Online Appendix—Not Intended for Publication

A Potential for Manipulation

In practice, the theoretical possibility that DOSE could be strategically manipulated—like most dynamic designs—is unlikely to be an important concern. 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 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. The short question sequence and general population sample in our survey means that neither of these conditions are likely to be met. Further, DOSE could easily be adapted to combat manipulation in other settings where there is greater likelihood of strategic behavior.

The incentive for participants to game adaptive procedures is small, and there is little evidence of such behavior in practice. Ray et al. (2012) find that 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. Further, they report that informing participants in a laboratory experiment that the question sequence is manipulable did not increase average earnings, and most participants stated that they did not try and manipulate the system. There is also little evidence of manipulation—identified by behavior changing significantly between early and later question rounds—in their data.

Strategic behavior is particularly unlikely in our survey, given that we ask only a short question sequence and participants are unlikely to have previous experience of economic experiments. The short question sequence means that there is little opportunity for learning about the adaptive nature of the question selection criteria, particularly as the type of choices. In addition, 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. There is no reason, therefore, to believe that they would recognize the opportunity to manipulate the question selection procedure.

DOSE can easily be adapted to address concerns about manipulation in settings where "gaming" is seen as particularly likely—for example, in the presence of large stakes. We suggest two possible remedies, but others are possible.[1] First, the actual question chosen for payment can be randomly selected from *all* possible questions after the personalized question sequence is completed. If that question has already been answered, the answer determines the payment. If not, the participant answers it, and this answer determines payoff. In the second remedy, the answers the participant provides to the DOSE questions determine parameter estimates, which are used to construct the choice to one of the unanswered questions, which is used for payment. Both designs mean that truthful response is incentive compatible—if the model of preferences used is correct. The latter, in the presence of risk aversion, should even increase incentives for consistent choices. However, both may reduce the strength of participant incentives as each question has a lower probability of influencing the final payoff.

An final approach would involve using DOSE to assess the extent of strategic manipulation among respondents. 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).

B 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

[1] The first remedy was suggested by Kate Johnson, and the second by Ian Krajbich.

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 4.3, 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 0.31 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 average coefficient using DOSE is 0.05 in the lab, which is 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 coefficient 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 2), 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 4.3, 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.[2] 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 4.4, 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 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

[2] 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)

stable over time due to reduced measurement error.[3]

C Additional Details of Survey Implementation

This subsection presents further details of the implementation of DOSE in the online survey, including both the question selection procedure and the estimation of the DOSE parameters presented in the main text. As discussed in Section 2.3, the survey included two DOSE survey modules. The first module focused on risk preferences, and consisted of 10 binary choices between a sure amount and a lottery.[4] The second module focused on time preferences and consisted of a further 10 binary choices between differing amounts at two different dates.[5]

Conceptually, the question selection differed from the outline presented in Section 2.3 in two ways. First, 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. Second, the survey questions were selected using the Kullback-Leibler information criterion suggested by El-Gamal and Palfrey (1996). The KL criterion they suggest captures the distance between the parameter vector k and all other vectors if k is the correct parameter vector. That is, it is:

$$KL(Q_i) = \sum_{k \in \mathcal{K}} \sum_{a \in A} \log \left(\frac{(1-p_k) l_k(a; Q_i)}{\sum_{j \neq k}^{\mathcal{K}} p_j l_j(a; Q_i)} \right) p_k l_k(a; Q_i) \tag{4}$$

This formula is very similar to (3). The main difference is that the likelihood of model k is

[3] 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.

[4] The set of potential questions used for the risk module included gains between \$1 and \$10 in increments of \$0.50, and sure amounts and losses varying ranging from \$0.50 to \$10 in increments of \$0.10. Questions were excluded if one choice was first order stochastically dominated for all values of the prior distribution.

[5] The questions for the time module were drawn from a question set consisting of monetary amounts between \$1 and \$10, in \$1 increments, and time periods of 0, 1, 3, 5, 7, 9, 10, 12, 16, 21, 28, 35, 42, 49, 56, 60, 70, 80, 90 days.

not included in the denominator—reflecting the fact that the El-Gamal and Palfrey (1996) KL criterion measures the divergence between model k and the other models.

In addition, some practical constraints were placed on the question procedure to account for the survey environment and to ensure that the questions asked provided information about all the parameters of interest. In the risk module, the first four questions were restricted to be lotteries over gains in order to focus the procedure on obtaining a precise estimate of ρ before moving onto estimates of λ. To make it harder for respondents to manipulate the procedure, the maximum prize was restricted to be no more than $7 in each even numbered round. Questions were also selected as if monetary amounts were 3 times the actual amounts offered in the lottery to improve discrimination of the risk and loss aversion parameters rather than the consistency parameter μ. In the time module, 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)).

