11$	1.66
50% of \$0, 50% of \$10	\$9	$3.11 < \rho < 6.58$	1.66
50% of \$0, 50% of \$10	\$10	$6.58 < \rho$	1.66

Notes: "CRRA range" is the implied range of CRRA coefficients implied by the choice of each lottery. "Estimated ρ" is the estimated value of the CRRA coefficient associated with the choice of each lottery used in the calculation of expected inaccuracy. Neither value is defined in the first row because the design does not allow the right hand side to be selected.

expected inaccuracy is given by:

$$E[|\hat{\rho} - \rho|] = \sum_{l=1}^{\mathcal{L}} p_l^{\mathcal{L}} \times |\hat{\rho}_l - \rho|$$

We also implemented an alternative simulation procedure for the Lottery Menu. Under this alternative, choice occurred according to to a multinomial logit probability distribution. That is, for each possible choice $k = 1, ..., 6$:

$$Prob(Choice = k) = \frac{exp(EU_k)^{\mu}}{\sum_{l=1}^{6} exp(EU_l)^{\mu}}$$

where EU_k is the expected utility of lottery k.

Double Multiple Price List (MPL)

Table D.3: Hypothetical MPL 2 used to estimate λ

Left hand choice	Right hand choice
50% of -$10, 50% of $10	-$10
50% of -$10, 50% of $10	-$9
⋮	⋮
⋮	⋮
50% of -$10, 50% of $10	$9
50% of -$10, 50% of $10	$10

We calculate the expected inaccuracy for the double MPL method using two hypothetical MPLs. MPL 1 offers participants a choice between a fixed 50:50 lottery between $0 and $10 and a series of fixed amounts. This MPL is used to elicit the estimate of the CRRA coefficient ρ. MPL 2 offers participants a choice between a 50:50 lottery between a loss of $10 and a gain of $10 and a series of fixed amounts. This second MPL is used to obtain the estimate of the loss aversion parameter λ. In both MPLs we enforce (in-line with the implementation in the surveys) that individuals could only switch once, and that individuals do not choose dominated options: the left hand side of MPL (the lottery) is chosen in the first row and the right hand side (the fixed amount) is chosen in the last row.

The row in which a participant first chooses the fixed amount (the right hand side) in MPL 1 implies a range of certainty equivalents and CRRA coefficients, as shown in Table D.2. We use the certainty equivalent at the midpoint of this range and the associated CRRA coefficient.

Similarly, the row in which a participant first chooses the fixed amount (the right hand side) of MPL 2 implies a range of certainty equivalents, as shown in Table D.3. We use the certainty equivalent at the midpoint of this range and use the estimated CRRA coefficient $\hat{\rho}$ estimated in MPL 1 to obtain the estimated loss aversion parameter, $\hat{\lambda}$. For comparability with the DOSE estimates, we truncate the range of $\hat{\lambda}$ and $\hat{\rho}$ to match the range of the prior used in the DOSE procedure.

The procedure for simulating behavior on these two MPLs was as follows. For each row r, the probability that a simulated individual defined by the parameter vector (ρ, λ, μ) first chooses the right hand side of the MPL in row r is calculated. This probability is defined by the logit probability (see (4)) comparing the lottery to the fixed amount offered in row r. To translate these binary choices into a probability distribution over the set of rows in the MPL we assume that individuals work either sequentially down or up an MPL, each with 50% probability. Suppose they work down the MPL. Then they first consider the choice between the lottery and the fixed amount in the first row in which they can choose the fixed amount (row 2 in our implementation). If they choose the fixed amount, they will always prefer the fixed amount lower in the MPL: thus this row is the "switching row". If, on the other hand, they prefer the lottery then they will move to the next row and consider the next binary choice. Alternatively, individuals may choose to work up the MPL by first considering the bottom row of the MPL, then the second-bottom, etc.

Now consider a MPL with $\mathcal{R}$ rows in which an individual can switch. Define the probability that the lottery is chosen in row r by individual i as q_r^i. This probability is defined by ρ, μ and, when losses are involved, λ. For simplicity we suppress the i indices. Define the probability row r is the switching row working down the MPL as p_r^D, and working up the MPL as p_r^U. Then these probabilities are given by:

$$\begin{aligned} p_r^D &= (1-q_r)\prod_{s=1}^{r-1} q_s \quad \text{and} \\ p_r^U &= (q_{r-1})\prod_{s=r}^{\mathcal{R}} (1-q_s) \end{aligned}$$

The expected inaccuracy for any parameter θ is then given by:

$$E[|\hat{\theta}_r - \theta|] = \sum_{r=1}^{\mathcal{R}} \left(0.5p_r^D + 0.5p_r^U\right) |\hat{\theta}_r - \theta|$$

where $\hat{\theta}_r$ is the estimated parameter associated with switching in row r. As discussed above, for ρ this is implied by the midpoint of the certainty equivalents defined by the switching row. For λ the value is defined both by the midpoint of the certainty equivalent and the estimated $\hat{\rho}$ from MPL 1.

D.2 Validating the MPL Simulation with Survey Data

In practice, choice data in the MPL is likely noisier than our simulations assume. To estimate the relative amount of noise in the survey, we compare simulated and real responses for three additional MPLs—a double MPL procedure (but with different payoffs), and an additional risk aversion MPL as implemented in Chapman et al. (2018).[21] Two of these MPLs elicited risk aversion, offering participants a choice between fixed amounts and 50:50 lotteries over gains: a lottery over \$0 and \$5 and a lottery between \$1 and \$4 respectively. The third MPL, which elicited loss aversion, offered participants a choice between fixed amounts and a 50:50 lottery between a loss of \$5 and a gain of \$5.

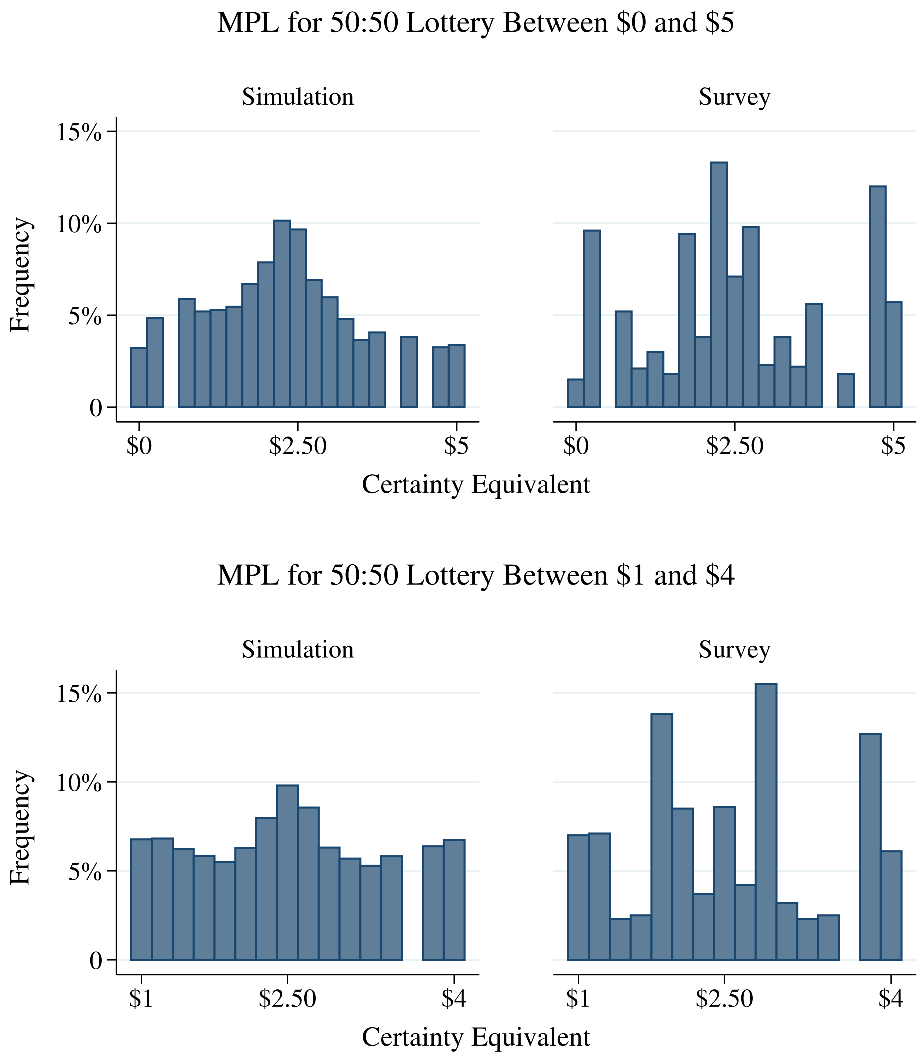
The comparison of the real and simulated responses suggests that the level of measurement error in our simulation was lower than that in the survey. The proportion of FOSD responses in the loss aversion MPL is much lower in this simulation than the real data: 20% rather than 37%. Further, the correlation between the certainty equivalents in the risk aversion MPLs—which is higher in the presence of less measurement error (Gillen et al., Forthcoming)—is higher in the simulation: 0.73 vs. 0.69. In addition, a significant degree of the correlation in the survey data is explained by participants repeatedly switching at the rows at the end of the MPLs: choices which are consistent, but unlikely to be accurate given the extreme parameter values they imply. Excluding such participants the correlation between the two MPLs falls to 0.50 in the survey, compared to 0.61 in the simulated data.

The reason that the simulations underestimate the measurement error in the survey MPLs appears to be that we do not account for participants' use of rules-of-thumb. As shown in Figure D.1, compared to our simulation participants in the survey were more likely to switch

[21] These MPLs have a different structure from those used in the results in Table D.4 and Figure D.2, but the simulation methodology is the same.

in rows of the MPL that are especially salient—such as the first or last rows, or those referring to the midpoint of the lottery.[22] These choices may be capturing framing effects or heuristics in the face of the large amount of information in an MPL—the simulations may thus miss an important source of error in the MPL by not accounting for such effects.

Figure D.1: MPL endpoints are chosen more frequently in real data.



Notes: The figure displays the real and simulated responses to the two risk aversion MPLs in Chapman et al. (2018).

[22] Amounts with zero choices in these histograms reflect the fact that, unlike the hypothetical MPLs in the previous section, the fixed amounts in these MPLs were not at regular intervals meaning that some values could not be chosen by the participants.

Table D.4: DOSE produces more accurate estimates.

	Average Inaccuracy	Spearman Rank Correlation with True Value
Loss Aversion		
DOSE 10 question	21%	0.86
DOSE 20 question	15%	0.91
Multiple Price List	36%	0.65
Risk Aversion		
DOSE 10 question	21%	0.66
DOSE 20 question	15%	0.79
Multiple Price List	37%	0.45
Lottery Menu	35%	0.28

Notes: Inaccuracy is the absolute distance from the true parameter value as a percentage of the true value.

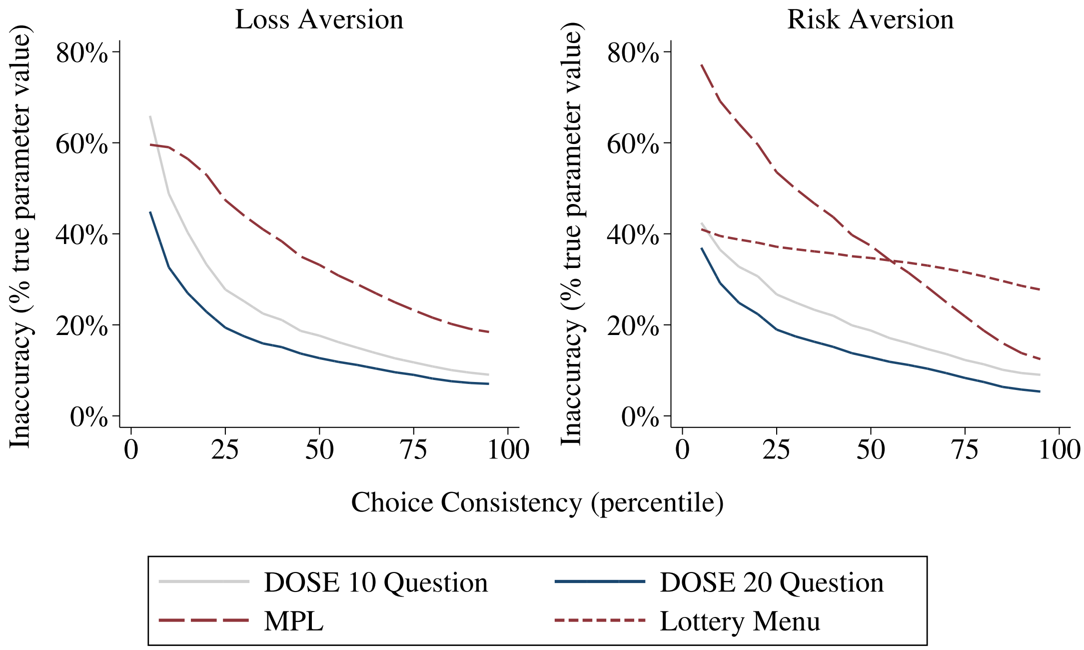
D.3 Comparison of DOSE with Other Elicitation Procedures

The estimates of risk and loss aversion from DOSE are approximately twice as accurate as those from the two other elicitation procedures, as shown in Table D.4. After 20 questions, DOSE obtains estimates of risk and loss aversion that are, on average, within 15% of the true parameter value. The average inaccuracy of the MPL and Lottery Menu procedures, in contrast, is at least 35%—much higher than even a 10-question DOSE procedure.

DOSE produces more accurate estimates than both other procedures regardless of participants' level of choice consistency (μ), as shown in Figure D.2. This figure repeats the parameter recovery analysis above, but assigns all simulated participants the same level of choice consistency, μ. We then vary μ across percentiles of the population distribution. The 20-question DOSE procedure always provides the most accurate estimates. Even the 10-question procedure performs better than either the MPL or Lottery Menu, except amongst extremely inconsistent participants.

Although the Lottery Menu procedure appears to perform better than the MPL for inconsistent participants, this advantage is not robust to the alternative simulation procedure

Figure D.2: DOSE is more accurate than other methods at all levels of choice consistency.



Notes: Estimates obtained using simulation procedure described in Appendix D.1, with all simulated participants at a point in the graph having the same value of μ.

in which choice in the Lottery Menu is made according to a multinomial logit distribution. Under this alternative assumption, the estimated inaccuracy is much higher both on average (as in Table D.4)—at 94%—and for the lowest consistency ventile—at 139% (for the highest ventile, it is 59%).

The high accuracy of the DOSE estimates also leads to higher correlations with the true parameter values than the other two procedures (column 2 of Table D.4). Thus, DOSE is less likely to miss associations between economic preferences and other characteristics through attenuation bias. The correlation between the true risk aversion parameter and the DOSE estimate is 0.79, compared to 0.45 with the MPL estimates and 0.28 for the Lottery Menu. For loss aversion, the DOSE procedure produces correlations above 0.85 with the true values, even after a 10-question procedure. This is reflected in our survey results, see Section 3.3.

Unlike the MPL, DOSE is able to elicit loss aversion estimates even when participants' choices violate First Order Stochastic Dominance (FOSD), although this is not an important factor in the simulation results of Table D.4. Because DOSE accounts for the possibility that a participant's choice is a mistake, the procedure can always recover parameter estimates. In the double MPL, on the other hand, participants may erroneously make choices on the second MPL (used to elicit loss aversion) that are First Order Stochastically Dominated given their choices on the first MPL (used to elicit risk aversion). This prevents estimation of the loss aversion parameter. In our simulation, the MPL could not recover estimates for 11% of participants—increasing to more than 50% of participants with low choice consistency.

In practice, the double MPL procedure is unable to elicit loss aversion for a significant proportion of the population, which may lead to biased conclusions about loss aversion. In particular, the double MPL used by Chapman et al. (2018) is unable to recover estimates for 37% of their participants, as measurement error and other factors lead to these participants appearing to make FOSD choices. The prevalence of these choice patterns was not random: loss aversion could only be computed for 50% of participants in the bottom quartile of cognitive ability, compared to 70% in the upper quartile. The results in this paper (see Section 3.2) indicate that high cognitive ability is associated with more loss aversion. This pattern of missing observations may thus lead to a biased over-estimate of loss aversion.[23]

D.4 Misspecification of the Utility Function

The DOSE estimates are robust to misspecification of the utility function. We run DOSE on the same 10,000 simulation subjects—each of whom has CRRA utility—but assuming a CARA utility function in the question selection procedure. We then compare the correlation between the risk aversion and loss aversion parameters under the different procedures, and demonstrate how the data collected can be re-estimated to elicit accurate CRRA utility

[23] The survey in Chapman et al. (2018) is similar to the one in this paper, although it did not utilize DOSE. Note this pattern of responses is consistent with the results in Section 3.2, as loss-tolerant individuals will wish to make choices that are close to violating FOSD, and measurement error can push them over the threshold. Low cognitive ability participants are more loss tolerant and make more FOSD choices.

parameters.[24]

Misspecifying the utility function does not lead to a loss of accuracy, as shown in Table D.5. For loss aversion, very similar estimates are obtained even when the CARA function is incorrectly used (see the top panel of the table). For risk aversion, we can recover the same estimates by re-estimating the correct utility function after the data collection process.

Further, the assumptions over parametric form are unlikely to be critical if researchers are interested in identifying correlations rather than the level of the risk and loss aversion estimates. Even without re-estimating, the Spearman correlation between the estimated CARA parameters and the true (CRRA) parameter values is very high—and notably higher than the correlations for either the MPL (0.45) or the Lottery Menu (0.28) procedures reported in Table D.4.

D.5 Allowing for Differential Risk Aversion over Gains and Losses

DOSE can obtain reasonably accurate estimates for a utility function with differential utility curvature between gains and losses, even if—as in our survey—no questions solely involving losses are asked. We simulate the DOSE procedure using the procedure outlined in Section D.1, but assuming that participants have a utility function with different power exponents in the gain and loss domain, as suggested by Prospect Theory (Kahneman and Tversky, 1979).[25] We then simulate a 20 question DOSE procedure using the same four

[24]Specifically we run DOSE on the simulation dataset assuming the following exponential (CARA) utility function:

$$u(x, \gamma_i, \lambda_i) = \begin{cases} \frac{1-e^{-\gamma_i x}}{\gamma_i} & \text{for } x \geq 0 \\ \lambda_i \left(\frac{e^{\gamma_i x}-1}{\gamma_i} \right) & \text{for } x < 0 \end{cases} \tag{6}$$

where λ represents loss aversion and γ captures risk aversion.

[25] Specifically, we estimate the following unrestricted power utility function:

$$u(x, \rho_i^+, \rho_i^-, \lambda_i) = \begin{cases} u(x) = x^{\rho_i^+} & \text{for } x \geq 0 \\ u(x) = -\lambda_i(-x)^{\rho_i^-} & \text{for } x < 0 \end{cases} \tag{7}$$

As with the previous simulations, we start by constructing a simulated dataset by estimating the procedure on the data from the 120 participants in Sokol-Hessner et al. (2009) and Frydman et al. (2011) with a discretized uniform prior (we use the same prior range for both risk aversion parameters as we use for ρ in the main estimates). The joint posterior from that procedure is then used to draw simulated participants.

Table D.5: DOSE estimates are robust to utility function misspecification.

	Average inaccuracy		Correlation with true value	
	10 question	20 question	10 question	20 question
Loss Aversion				
CRRA (Not misspecified)	21%	14%	0.84	0.89
CARA (Misspecified)	23%	16%	0.84	0.90
CARA re-estimated as CRRA	21%	15%	0.85	0.91
Risk Aversion				
CRRA (Not misspecified)	21%	16%	0.66	0.79
CARA (Misspecified)	n.a.	n.a.	0.58	0.75
CARA re-estimated as CRRA	21%	15%	0.67	0.77

Notes: Inaccuracy is defined as the absolute distance from the true parameter value displayed as a percentage of the true value. “Correlation with true value” displays the Spearman correlation coefficient between the true parameter and the estimated parameters.

parameter power utility function in the question selection procedure.

DOSE extracts meaningful information about all four parameters, although with less precision than in the three parameter model, as shown in Table D.6. For risk aversion over gains the accuracy of the estimates are similar to those for the three parameter model. For the two parameters regarding choices over losses, however, the estimates are noisier. The inaccuracy and correlation with true estimates for the curvature over losses are comparable to those for the Multiple Price List in the three parameter model. The loss aversion parameter has higher correlations. Further, although the average inaccuracy is very high, this is largely an artefact of the fact that there a number of very small values of λ in the simulation. Excluding the smallest 10% of values of λ (corresponding to values of less than 0.5), the estimated inaccuracy is 36%.

Table D.6: DOSE estimates of the 4 parameter model are less accurate than the 3 parameter model.

	CRRA		CRRA with ρ^+ and ρ^-	
	Average inaccuracy	Correlation w/true value	Average inaccuracy	Correlation w/true value
Loss Aversion				
DOSE 10 question	21%	0.84	85%	0.61
DOSE 20 question	14%	0.89	63%	0.73
Risk Aversion over Losses				
DOSE 10 question	n.a.	n.a.	38%	0.31
DOSE 20 question	n.a.	n.a.	31%	0.51
Risk Aversion over Gains				
DOSE 10 question	21%	0.66	21%	0.60
DOSE 20 question	16%	0.79	16%	0.73

Notes: Inaccuracy is defined as the absolute distance from the true parameter value displayed as a percentage of the true value. "Correlation with true value" displays the Spearman correlation coefficient between the true parameter and the estimated parameters.

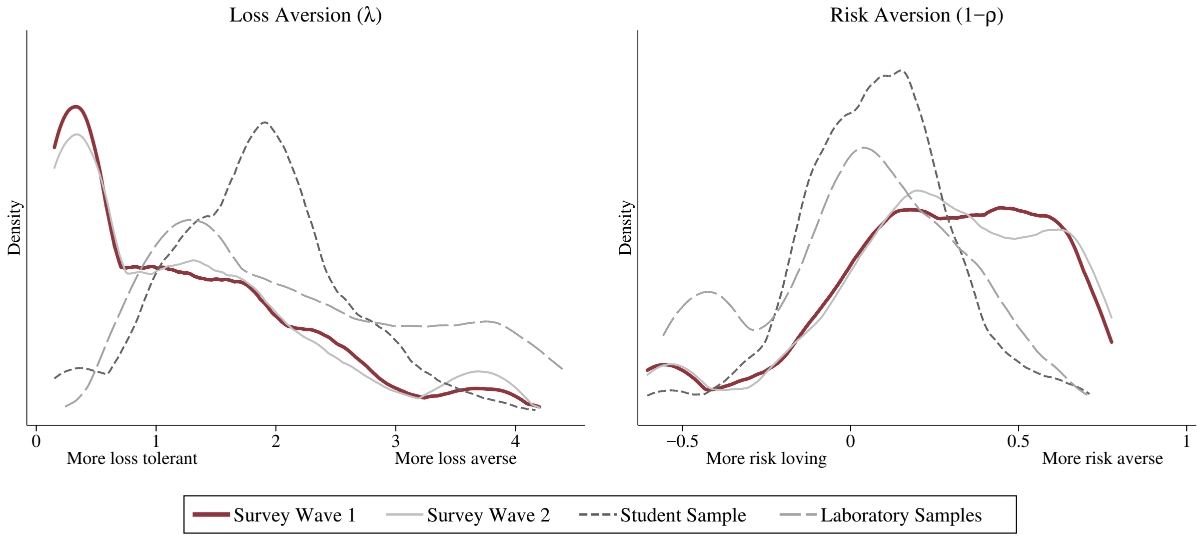
E Additional Laboratory Data

The level of loss aversion amongst participants using DOSE in the laboratory is similar to our student survey, and close to the "standard" estimate of 2 (Fehr-Duda and Epper, 2012). We take advantage of four previous laboratory studies that used DOSE (based on our original working paper) to estimate risk and loss aversion in much the same way we do. Each of these studies used DOSE to ask participants a personalized sequence of binary choices with the same structure as the questions in our survey. We take those individual choices, and then re-estimate the individual risk and loss aversion parameters using the same procedure and priors as used for our main estimates. The level of loss aversion amongst these laboratory subjects is similar to that our student survey—but very different to the general population. When implemented in a standard experimental setting, DOSE estimates are in-line with previous studies—our findings thus appear to be a result of our general population sample, rather than our elicitation procedure and survey methodology.

We estimate risk and loss aversion for a total of 439 students across the four laboratory experiments. Each of these participants answered 30 or 40 binary choices selected by DOSE from a set of approximately 140 potential questions based on those using in Sokol-Hessner et al. (2009) and Frydman et al. (2011). 58 participants completed DOSE at Caltech (Krajbich et al., 2017), 172 at Claremont McKenna (108 collected by Clay et al. (2017) and 64 by Clay et al. (2016)), and 209 students at UCLA (this data was generously provided to us by Alec Smith). Participants at Caltech and UCLA were paid using the two incentive compatible methods described in Appendix B.2. Under the first method, implemented in the Caltech study, the parameters estimated from the DOSE module were used to implement payment on a randomly selected question from the full set of possible questions. In the second method, used at UCLA, one question from the full set of possible questions was randomly selected at the end of the DOSE sequence. If the question had been answered previously, that answer was used to determine pay. If the question had not been chosen by DOSE, it was posed to the subjects and then payment was made. Participants at Claremont were not paid for the DOSE module.