After receiving the survey responses, the individual-level estimates presented in the paper were obtained by performing the Bayesian updating procedure on the answers to the questions selected by the procedure above. This re-estimation allowed us to use a more refined prior, and also to use more of the information obtained during the risk module in estimating individual discount factors. The re-estimation used a discrete uniform prior, consisting of 100 points. The range of the prior was constructed based on previous participant estimates obtained by Sokol-Hessner et al. (2009) and Frydman et al. (2011). We allow for values of ρ between 0.2 and 1.7, for λ between 0 and 4.6, for δ (and β) between 0.2 and 1.0 and for μ between 0 and 8.0. 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. In practice, however, we found very little evidence of present bias either in the DOSE module or in the time MPLs.

As such, the prior used to obtain estimates did not include a present bias parameter.

Implementing the survey through YouGov's online platform precluded using DOSE to choose questions in real time. Instead, simulated responses were used to map out all possible sets of binary choices in advance. That tree was then used to route respondents through the survey. Mapping such a tree with the 100 point discretized prior was infeasible given both computational constraints and the limitations of YouGov's interface (since mapping such a tree over 20 questions would involve over 500,000 routes through the survey).

Given these constraints, question selection was implemented separately for the risk and time modules, and used a coarser prior.[6] To utilize the information about the curvature of the utility function from the risk-loss module, respondents were assigned to one of ten prior distributions over ρ, based on their estimated ρ from the risk-loss module. These ten distributions focus on the mass points of the prior used in the risk preference module.

D Robustness Checks

This section presents extended survey results and robustness tests.

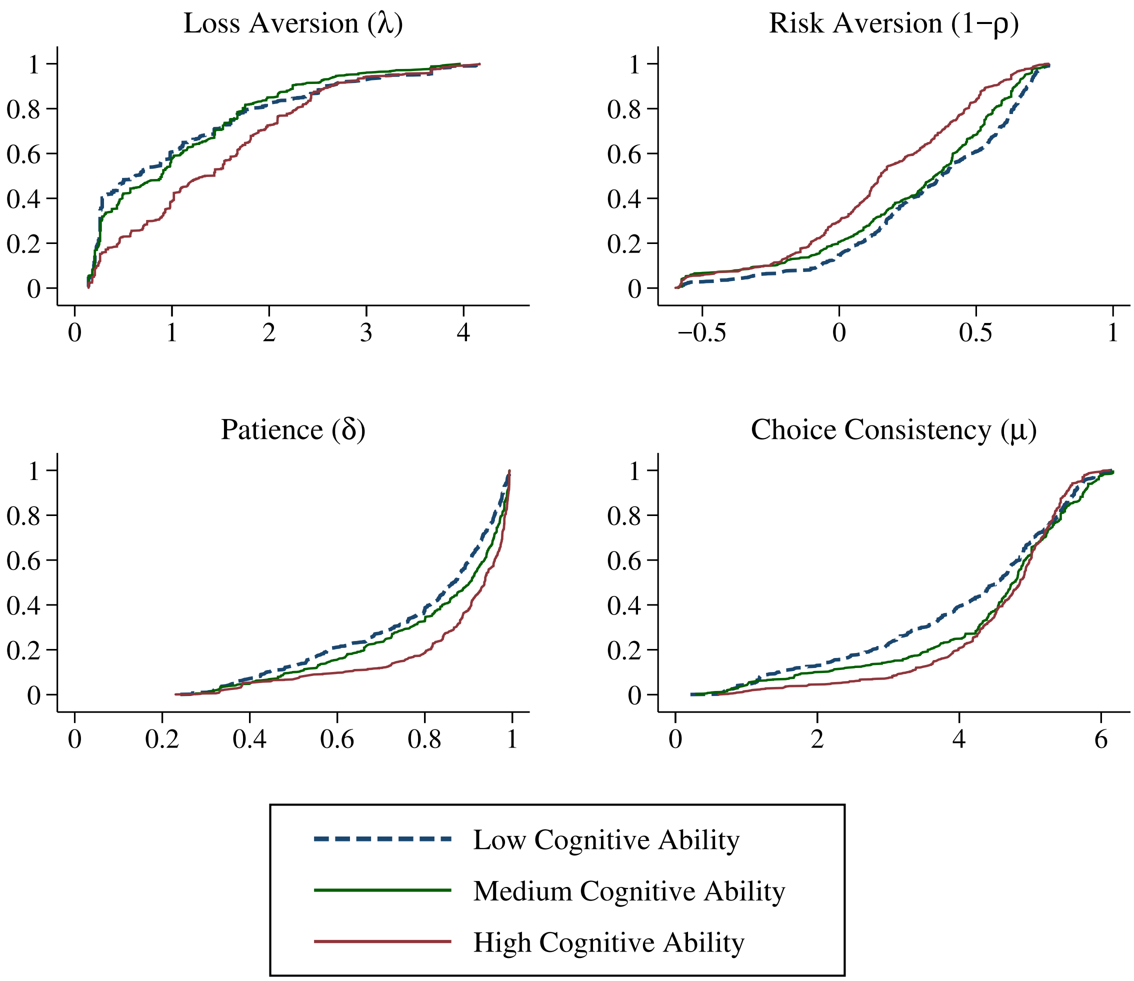
D.1 Robustness of Correlations with Economic Preferences

In this subsection we present extended versions of the correlation tables in Section 4, including a wider range of individual characteristics and comparisons with alternative measures of risk and time preferences.