Comparing the distribution of preference parameters, as in Figure E.1, across our three samples suggests that our findings are not an artifact either of the DOSE methodology or our implementation in an online survey. Amongst students, the extent of loss tolerance is similar whether DOSE was implemented in the laboratory or online (10% of participants in each sample). Further, the median loss aversion parameter in both cases—$\lambda = 1.99$ and $\lambda = 1.84$ respectively—is similar to the average level of loss aversion reported in laboratory studies using alternative elicitation methods (see summary in Booij et al., 2010, Table 1). The general population, in contrast, has a much higher level of loss tolerant participants

Figure E.1: The general population is more loss averse and risk tolerant than students.



Notes: The figure displays the kernel density of each parameter, plotted using Epanechnikov kernel with bandwidth chosen by rule-of-thumb estimator. "Student Sample" refers to students completing DOSE in an online survey, as discussed in Section 3.1. "Laboratory Samples" refers to students completing DOSE in previous laboratory experiments—see text above for details.

(52%) and lower median loss aversion ($\lambda = 0.99$). The general population is also more risk averse than the lab-based populations, in-line with prior research (see Snowberg and Yariv, 2018, and references therein): the median CRRA coefficient (1-ρ) in the general population is 0.29 vs. 0.05–0.09 in the two student samples.

F Robustness Checks

This section presents extended survey results and robustness tests. In the first subsection, we demonstrate that our main results are robust to misspecification, with a detailed discussion of the robustness tests discussed in Section 4.1. We then show that the correlations between economic preferences and cognitive ability presented in Section 3.2 are robust to controlling for other economic characteristics (F.2). The relationship between cognitive ability and expected-value making choices is also robust to controlling for individual characteristics: high cognitive ability individuals make fewer such choices over lotteries including losses.

In Appendix F.4 we provide additional evidence that neither survey fatigue or inattention appear to explain our results, extending the analysis in Section 4.4. The following subsection (F.5) shows that DOSE choice consistency measure can identify inattention (Section 3.3) even when restricting the sample to those answering quickly—a standard way of identifying inattention in the survey. The final subsection then shows that the results are very similar across the two waves of our survey.

Section 3.3

F.1 Model (Mis-)Specification

This subsection contains additional information related to the tests of model misspecification in Section 4.1. First we show that the categorization by DOSE in our main estimates clearly captures participants' choices, indicating that the finding of loss tolerance is not a consequence of functional form. Consequently re-estimating the data with CARA utility or allowing differential risk aversion over gains and losses does not significantly change our results: the proportion loss tolerant remains similar regardless of the specification. Finally, we show that our conclusions are robust to implementing either a probit error specification or the random parameter model suggested by Apesteguia and Ballester (2018).

In Table F.1 we classify participants according to their estimated parameter values—for instance, a participant is "loss averse, risk averse" if they have both $\lambda > 1$ and $\rho < 1$—and we examine how the frequency of lotteries accepted varies according to the expected value (relative to a sure amount) and whether the lottery involved a loss. The pattern of behavior is as would be expected. Loss tolerant participants nearly always choose lotteries with losses, and risk loving participants nearly always choose lotteries over gains. Loss averse and risk averse participants, in contrast, are much less likely to accept such lotteries.

Given this behavior, it is not surprising that the extent of loss tolerance does not vary significantly when we re-estimate the data using alternative utility functions. Using our preferred CRRA specification ((3) in Section 2.2), 52% are loss tolerant with a median loss aversion of 0.99. Using the CARA model (see (6) in Appendix D.4) 48% are loss tolerant,

Table F.1: DOSE classification reflects clear pattern of choices.

	% Lotteries Accepted			
	Lotteries with Losses		Lotteries with Only Gains	
	EV≤Sure	EV>Sure	EV≤Sure	EV>Sure
Classification by DOSE				
Loss averse, risk averse	14%	46%	8%	41%
Loss averse, risk loving	2%	51%	53%	98%
Loss tolerant, risk averse	82%	99%	5%	41%
Loss Tolerant, Risk Loving	82%	100%	78%	96%

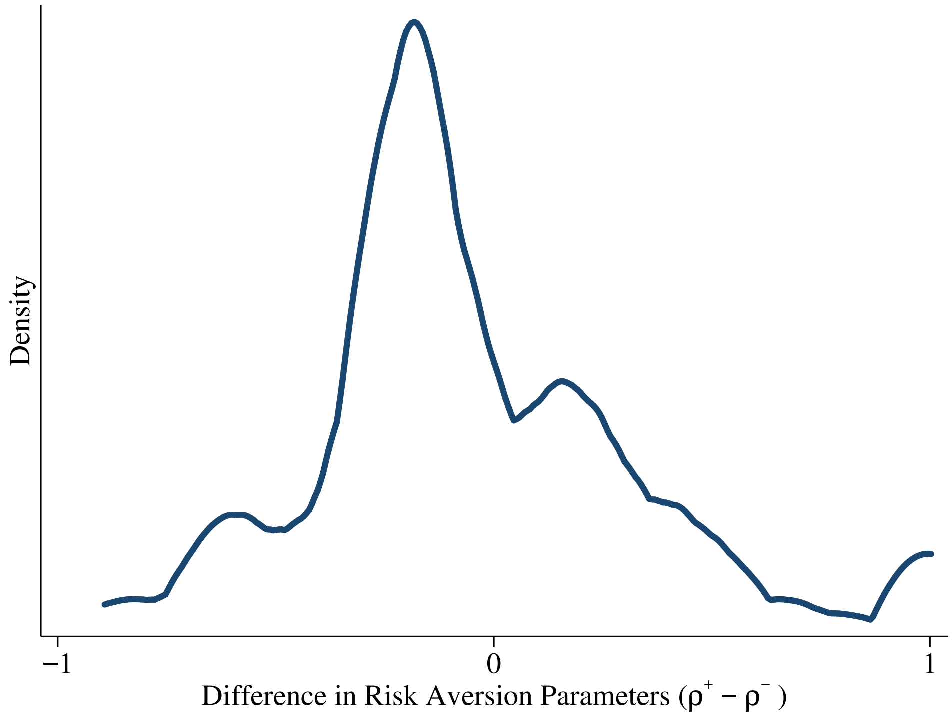
Notes: The table displays the unweighted percentage of lotteries accepted, categorizing participants according to their estimated DOSE parameters. "EV"=Expected Value of lottery and "Sure"= the sure amount offered in each lottery.

and the median loss aversion parameter is 1.02. Allowing for differential curvature (see (7) in Appendix D.4), 44% are loss tolerant, and the median loss aversion parameter is 1.16. The overall distributions of parameters are also very similar across the three models, as we have seen in Figure 6.

The re-estimated survey data also provides some support for our assumption of a CRRA utility with the same curvature over gains and losses in (3). As shown in Figure F.1, on average the curvature of the risk aversion parameter is similar across the two domains. In particular the mean difference in the two parameters ($\rho^+ - \rho^- < 1$) is -0.04, and the median difference is -0.12. These results are consistent with previous findings that utility over losses is closer to linearity (Booij et al., 2010). However, it is clear from the figure that there is considerable individual heterogeneity that is not captured by these average estimates.

Similar exercises demonstrate that our results are robust to allowing for alternative error processes, as shown in Figures F.2 and F.3. Figure F.2 shows our results for risk and loss aversion using a probit (rather than logit) error function. That is, we assume that the probability that the lottery is chosen follows a standard normal distribution over the utility difference between the lottery and sure amount. The estimated parameters are extremely

Figure F.1: On average the risk aversion parameter is similar in the loss and gain domains.

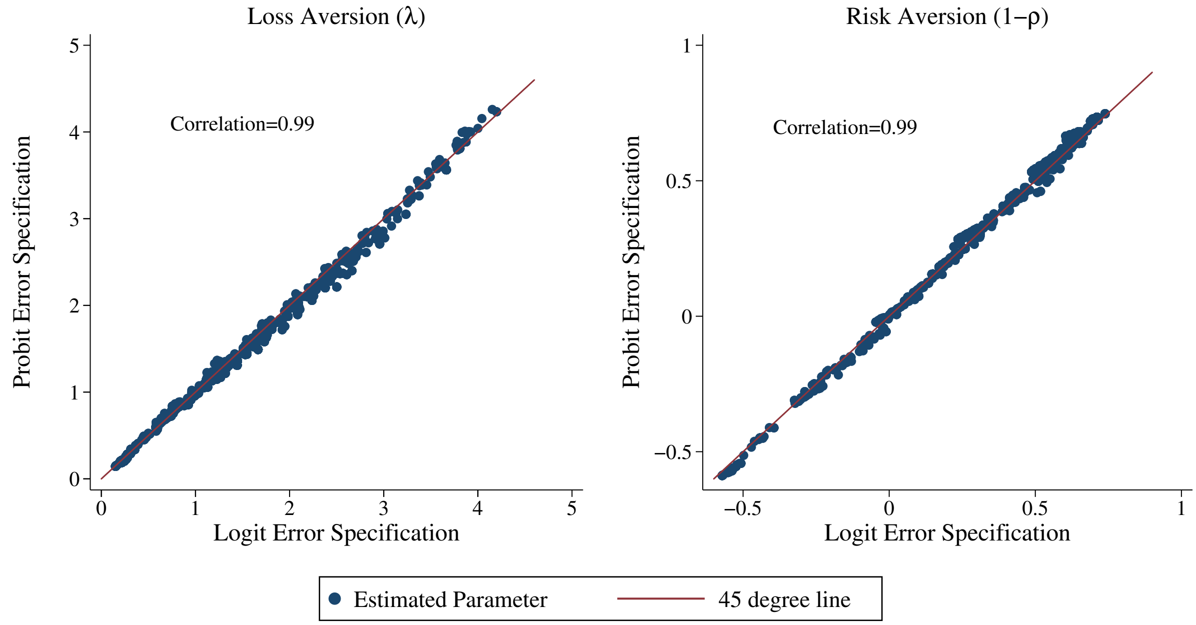


Notes: The figure displays the density of the difference in the risk aversion parameters over gains and losses $(\rho^{+} - \rho^{-})$ from (7).

similar, with the median loss aversion parameter (0.99) and the percentage loss tolerant (52%) the same as with the logit error specification.

Our results are also similar when implementing a random-parameter model (RPM), as suggested by Apesteguia and Ballester (2018) to avoid potential biases associated with the logit (and similar) error functions. In particular, they show that with a logit error specification the relationship between the level of risk aversion and the probability of accepting a lottery is non-monotonic—potentially creating issues for identification of the risk aversion parameter. Those non-monotonicities do not affect loss aversion directly because the difference between the utility of a lottery with a loss and a sure amount is monotonically decreasing in loss aversion, meaning that higher loss tolerance is associated with an increasing probability of accepting a lottery (for a given level of risk aversion). However, they could impact our loss aversion estimates indirectly through the risk aversion parameter. As such, we re-estimate our results using a RPM model, finding that our main conclusions are unchanged.

Figure F.2: Using the probit error specification leads to very similar parameter estimates.

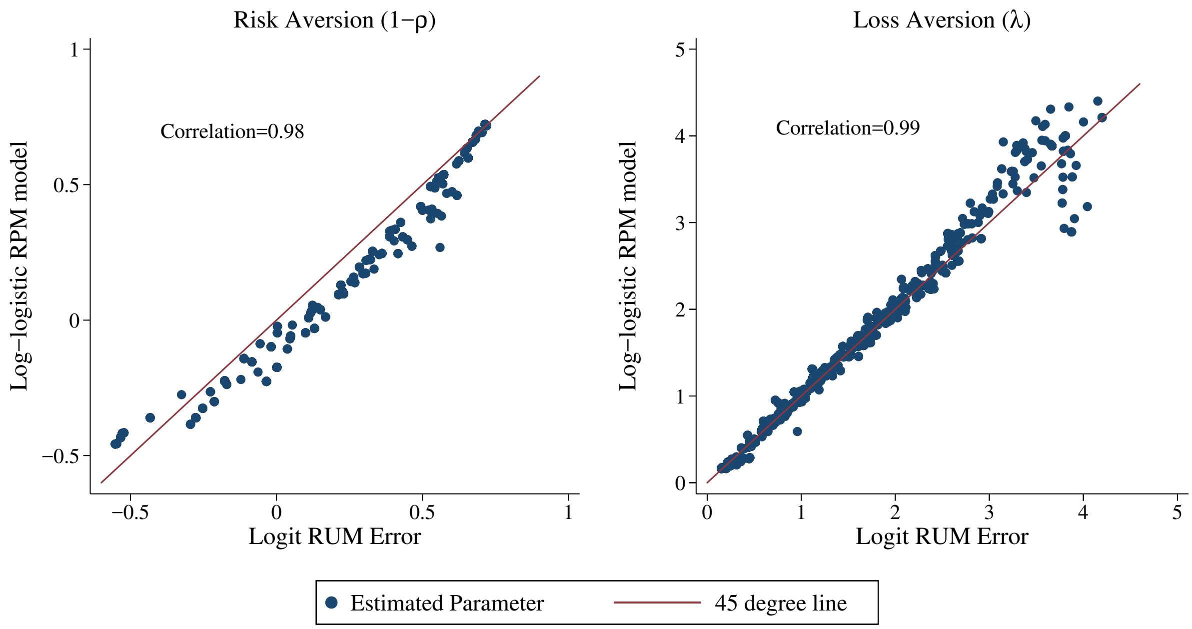


Notes: The figure displays the distribution of the loss and risk aversion estimates under the logit error specification used in the main estimates (x-axis) and a probit error specification (y-axis).

To implement the RPM, we first estimated risk aversion and then, in a second stage, estimated loss aversion. We choose this two stage approach as, first, implementing a RPM in a two parameter model is not straightforward (Apesteguia and Ballester (2018) examine only a one parameter model) and, second, it is only risk aversion that is directly affected by the non-mononocities (as discussed in the previous paragraphs). The risk aversion parameter is largely identified off of questions where the lottery contains only gains. As such we used a RPM on those questions to identify the risk aversion parameter. We then fix individual-level risk aversion parameters using the RPM estimates, and implement the logit specification to estimate loss aversion using the remaining questions—those that ask participants to choose between a lottery with both a loss and a gain and a sure amount of zero.

As shown in Figure F.3, our conclusions about loss aversion are unchanged when using the RPM model, although our risk aversion estimates are slightly different. The left hand panel of the figure displays compares the risk aversion parameters estimated using RPM with those estimated using the logit model: the two are highly correlated (0.98), however the RPM estimates are (almost) uniformly lower numerically. These numerical differences in the risk aversion parameter do not, however, significantly impact our loss aversion estimates, as shown in the right hand panel of Figure F.3. The correlation with our main estimates is again very high (0.99), but there is little evidence of a level shift. Substantively, our mixed RPM/RUM model produces a median loss aversion parameter of 1.02, with 49% loss tolerant, versus our preferred specification in which the median loss aversion parameter is 0.99 and 52% are loss tolerant.

Figure F.3: Estimates of loss aversion are similar using a random parameter model.



Notes: The left-hand panel of the figure displays the DOSE estimates of risk aversion using only lotteries over gains under the Random Parameter Model (RPM; y-axis) and Random Utility Model (RUM; x-axis). The right-hand panel displays estimates of loss aversion from the same two models.

F.2 Robustness of Correlations with Economic Preferences

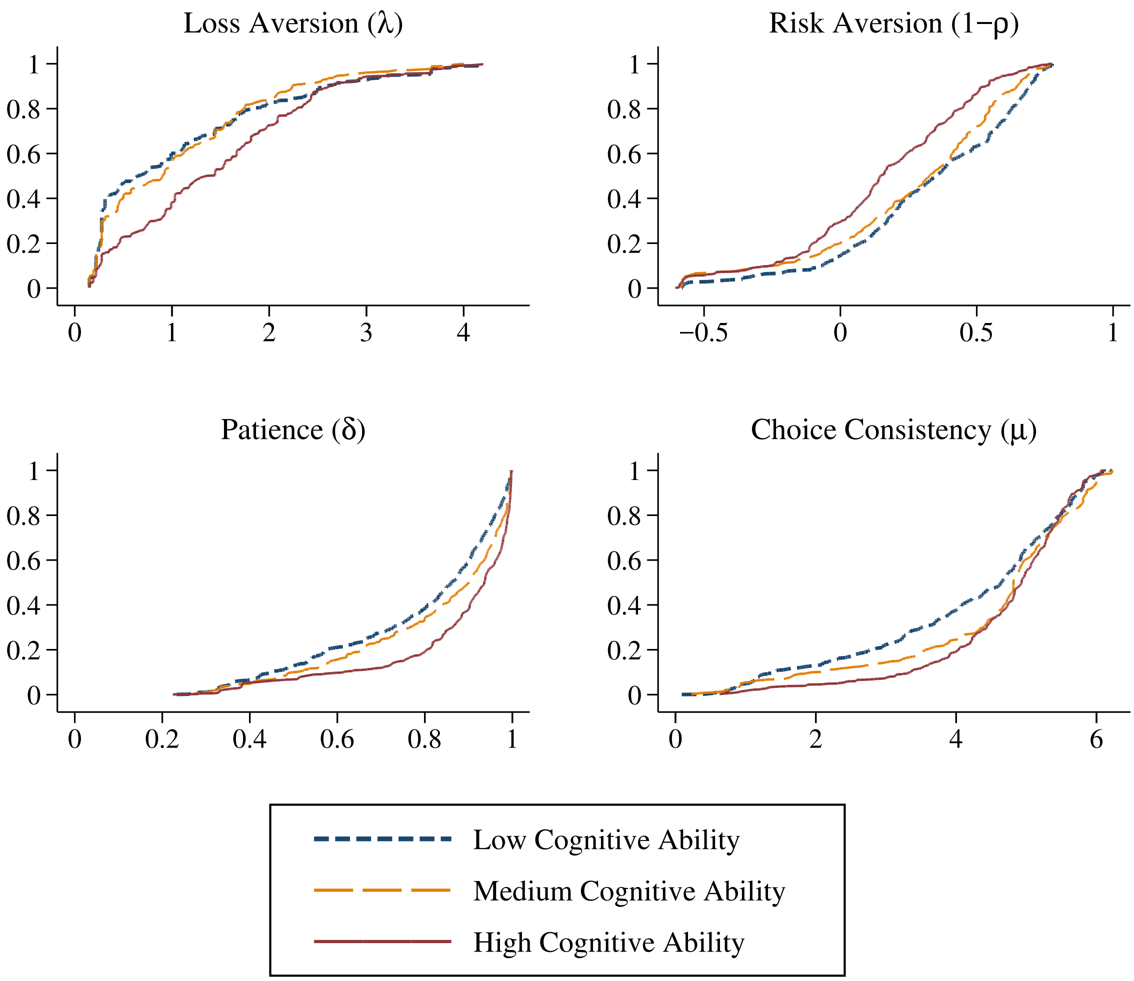
In this subsection we present extended versions of the correlation tables in Section 3, and robustness tests of the relationship between cognitive ability and economic preferences. The relationships between economic preferences and cognitive ability are robust to controlling for other individual characteristics, including both income and education. Indeed, most of the correlation between risk and loss aversion and education is explained by differences in cognitive ability. Finally, we turn to our choice data and investigate the relationship between cognitive ability and choosing the option that maximizes expected value. Consistent with the evidence in Figure 3, high cognitive ability participants are more likely to make EV-maximizing choices when the question offers only gains, but less likely to do so when the lottery contains a loss.

The distribution of preferences for high cognitive ability participants differs significantly from the rest of the population, as shown in Figure F.4. For each of the three economic preferences, the low and medium cognitive ability participants appear quite similar—but there is a first order stochastic dominance relation with high cognitive ability participants. Further, there is no evidence that the correlations between cognitive ability and preferences (discussed in Section 3.2) are driven by high cognitive ability individuals clustering at values near risk- or loss-neutrality.

In Table F.2 we compare the correlations when using the DOSE measure of risk aversion (column 1) and time preference (column 5) with the other risk and time measures in our survey. For risk aversion these alternative measures included two MPL modules, one relating to Willingness-to-Pay for a lottery (which we use in Section 3.3), and one relating to Willingness-to-Accept, as well as a risky project measure (Gneezy and Potters, 1997). For time preferences, as discussed in Section 2, we included two MPLs as well as the DOSE module.

The pattern of correlations is much stronger when using the DOSE measure than either MPL measure. As discussed in Section 3.3, the weak correlations with the MPL (WTP) measure are consistent with attenuation bias due to higher measurement error in the MPL. The weak pattern of correlations with the MPL (WTA) measure could also be explained by attenuation bias or could result from the WTA measure capturing a different dimension of risk preferences to the other risk measures in our survey (see Chapman et al., 2017). The risky project measure, which may suffer from less attenuation bias than the MPLs due to its simplicity, identifies a similar pattern of correlations to the DOSE risk aversion

Figure F.4: Economic preferences among low and high cognitive ability participants are clearly different.



Notes: Figures display the cumulative density of each preference parameter in the Wave 1 survey. Low, medium, and high cognitive ability are defined by the terciles of the distribution.

measure. The correlations between the risky project measure and individual characteristics consistently have the same sign, degree of statistical significance and magnitude as those with the DOSE estimates. The main exception are the correlations with cognitive ability, where DOSE identifies much stronger correlations than the project measure.

Loss aversion is also correlated with other individual characteristics not presented in the main text, as shown in Table F.3. More loss averse individuals are more likely to attend church, less likely to be white, and more likely to own a home. Further, the two component

Table F.2: Comparison of correlations between different risk and time measures and individual characteristics

	Risk Aversion				Patience	
	DOSE	MPL (WTP)	MPL (WTA)	Risky Project	DOSE	MPL
Cognitive ability	-0.21***	-0.04	0.01	-0.07***	0.18***	0.19***
	(.028)	(.028)	(.028)	(.029)	(.029)	(.027)
IQ	-0.18***	-0.05	0.00	-0.07**	0.14***	0.17***
	(.029)	(.028)	(.029)	(.030)	(.032)	(.029)
CRT	-0.18***	-0.02	0.02	-0.05*	0.18***	0.15***
	(.029)	(.027)	(.025)	(.029)	(.036)	(.025)
Income	-0.14***	-0.06*	0.00	-0.13***	0.11***	0.08***
	(.034)	(.034)	(.031)	(.035)	(.034)	(.032)
Education	-0.10***	-0.02	0.03	-0.09***	0.17***	0.11***
	(.033)	(.034)	(.028)	(.032)	(.037)	(.033)
Male	-0.10***	-0.06**	0.01	-0.12***	-0.02	0.02
	(.032)	(.030)	(.030)	(.032)	(.035)	(.034)
Age	0.01	0.05	-0.04	-0.02	0.18***	0.15***
	(.032)	(.031)	(.028)	(.032)	(.036)	(.034)
Stock Investor	-0.11***	-0.06**	-0.00	-0.12***	0.10***	0.09***
	(.029)	(.029)	(.026)	(.030)	(.031)	(.030)
Non-white	-0.07**	0.02	-0.04	-0.08***	0.13***	0.12***
	(.032)	(.030)	(.030)	(.032)	(.035)	(.034)
Own Home	0.10***	-0.01	0.01	-0.06*	-0.18***	-0.16***
	(.034)	(.032)	(.032)	(.034)	(.037)	(.038)
Employed	-0.03	-0.04	0.01	-0.09***	0.03	0.03
	(.031)	(.029)	(.029)	(.031)	(.035)	(.034)
Church Attendance	0.09***	-0.02	0.04	-0.04	0.01	-0.03
	(.032)	(.029)	(.030)	(.031)	(.034)	(.035)
Marital Status	-0.03	-0.02	-0.02	-0.01	-0.12***	-0.09***
	(.033)	(.031)	(.031)	(.032)	(.037)	(.036)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors are calculated by regressing the (standardized) preference parameter on each (standardized) individual characteristic, and are presented in parentheses. The survey contained two MPLs to measure each of WTA, WTP and time discounting. Correlations are estimated by stacking the two and clustering standard errors by participant.

parts of our cognitive ability measure have similar correlations with each of the economic preference variables, demonstrating that it is appropriate to combine the two.

Table F.3: Additional correlations between estimated DOSE parameters and individual characteristics

	Loss Aversion (λ)	Risk Aversion ($1-\rho$)	Patience (δ)	Choice Consistency (μ)
Cognitive Ability	0.21***	-0.21***	0.18***	0.16***
	(.030)	(.028)	(.029)	(.026)
IQ	0.18***	-0.18***	0.14***	0.13***
	(.033)	(.029)	(.032)	(.028)
CRT	0.19***	-0.18***	0.18***	0.15***
	(.029)	(.029)	(.036)	(.025)
Income	0.15***	-0.14***	0.11***	0.06*
	(.032)	(.034)	(.034)	(.033)
Education	0.13***	-0.10***	0.17***	0.11***
	(.032)	(.033)	(.037)	(.031)
Male	0.07**	-0.10***	-0.02	0.00
	(.033)	(.032)	(.035)	(.033)
Age	-0.11***	0.01	0.18***	0.06*
	(.033)	(.032)	(.036)	(.036)
Stock Investor	0.06**	-0.11***	0.10***	-0.01
	(.031)	(.029)	(.031)	(.032)
Non-white	-0.14***	0.10***	-0.18***	-0.07*
	(.033)	(.034)	(.037)	(.038)
Own Home	0.06*	-0.07**	0.13***	0.02
	(.033)	(.032)	(.035)	(.033)
Employed	0.06*	-0.03	0.03	0.04
	(.032)	(.031)	(.035)	(.031)
Church Attendance	-0.06*	0.09***	0.01	-0.02
	(.033)	(.032)	(.034)	(.032)
Marital Status	0.04	-0.03	-0.12***	-0.07**
	(.035)	(.033)	(.037)	(.033)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors are calculated by regressing the (standardized) preference parameter on each (standardized) individual characteristic, and are presented in parentheses.

The correlations with cognitive ability are robust to the inclusion of the other sociodemographic controls in Table 1—see Table F.4 and Table F.5. For each of the four preference parameters, the first specification includes only the attributes—age and sex—that are not

Table F.4: The correlations between cognitive ability and economic preferences are robust to the inclusion of demographic controls.

	Loss Aversion		Risk Aversion		Patience		Choice Consistency	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Ability	0.20***	0.15***	-0.20***	-0.17***	0.21***	0.17***	0.17***	0.15***
	(.031)	(.034)	(.028)	(.033)	(.028)	(.031)	(.028)	(.029)
Male	0.05	0.07	-0.12*	-0.14**	-0.12*	-0.14*	-0.07	-0.06
	(.067)	(.066)	(.063)	(.067)	(.066)	(.073)	(.067)	(.065)
Age	-0.09***	-0.09**	-0.01	-0.00	0.19***	0.18***	0.07**	0.04
	(.032)	(.036)	(.032)	(.036)	(.035)	(.041)	(.035)	(.040)
Education		0.06		-0.01		0.10**		0.05
		(.038)		(.037)		(.044)		(.033)
Income		0.08**		-0.08**		0.04		0.04
		(.036)		(.036)		(.037)		(.038)
Stock Investor		0.06		-0.08		-0.07		-0.19*
		(.081)		(.074)		(.079)		(.097)
Obs.	2000	1740	2000	1740	2000	1740	2000	1740
Adj. R^2	0.05	0.07	0.05	0.06	0.07	0.07	0.03	0.03

Note: Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. All continuous variables are standardized. Robust standard errors are displayed in parentheses. Missing observations are due to unreported incomes.

potentially endogenous to cognitive ability. The second specification then includes the remaining variables including, of most interest, education and income. In the specifications in Table F.4 all variables are included as continuous measures (except stock investor and male). The coefficients are standardized, and so are comparable to the correlations in Table 1. In all specifications the coefficient for cognitive ability is still strongly statistically significant and, compared to the other controls, large—although slightly lower than the simple correlations.

The results are similar when including all characteristics as categorical variables, as shown in Table F.5. These specifications allow for potential non-monotonic relationships, as well as having the added advantage of allowing us to include participants that did not report their income. The relationship with cognitive ability appears to be monotonic although, interestingly, the association with loss aversion seems limited to the top tercile of ability.

The results in Table F.4 and Table F.5 suggest that much of the correlation between education and both risk and loss aversion is explained by cognitive ability. To test that it is cognitive ability, and not one of the other controls, that weakens the association we

Table F.5: Participants in top tercile of cognitive ability are more loss averse.

	Loss Aversion		Risk Aversion		Patience		Choice Consistency	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive ability:								
Middle Tercile	-0.03	-0.03	-0.18**	-0.17**	0.19**	0.17**	0.26***	0.24***
	(.082)	(.081)	(.078)	(.078)	(.081)	(.080)	(.086)	(.083)
Top Tercile	0.35***	0.27***	-0.49***	-0.42***	0.48***	0.38***	0.40***	0.37***
	(.083)	(.085)	(.071)	(.075)	(.080)	(.083)	(.078)	(.076)
Age:								
36–50	-0.23**	-0.23**	0.04	0.05	0.15	0.14	0.09	0.07
	(0.106)	(0.103)	(.096)	(.095)	(0.110)	(0.109)	(.097)	(.096)
51–64	-0.32***	-0.35***	0.06	0.11	0.34***	0.32***	0.21**	0.21**
	(.092)	(.092)	(.092)	(.092)	(.098)	(.099)	(.094)	(.098)
65+	-0.26***	-0.28***	0.01	0.03	0.46***	0.50***	0.13	0.15
	(.095)	(.099)	(.094)	(.096)	(.098)	(0.100)	(0.100)	(0.106)
Male	0.08	0.08	-0.12*	-0.11*	-0.11*	-0.10	-0.06	-0.05
	(.067)	(.065)	(.062)	(.062)	(.066)	(.064)	(.067)	(.065)
Education:								
Some College		-0.02		-0.09		0.30***		0.04
		(.077)		(.075)		(.085)		(.076)
4-year College		0.15*		-0.07		0.26***		0.16**
		(.085)		(.079)		(.087)		(.077)
Income:								
2nd Quartile		-0.03		0.10		-0.12		0.11
		(.088)		(.095)		(0.102)		(.096)
3rd Quartile		0.20**		0.05		-0.09		0.05
		(.092)		(.091)		(0.109)		(.091)
4th Quartile		0.23**		-0.19**		0.14		0.04
		(0.100)		(.096)		(.093)		(0.101)
Unreported		0.34***		-0.04		-0.00		-0.26**
		(0.119)		(0.115)		(0.108)		(0.120)
Stock Investor		0.06		-0.12*		-0.01		-0.17*
		(.074)		(.068)		(.071)		(.089)
Obs.	2000	2000	2000	2000	2000	2000	2000	2000
Adj. R^2	0.05	0.07	0.05	0.06	0.06	0.08	0.03	0.05

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. All dependent variables are standardized. Robust standard errors are displayed in parentheses.

carry out additional specifications adding the variables one at a time—see Table F.6. For each preference parameter, we start by adding education and income separately, then both together and, finally, add cognitive ability. It is only when cognitive ability is added that the magnitude of the coefficient with education diminishes significantly—suggesting that cognitive ability jointly determines educational outcomes and these two preferences.

Table F.6: Much of the relationship between education and risk preferences is explained by differences in cognitive ability.

	Loss Aversion				Risk Aversion				Patience				Choice Consistency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Cognitive Ability:																
Middle Tercile				-0.03				-0.17**				0.17**				0.24***
				(.081)				(.078)				(.080)				(.083)
Top Tercile				0.27***				-0.42***				0.38***				0.37***
				(.085)				(.075)				(.083)				(.076)
Education:																
Some College	0.04		0.03	-0.02	-0.17**		-0.15**	-0.09	0.35***		0.35***	0.30***	0.11		0.09	0.04
	(.078)		(.077)	(.077)	(.077)		(.076)	(.075)	(.082)		(.083)	(.085)	(.080)		(.077)	(.076)
4-year College	0.32***		0.22***	0.15*	-0.28***		-0.17**	-0.07	0.40***		0.35***	0.26***	0.23***		0.25***	0.16**
	(.078)		(.084)	(.085)	(.077)		(.078)	(.079)	(.080)		(.084)	(.087)	(.074)		(.078)	(.077)
Income:																
2nd Quartile		-0.00	-0.01	-0.03		0.07	0.08	0.10		-0.07	-0.10	-0.12		0.15	0.13	0.11
		(.087)	(.087)	(.088)		(.098)	(.097)	(.095)		(0.107)	(0.103)	(0.102)		(.097)	(.096)	(.096)
3rd Quartile		0.26***	0.22**	0.20**		-0.00	0.03	0.05		0.01	-0.07	-0.09		0.12	0.07	0.05
		(.094)	(.093)	(.092)		(.093)	(.092)	(.091)		(0.112)	(0.109)	(0.109)		(.097)	(.094)	(.091)
Top Quartile		0.34***	0.26***	0.23**		-0.28***	-0.23**	-0.19**		0.27***	0.17*	0.14		0.14	0.07	0.04
		(.098)	(0.101)	(0.100)		(.096)	(.096)	(.096)		(.095)	(.096)	(.093)		(0.102)	(0.101)	(0.101)
Unreported		0.38***	0.36***	0.34***		-0.06	-0.06	-0.04		0.01	0.01	-0.00		-0.24*	-0.25**	-0.26**
		(0.122)	(0.121)	(0.119)		(0.115)	(0.114)	(0.115)		(0.111)	(0.110)	(0.108)		(0.126)	(0.125)	(0.120)
Stock Investor		0.11	0.07	0.06		-0.17***	-0.15**	-0.12*		0.06	0.01	-0.01		-0.11	-0.15*	-0.17*
		(.072)	(.074)	(.074)		(.067)	(.068)	(.068)		(.071)	(.072)	(.071)		(.086)	(.088)	(.089)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Obs.	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000
Adj. R^2	0.04	0.05	0.06	0.07	0.02	0.03	0.03	0.06	0.06	0.04	0.06	0.08	0.01	0.02	0.03	0.05

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. All dependent variables are standardized. Robust standard errors are displayed in parentheses.

F.3 Cognitive Ability and Expected Value Maximizing Choices

The relationship between expected value maximizing choices and cognitive ability discussed in Section 3.2 is robust to controlling for other socio-demographic characteristics, as shown in in Table F.7. Each observation in these regressions is an individual choice, with the (binary) dependent variable indicating whether a participant chose an option that maximized expected value. Specifications (1)–(4) relate to lotteries with losses, while specifications (5)–(8) relate to those with only gains. The omitted category for cognitive ability is low cognitive ability individuals.

The results of the regressions clearly reflect the pattern of choices presented in Figure 4. High cognitive ability participants are consistently more likely to choose an option with the highest expected value if that option involves accepting a lottery over gains or rejecting a lottery over losses. However, they are less likely to do so when the EV-maximizing option involves either accepting lottery over losses (that is, one with negative expected value) or accepting a sure amount over gains. This finding is robust to including controls for individual characteristics (specifications (2) and (6)), and question characteristics (specifications (3) and (7)). Finally, in the last specification, we allow for the relationship between education and making an EV-maximizing choice to vary according to whether that choice is a lottery or not. Again, the relationship with cognitive ability is largely unchanged.

Table F.7: High cognitive ability participants make fewer EV-maximizing decisions for lotteries with negative expected value.

	DV = Made Expected-Value-Maximizing Choice							
	Lotteries with Losses				Lotteries with Only Gains			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EV≤Sure Amount								
x Medium Cognitive Ability	0.062***	0.064***	0.041**	0.040**	-0.065**	-0.073**	-0.075**	-0.074**
	(.023)	(.023)	(.020)	(.020)	(.032)	(.032)	(.031)	(.031)
x High Cognitive Ability	0.171***	0.168***	0.114***	0.104***	-0.044	-0.063**	-0.070**	-0.066**
	(.025)	(.025)	(.023)	(.023)	(.031)	(.031)	(.030)	(.031)
EV>Sure Amount	0.264***	0.269***	0.152***	0.159***	-0.289***	-0.292***	-0.244***	-0.267***
	(.023)	(.022)	(.023)	(.026)	(.025)	(.025)	(.029)	(.033)
x Medium Cognitive Ability	0.013	0.013	-0.010	-0.009	0.040**	0.036**	0.027	0.026
	(.020)	(.020)	(.019)	(.019)	(.018)	(.018)	(.017)	(.017)
x High Cognitive Ability	-0.088***	-0.087***	-0.107***	-0.101***	0.120***	0.104***	0.087***	0.084***
	(.018)	(.019)	(.018)	(.018)	(.017)	(.018)	(.018)	(.018)
Some College		0.007	0.010	-0.001		-0.000	0.001	-0.042
		(.014)	(.012)	(.020)		(.016)	(.015)	(.030)
x EV>Sure Amount				0.018				0.066*
				(.029)				(.035)
4-year College		0.015	0.022	0.066***		0.020	0.026	0.017
		(.015)	(.014)	(.024)		(.017)	(.016)	(.032)
x EV≤Sure Amount				-0.069**				0.014
				(.031)				(.037)
Age (Standardized)		0.035***	-0.002	-0.002		0.003	-0.012*	-0.012*
		(.006)	(.006)	(.006)		(.007)	(.007)	(.007)
Male		0.012	0.010	0.010		0.004	0.003	0.004
		(.011)	(.010)	(.010)		(.013)	(.013)	(.013)
Income: 2nd Quartile		0.032**	0.031**	0.031**		0.058***	0.057***	0.057***
		(.016)	(.015)	(.015)		(.019)	(.018)	(.018)
Income: 3rd Quartile		-0.003	0.000	-0.001		0.033*	0.035*	0.036*
		(.016)	(.015)	(.015)		(.019)	(.019)	(.019)
Income: 4th Quartile		-0.011	-0.005	-0.005		0.067***	0.068***	0.067***
		(.017)	(.016)	(.016)		(.021)	(.020)	(.020)
Income: Unstated		-0.006	-0.010	-0.010		0.022	0.020	0.020
		(.020)	(.018)	(.018)		(.022)	(.021)	(.021)
Lottery Prize ($)			0.031***	0.031***			-0.004	-0.004
			(.002)	(.002)			(.004)	(.004)
|EV - Sure Amount| ($)			-0.021***	-0.020***			-0.023***	-0.023***
			(.004)	(.004)			(.007)	(.007)
Response Time: Quartile 2			0.144***	0.143***			0.061***	0.062***
			(.013)	(.013)			(.020)	(.020)
Response Time: Quartile 3			0.220***	0.221***			0.082***	0.084***
			(.014)	(.014)			(.019)	(.019)
Response Time: Quartile 4			0.216***	0.217***			0.130***	0.131***
			(.015)	(.015)			(.020)	(.020)
Obs.	11154	11154	11154	11154	8846	8846	8846	8846
Adj. R^2	0.04	0.04	0.10	0.11	0.04	0.05	0.06	0.06

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors are clustered by participant and displayed in parentheses.

F.4 Additional Tests of Fatigue and Inattention

Neither survey fatigue or inattention appear to explain our results. The pattern of economic preferences does not change according to the position of the DOSE module, suggesting that participants did not become fatigued later in the survey. Nor do participants appear to choose more randomly later in the DOSE module, as would be expected if they get bored of the questions. There is also little evidence that participants get tired and click through the survey quickly without paying attention: there is little correlation between response time and our three economic preference parameters. Further, we observe the same relationships between economic preferences and individual characteristics when excluding the fastest participants. Thus, boredom causing unusually fast choice, which then leads to our results does not seem a plausible explanation for our conclusions.

It does not appear that survey fatigue affected our estimate of any of the four DOSE parameters, as shown in Figure F.5. The position of the two DOSE modules was randomized across participants, and could appear in the third, fourth, fifth or sixth position in the survey (the two modules always appeared together). The distribution of parameters is similar regardless of the position in both waves of the survey. Thus it does not appear that participants' behavior changed when they took DOSE further into the survey, suggesting that fatigue is not a major factor in our results.

We can also look for signs of fatigue within the DOSE module itself by observing that if participants start choosing randomly then the Bayesian prior will be "surprised" more often. That is, at each point in the question selection process, DOSE's prior has an alternative that it thinks the participant will choose. When the participant chooses an option the procedure thinks should only be a 30% chance, we encode that as $50 - 30 = 20$ percentage points surprising. If, on the other hand, the participant chooses an option the procedure things is 70%, then we encode this as $50 - 70 = -20\%$ surprising. If participants are getting fatigued, we would expect them to start choosing randomly, that is, make more surprising choices (on average) later in the module.

As can be seen in Figure F.6, the surprise actually decreases slightly across DOSE questions, leading to the conclusion that, if anything, participants are making more consistent choices towards the end of the DOSE procedure than at the beginning. Further, this pattern is consistent regardless of cognitive ability and whether participants are classified as loss tolerant. As such, fatigued participants making choices at random do not appear to be driving the high degree of loss tolerance we observe.

While there is little evidence of survey fatigue, it is possible that some participants were inattentive throughout the survey. We examine this possibility next by re-analyzing our data while leaving out those who were most likely to have given up and rushed through the survey. The same results hold, implying that neither confusion nor inattentiveness, nor giving up, is likely to explain many of the choices we see.

One might expect that bored or inattentive participants would just click through screens quickly but, as shown in Figure F.7, our results are largely unchanged when removing the fastest responses. In this figure, we first look at the slowest 80% of participants, then the slowest 60%, and so on. The distributions overlap almost entirely—and the median loss aversion parameter consistently remains very similar in the whole sample (1.03 or below): fast response or inattention cannot be said to be the explanation for loss tolerance.

Correlations between the economic preferences and other characteristics are also robust to removing "fast responders", as shown in the following four tables. Tables F.8 and F.9 present the results when removing the fastest 50% of participants on the survey. Tables F.10 and F.11 then do the same, but removing the fastest 50% on the DOSE module. In both cases, the relationships are similar to in the whole sample (Table F.3).

A final manifestation of inattention might be choosing the same option in each question: either the lottery (always listed first), or the sure amount (always listed second). However, there is little evidence of this pattern of inattention in our results: fewer than 6% participants chose the same option in all ten question rounds. While we cannot rule out that some participants rapidly clicked through the DOSE module, such behavior does not appear to

Table F.8: Correlations with individual characteristics are similar when removing fastest 50% of participants on entire survey.

	Loss Aversion (λ)	Risk Aversion ($1-\rho$)	Patience (δ)	Choice Consistency (μ)
Cognitive Ability	0.22***	-0.21***	0.17***	0.18***
	(.039)	(.039)	(.042)	(.035)
IQ	0.20***	-0.18***	0.09*	0.15***
	(.039)	(.040)	(.050)	(.035)
CRT	0.18***	-0.18***	0.23***	0.16***
	(.045)	(.043)	(.040)	(.035)
Income	0.12***	-0.14***	0.10**	0.11***
	(.045)	(.046)	(.045)	(.040)
Education	0.07	-0.14***	0.15***	0.10***
	(.041)	(.044)	(.043)	(.039)
Male	0.08*	-0.12***	-0.04	0.02
	(.043)	(.045)	(.051)	(.043)
Age	-0.08*	0.04	0.18***	-0.00
	(.047)	(.047)	(.057)	(.046)
Stock Investor	0.13***	-0.13***	0.08*	-0.01
	(.043)	(.042)	(.046)	(.042)
Non-white	-0.16***	0.10**	-0.18***	-0.02
	(.045)	(.049)	(.055)	(.046)
Own Home	0.07*	-0.07*	0.11**	0.05
	(.044)	(.044)	(.052)	(.045)
Employed	0.08**	-0.03	-0.01	0.07*
	(.043)	(.044)	(.050)	(.039)
Church Attendance	0.00	0.08*	-0.00	0.06
	(.044)	(.044)	(.050)	(.045)
Marital Status	0.02	0.00	-0.14***	-0.06
	(.048)	(.047)	(.059)	(.042)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors are calculated by regressing the (standardized) preference parameter on each (standardized) individual characteristic, and are presented in parentheses.

affect our results.

Table F.9: Correlations with individual characteristics are similar when removing fastest 50% of participants on entire survey and including controls.

	Loss Aversion		Risk Aversion		Patience		Choice Consistency	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Ability	0.21***	0.18***	-0.19***	-0.12**	0.20***	0.18***	0.18***	0.16***
	(.039)	(.047)	(.041)	(.048)	(.042)	(.046)	(.036)	(.040)
Male	0.07	0.11	-0.16*	-0.22**	-0.14	-0.20*	-0.03	-0.11
	(.085)	(.092)	(.087)	(.091)	(.097)	(0.105)	(.087)	(.090)
Age	-0.05	-0.09*	0.01	0.02	0.20***	0.20***	0.02	-0.04
	(.045)	(.055)	(.048)	(.052)	(.057)	(.066)	(.046)	(.049)
Education		-0.03		-0.03		0.10**		0.05
		(.046)		(.053)		(.051)		(.043)
Income		0.03		-0.05		0.03		0.08*
		(.051)		(.049)		(.051)		(.044)
Stock Investor		0.21*		-0.12		-0.11		-0.19*
		(0.109)		(.096)		(0.106)		(0.102)
Obs.	993	861	993	861	993	861	993	861
Adj. R^2	0.05	0.07	0.05	0.05	0.07	0.08	0.03	0.04

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. All dependent variables are standardized. Robust standard errors are displayed in parentheses. Missing observations are due to unreported incomes.

F.5 Choice Consistency and Response Time

We now show that that controlling for choice consistency helps identify a pattern of correlations even when restricting the sample to those answering very fast—and so who might be thought to be paying little attention. In the left hand panel of Figure F.8 we show the pattern of correlations restricting the sample to first those answering the risk MPL module quickly and in the right hand panel we present the correlations for those answering the whole survey quickly (quickly being defined as below the respective median). In both cases we compare the correlations for all participants to those in the high consistency group.

In both panels there is more evidence of correlations after restricting the sample to high consistency participants. The magnitude of the correlations in frequently higher, and several emerge as statistically significant once only high consistency participants are considered. The magnitude of the correlations is, in fact, similar to those in Figure 5, although the standard errors are larger (explained by the fact the sample is half as large). The choice consistency measure appears, then, to be distinguishing participants that answer accurately but rapidly—

Table F.10: Correlations with individual characteristics are similar when removing fastest 50% of participants on DOSE module.

	Loss Aversion (λ)	Risk Aversion ($1-\rho$)	Patience (δ)	Choice Consistency (μ)
Cognitive Ability	0.26***	-0.21***	0.18***	0.16***
	(.037)	(.039)	(.041)	(.036)
IQ	0.25***	-0.20***	0.12***	0.12***
	(.037)	(.041)	(.049)	(.036)
CRT	0.19***	-0.16***	0.22***	0.15***
	(.044)	(.043)	(.040)	(.037)
Income	0.12***	-0.12***	0.10**	0.07*
	(.044)	(.046)	(.046)	(.041)
Education	0.13***	-0.14***	0.16***	0.09*
	(.041)	(.039)	(.045)	(.047)
Male	0.08*	-0.10**	-0.04	-0.01
	(.042)	(.044)	(.050)	(.046)
Age	-0.05	0.07	0.19***	0.04
	(.047)	(.045)	(.056)	(.057)
Stock Investor	0.12***	-0.09***	0.15***	-0.03
	(.039)	(.037)	(.038)	(.040)
Non-white	-0.18***	0.12***	-0.19***	-0.01
	(.044)	(.047)	(.056)	(.053)
Own Home	0.08*	-0.11***	0.13***	0.00
	(.043)	(.043)	(.051)	(.048)
Employed	0.12***	-0.03	0.02	0.04
	(.042)	(.042)	(.049)	(.042)
Church Attendance	-0.03	0.05	0.03	-0.02
	(.043)	(.044)	(.049)	(.046)
Marital Status	-0.02	0.01	-0.18***	-0.07
	(.047)	(.046)	(.056)	(.048)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors are calculated by regressing the (standardized) preference parameter on each (standardized) individual characteristic, and are presented in parentheses.

whose responses include meaningful information—from those that answer quickly due to a lack of care or attention.

Table F.11: Correlations with individual characteristics are similar when removing fastest 50% of participants on DOSE module and including controls.

	Loss Aversion		Risk Aversion		Patience		Choice Consistency	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Ability	0.26***	0.20***	-0.20***	-0.16***	0.23***	0.19***	0.17***	0.12***
	(.038)	(.043)	(.040)	(.050)	(.040)	(.043)	(.042)	(.035)
Male	0.05	0.13	-0.12	-0.13	-0.16*	-0.20**	-0.08	-0.10
	(.083)	(.091)	(.084)	(.090)	(.091)	(0.102)	(.098)	(.090)
Age	-0.01	-0.06	0.04	0.01	0.22***	0.18***	0.06	-0.03
	(.045)	(.054)	(.046)	(.053)	(.054)	(.065)	(.058)	(.052)
Education		0.05		-0.09*		0.09*		0.03
		(.045)		(.048)		(.053)		(.042)
Income		0.01		-0.03		0.01		0.06
		(.048)		(.050)		(.049)		(.040)
Stock Investor		0.12		-0.03		0.03		-0.17*
		(.097)		(.079)		(.084)		(.089)
Obs.	1012	875	1012	875	1012	875	1012	875
Adj. R^2	0.07	0.08	0.05	0.05	0.09	0.08	0.03	0.03

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. All dependent variables are standardized. Robust standard errors are displayed in parentheses. Missing observations are due to unreported incomes.

F.6 Results using Wave 2 Survey Sample

The survey results are similar when using data from the second wave of the survey. As shown in Tables F.12 and F.13, the correlations with other sociodemographic variables are of similar magnitude and direction to those in the first wave. Similarly, the pattern of choices of low and high cognitive ability participants follow a similar pattern (Figure F.9 and Table F.14).

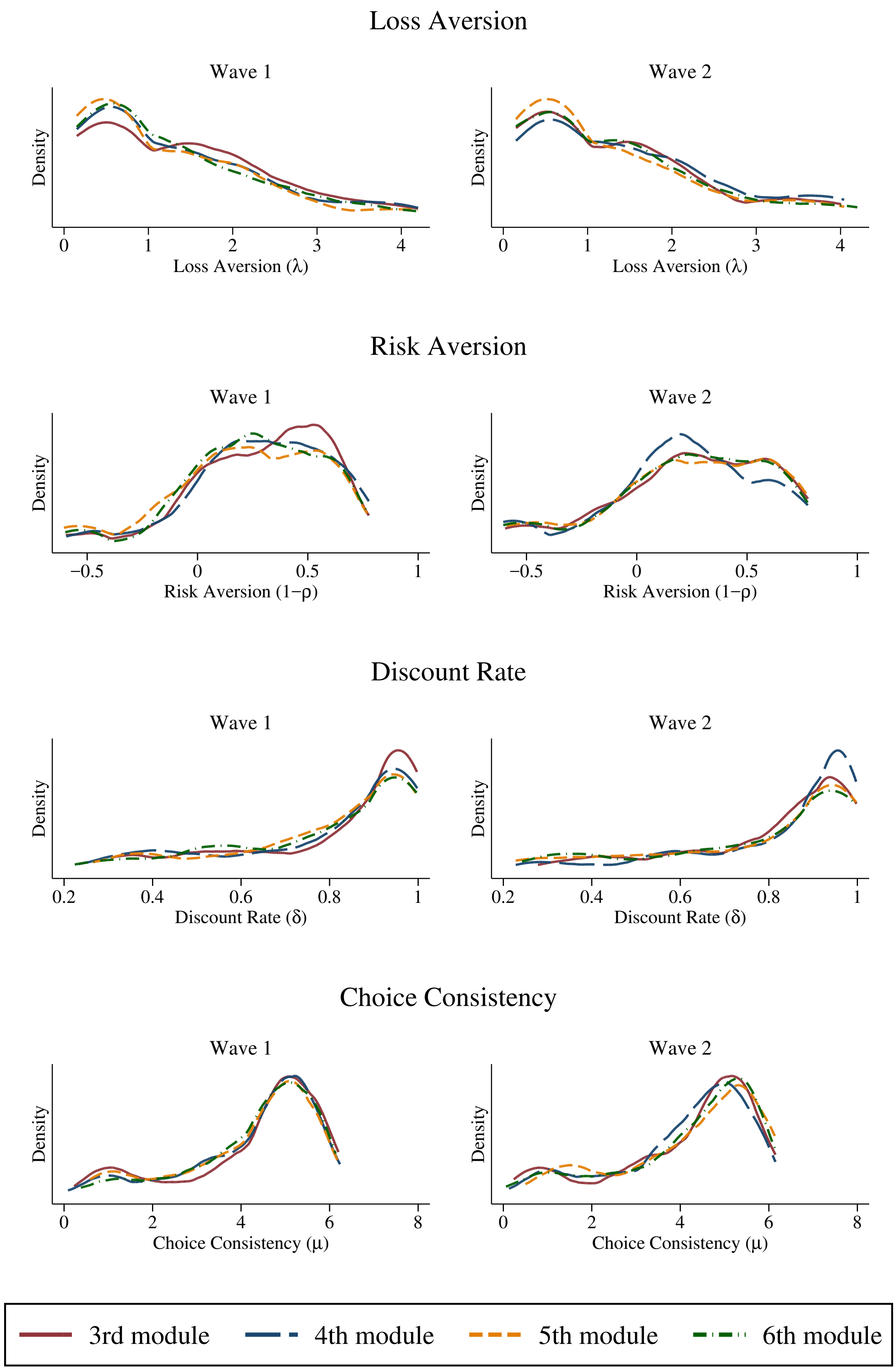
The only notable exception is that restricting the Wave 2 sample to high consistency participants does not recover a statistically significant correlation between the risk aversion MPL measure and cognitive ability (see Figure F.10)—a difference that is probably explained by differential attrition. The first and third panels of the figure are very similar to those in Figure 5: the DOSE risk aversion estimates are consistently correlated with individual characteristics, and the MPL measure is not. The middle panel shows that there is still a pattern of higher correlations between the MPL measure and other individual characteristics after removing inconsistent participants; however the correlation with cognitive ability (and also stock ownership) is not distinguishable from zero at conventional levels. The reason appears to be that lower cognitive ability is associated with higher drop out rates between survey waves. In contrast, none of the DOSE measures—or education where we do see a higher correlation—is correlated with attrition. The reduced variability in the sample then reduces the ability to identify a genuine correlation.

Table F.12: Correlations between estimated DOSE parameters and individual characteristics in Wave 2 data are similar to those in Wave 1.

	Loss Aversion (λ)	Risk Aversion ($1-\rho$)	Patience (δ)	Choice Consistency (μ)
Cognitive Ability	0.26***	-0.23***	0.25***	0.18***
	(0.033)	(0.032)	(0.031)	(0.034)
IQ	0.22***	-0.20***	0.21***	0.15***
	(0.033)	(0.035)	(0.031)	(0.035)
CRT	0.23***	-0.21***	0.23***	0.17***
	(0.035)	(0.033)	(0.032)	(0.034)
Income	0.12***	-0.19***	0.20***	0.15***
	(0.039)	(0.042)	(0.048)	(0.040)
Education	0.14***	-0.15***	0.21***	0.17***
	(0.034)	(0.036)	(0.040)	(0.035)
Male	0.06	-0.04	0.02	0.01
	(0.040)	(0.040)	(0.043)	(0.042)
Age	-0.11***	0.00	0.08*	0.08*
	(0.042)	(0.040)	(0.047)	(0.043)
Stock Investor	0.04	-0.16***	0.17***	0.08**
	(0.035)	(0.038)	(0.038)	(0.037)
Non-white	-0.13***	0.11***	-0.23***	-0.17***
	(0.044)	(0.042)	(0.049)	(0.048)
Own Home	-0.00	-0.10***	0.13***	0.05
	(0.041)	(0.040)	(0.044)	(0.043)
Employed	-0.01	-0.06	0.11***	0.06
	(0.040)	(0.040)	(0.041)	(0.041)
Church Attendance	0.00	0.04	-0.03	0.00
	(0.040)	(0.041)	(0.039)	(0.039)
Marital Status	0.09**	-0.02	-0.05	-0.12***
	(0.042)	(0.043)	(0.043)	(0.043)

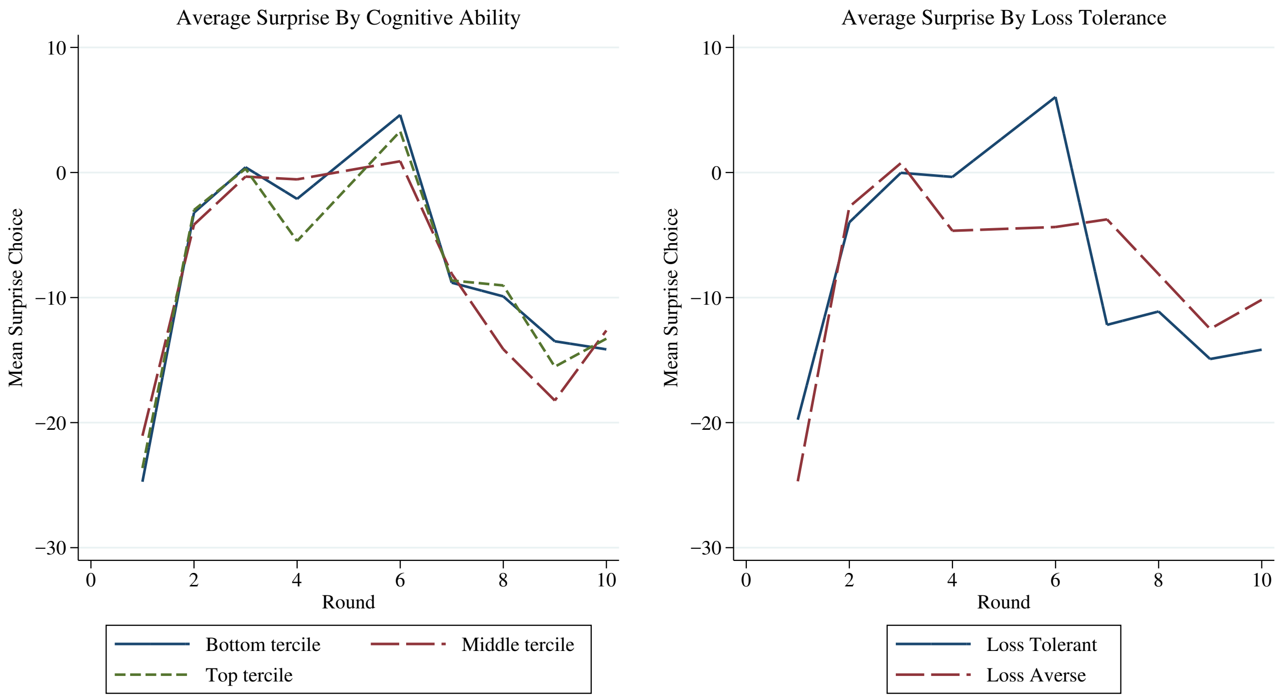
Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors are calculated by regressing the (standardized) preference parameter on each (standardized) individual characteristic, and are presented in parentheses.

Figure F.5: The parameter distributions are similar regardless of the position in the survey.



Notes: Figure displays the kernel density of each parameter using an Epanechnikov kernel, according to the position of the DOSE module in the survey.

Figure F.6: No evidence of fatigue within the DOSE module.



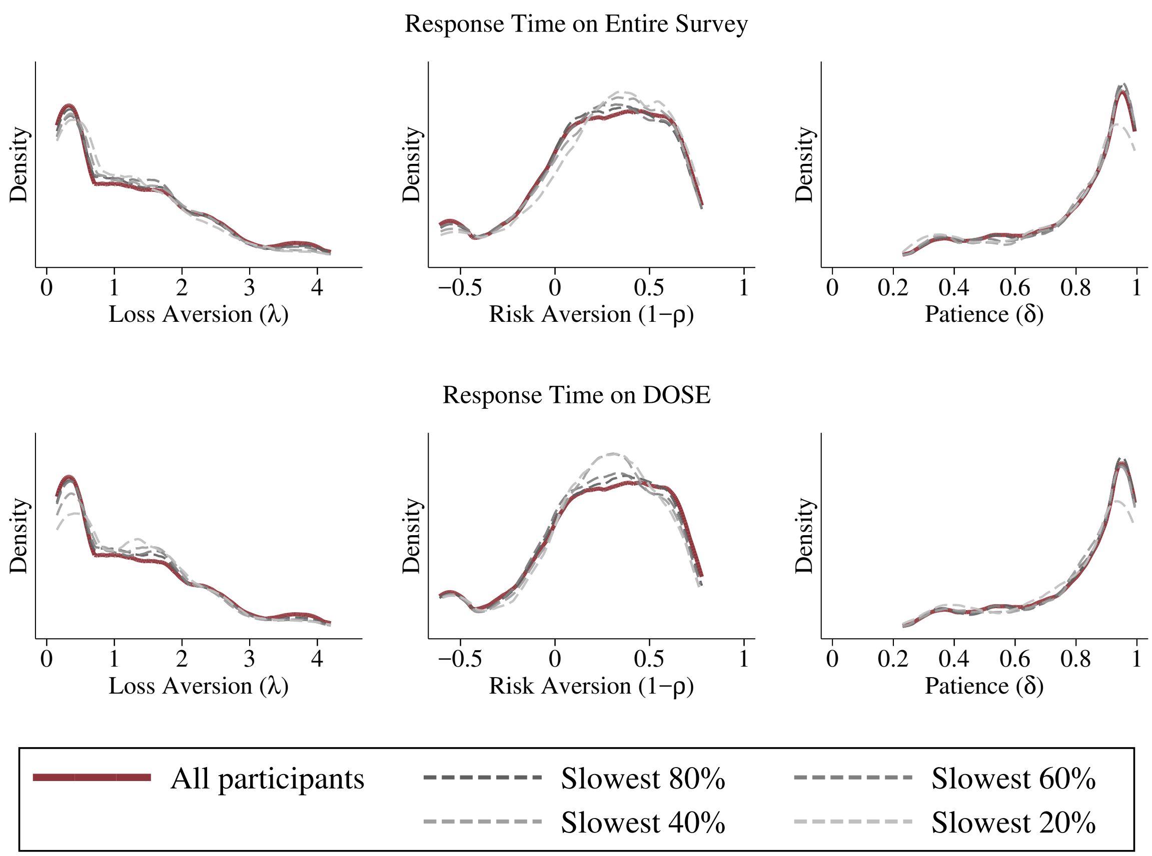
Notes: The figure plots the mean "surprise" of individuals' choice in each round, where surprise is calculated using the DOSE priors before each question. Questions with losses were allowed only from round 5 onwards, leading to considerable updating and a high level of surprise. Consequently data from round 5 is excluded from the figure.

Table F.13: Correlations between cognitive ability and economic preferences in Wave 2 are similar after including demographic controls.

	Loss Aversion		Risk Aversion		Patience		Choice Consistency	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Ability	0.25***	0.22***	-0.23***	-0.18***	0.27***	0.20***	0.19***	0.16***
	(0.035)	(0.042)	(0.034)	(0.035)	(0.034)	(0.035)	(0.037)	(0.037)
Male	0.04	0.10	-0.01	-0.04	-0.04	-0.01	-0.03	-0.10
	(0.079)	(0.082)	(0.079)	(0.079)	(0.083)	(0.084)	(0.085)	(0.090)
Age	-0.09**	-0.05	-0.02	0.02	0.11**	0.04	0.10**	0.09**
	(0.040)	(0.043)	(0.039)	(0.043)	(0.046)	(0.052)	(0.042)	(0.046)
Education		0.07*		-0.03		0.08*		0.10**
		(0.040)		(0.042)		(0.043)		(0.039)
Income		0.04		-0.08*		0.08		0.07
		(0.041)		(0.043)		(0.049)		(0.043)
Stock Investor		-0.06		-0.21**		0.11		-0.05
		(0.080)		(0.100)		(0.097)		(0.093)
Obs.	1465	1271	1465	1271	1465	1271	1465	1271
Adj. R^2	0.07	0.08	0.05	0.08	0.07	0.09	0.04	0.06

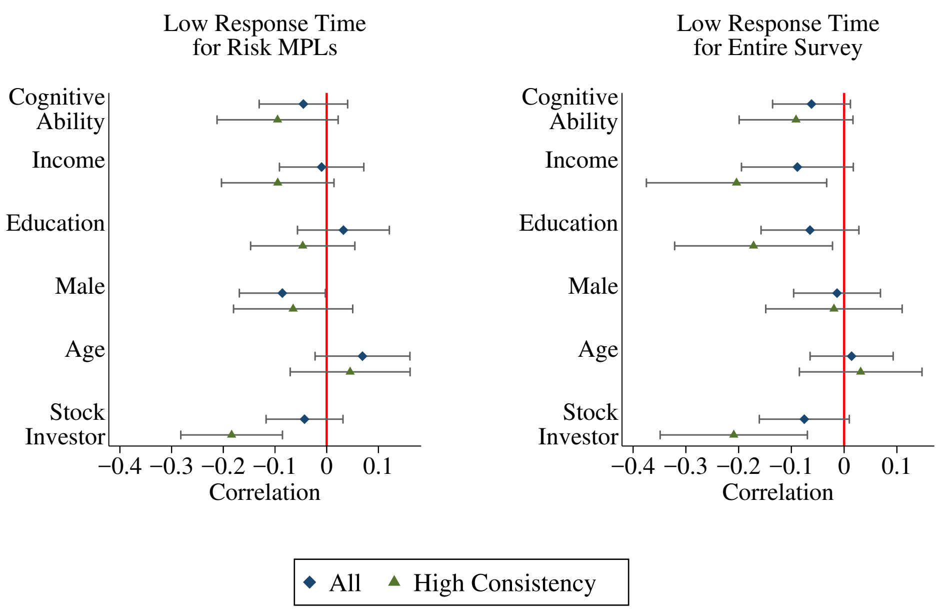
Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. All continuous variables are standardized. Robust standard errors are displayed in parentheses. Missing observations are due to unreported incomes.

Figure F.7: Distributions are similar when removing participants with short response times.



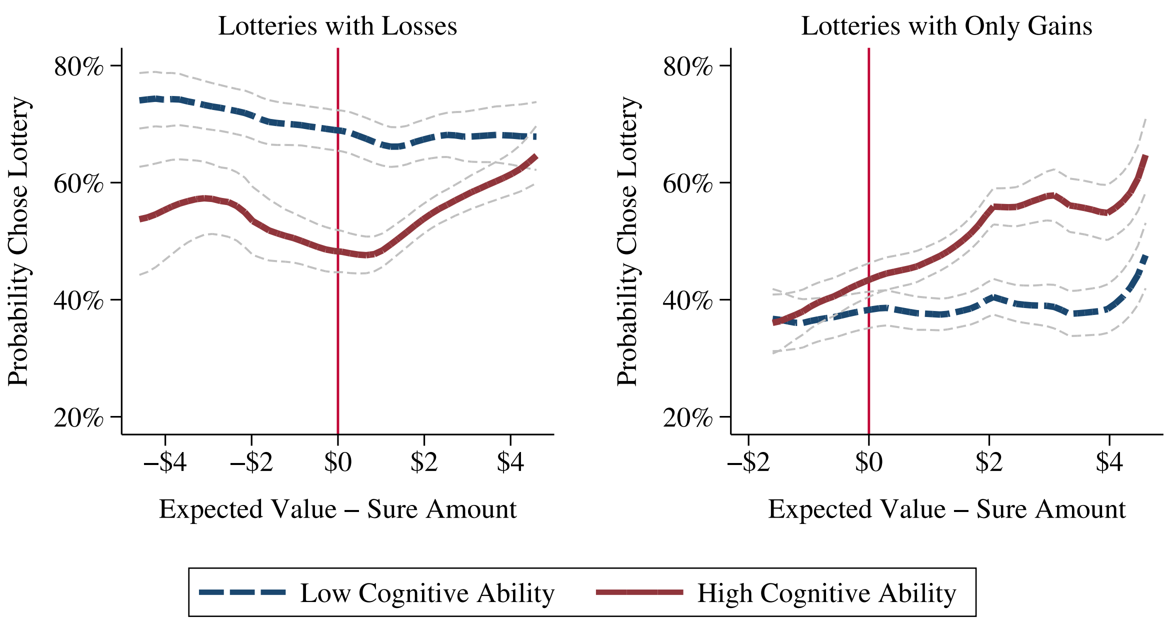
Notes: Plotted using Epanechnikov kernel, with bandwidth chosen by rule-of-thumb estimator.

Figure F.8: Accounting for choice consistency leads to a clearer pattern of correlations even after removing very fast responses.



Notes: The left panel includes only participants below the median response time on the risk MPL module. The right panel includes only participants below the median response time on the entire survey. "High Consistency" refers to those with choice consistency above the median. The survey contained two MPL measures of risk preference. Correlations are estimated by stacking the two and clustering standard errors by participant.

Figure F.9: Pattern of individual choices is similar in the Wave 2 data.



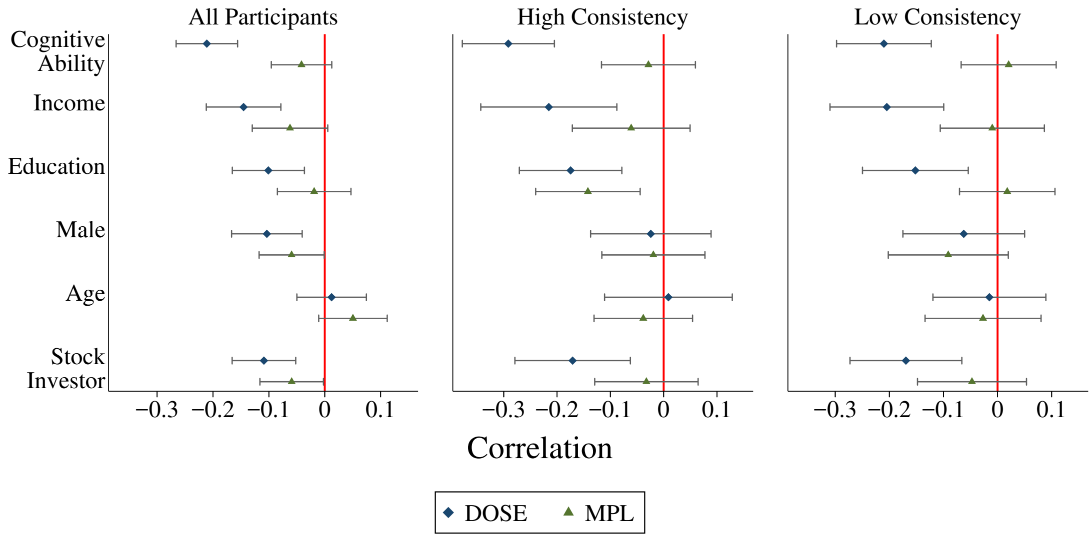
Notes: Figure displays the Nadaraya-Watson estimator (local mean smoothing) estimator (bandwidth 1) with Epanechnikov kernel. Grey dotted lines represent 95% confidence intervals, constructed with 10,000 clustered bootstrap replications. "Lotteries with losses" identifies questions where participants chose between \$0 for sure and a 50:50 lottery between a gain and a loss (both amounts varying). "Lotteries with only gains" identifies questions where participants chose between a varying, strictly positive, sure amount and a 50:50 lottery between a varying gain and \$0. Participants were asked a personalized question sequence, and so the set of possible choices varied across individuals. High and low cognitive ability refer to the top and bottom terciles respectively.

Table F.14: High cognitive ability participants make fewer EV-maximizing decisions for lotteries with negative expected value using Wave 2 data.

	DV = Made Expected-Value-Maximizing Choice							
	Lotteries with Losses				Lotteries with Only Gains			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EV≤Sure Amount								
x Medium Cognitive Ability	0.056**	0.058**	0.049**	0.047**	-0.032	-0.036	-0.037	-0.033
	(.026)	(.026)	(.024)	(.024)	(.038)	(.038)	(.037)	(.037)
x High Cognitive Ability	0.172***	0.171***	0.127***	0.120***	0.018	0.008	0.002	0.011
	(.031)	(.032)	(.029)	(.030)	(.040)	(.041)	(.040)	(.041)
EV>Sure Amount	0.260***	0.261***	0.151***	0.156***	-0.272***	-0.273***	-0.230***	-0.246***
	(.027)	(.027)	(.027)	(.031)	(.032)	(.032)	(.037)	(.042)
x Medium Cognitive Ability	-0.035	-0.032	-0.037*	-0.036	0.036*	0.032	0.031	0.029
	(.023)	(.023)	(.022)	(.022)	(.020)	(.020)	(.019)	(.019)
x High Cognitive Ability	-0.115***	-0.109***	-0.124***	-0.120***	0.138***	0.127***	0.115***	0.110***
	(.023)	(.023)	(.022)	(.023)	(.022)	(.023)	(.022)	(.022)
Some College		0.023	0.020	0.015		0.025	0.022	0.005
		(.017)	(.016)	(.025)		(.019)	(.018)	(.037)
x EV>Sure Amount				0.008				0.025
				(.036)				(.042)
4-year College		-0.018	-0.016	0.011		0.013	0.012	-0.018
		(.018)	(.017)	(.028)		(.021)	(.020)	(.040)
x EV≤Sure Amount				-0.040				0.045
				(.037)				(.044)
Age (Standardized)		0.034***	-0.005	-0.005		0.003	-0.008	-0.008
		(.007)	(.007)	(.007)		(.008)	(.009)	(.009)
Male		0.004	-0.000	-0.001		-0.003	-0.004	-0.004
		(.013)	(.012)	(.012)		(.016)	(.015)	(.015)
Income: 2nd Quartile		0.009	0.008	0.009		-0.013	-0.013	-0.013
		(.020)	(.019)	(.019)		(.023)	(.022)	(.022)
Income: 3rd Quartile		0.003	0.008	0.008		0.002	0.001	0.001
		(.020)	(.018)	(.018)		(.022)	(.022)	(.022)
Income: 4th Quartile		0.045**	0.048***	0.048**		0.045*	0.043*	0.043*
		(.020)	(.019)	(.019)		(.025)	(.024)	(.024)
Income: Unstated		0.009	0.003	0.003		-0.003	-0.004	-0.003
		(.023)	(.022)	(.022)		(.025)	(.025)	(.025)
Lottery Prize ($)			0.030***	0.030***			-0.002	-0.002
			(.003)	(.003)			(.005)	(.005)
|EV - Sure Amount| ($)			-0.018***	-0.018***			-0.026***	-0.026***
			(.004)	(.004)			(.008)	(.008)
Response Time: Quartile 2			0.136***	0.136***			0.049**	0.048**
			(.015)	(.015)			(.023)	(.023)
Response Time: Quartile 3			0.207***	0.208***			0.095***	0.096***
			(.016)	(.016)			(.023)	(.023)
Response Time: Quartile 4			0.198***	0.199***			0.082***	0.082***
			(.018)	(.018)			(.025)	(.025)
Obs.	8151	8151	8151	8151	6499	6499	6499	6499
Adj. R^2	0.03	0.04	0.09	0.09	0.05	0.05	0.06	0.06

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors are clustered by participant and displayed in parentheses.

Figure F.10: Correlations with MPL and DOSE risk aversion measures using Wave 2 data.



Notes: Figure displays correlations between the DOSE and MPL measures of risk aversion and individual characteristics. The left hand panel includes all participants in the Wave 2 survey, while the middle (right) panel restricts the sample to those above (below) the median in the choice consistency variable. The survey contained two MPL measures of risk preference. Correlations are estimated by stacking the two and clustering standard errors by participant.