The distribution of preferences for high cognitive ability participants differs significantly from the rest of the population, as shown in Figure D.1. 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, as discussed in Section 4.2, there is no evidence that the correlations between

[6]The prior included 12 mass points for ρ, 20 for λ and 4 for μ.

Figure D.1: 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.

cognitive ability and preferences are driven by high cognitive ability individuals clustering at values near risk- or loss-neutrality.

In Table D.1 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 the main paper), 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 4.3, 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 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 D.2. 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 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.

The correlations with cognitive ability are robust to the inclusion of the other sociodemographic controls in Table 2—see Table D.3 and Table D.4. For each of the four preference parameters, the first specification includes only the attributes—age and sex—that are not 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 D.3 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 2. In all specifications the coefficient for cognitive ability is still strongly statistically significant and,

Table D.1: 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***
	(.030)	(.027)	(.025)	(.029)	(.036)	(.025)
Income	-0.15***	-0.06*	0.00	-0.13***	0.12***	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	0.02	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***
	(.033)	(.032)	(.032)	(.034)	(.037)	(.038)
Employed	-0.04	-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.

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 D.4. These specifications allow for potential non-monotonic relationships, as well

Table D.2: 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***	0.15***
	(.030)	(.028)	(.029)	(.026)
IQ	0.18***	-0.18***	0.14***	0.12***
	(.033)	(.029)	(.032)	(.028)
CRT	0.19***	-0.18***	0.18***	0.14***
	(.029)	(.030)	(.036)	(.025)
Income	0.15***	-0.15***	0.12***	0.06*
	(.032)	(.034)	(.034)	(.033)
Education	0.13***	-0.10***	0.17***	0.11***
	(.032)	(.033)	(.037)	(.032)
Male	0.08**	-0.10***	-0.02	0.01
	(.033)	(.032)	(.035)	(.033)
Age	-0.10***	0.02	0.18***	0.05
	(.033)	(.032)	(.036)	(.036)
Stock Investor	0.06**	-0.11***	0.10***	-0.02
	(.031)	(.029)	(.031)	(.032)
Non-white	-0.14***	0.10***	-0.18***	-0.05
	(.033)	(.033)	(.037)	(.038)
Own Home	0.06*	-0.07**	0.13***	0.01
	(.033)	(.032)	(.035)	(.033)
Employed	0.06*	-0.04	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.

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.

Table D.3: 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***	-0.19***	-0.17***	0.21***	0.17***	0.16***	0.14***
	(0.031)	(0.034)	(0.028)	(0.033)	(0.028)	(0.031)	(0.028)	(0.029)
Male	0.05	0.07	-0.12**	-0.15**	-0.12*	-0.14*	-0.05	-0.04
	(0.067)	(0.066)	(0.063)	(0.067)	(0.067)	(0.073)	(0.068)	(0.065)
Age	-0.09***	-0.09**	-0.00	0.00	0.19***	0.18***	0.06*	0.03
	(0.032)	(0.037)	(0.031)	(0.036)	(0.035)	(0.041)	(0.035)	(0.040)
Education		0.06		-0.00		0.10**		0.05
		(0.038)		(0.036)		(0.044)		(0.033)
Income		0.08**		-0.09**		0.04		0.05
		(0.036)		(0.036)		(0.037)		(0.038)
Stock Investor		0.05		-0.08		-0.07		-0.20**
		(0.081)		(0.074)		(0.079)		(0.097)
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.

The results in Table D.3 and Table D.4 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 carry out additional specifications adding the variables one at a time—see Table D.5. 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 D.4: 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	-0.02	-0.03	-0.18**	-0.17**	0.18**	0.17**	0.25***	0.24***
	(.082)	(.080)	(.078)	(.077)	(.082)	(.080)	(.086)	(.083)
Top Tercile	0.35***	0.27***	-0.48***	-0.41***	0.48***	0.38***	0.38***	0.35***
	(0.083)	(.085)	(.071)	(.076)	(.080)	(.083)	(.078)	(.076)
Age:								
36–50	-0.23**	-0.23**	0.04	0.06	0.15	0.14	0.08	0.05
	(0.106)	(0.103)	(0.095)	(0.094)	(0.110)	(0.109)	(0.098)	(0.096)
51–64	-0.32***	-0.35***	0.07	0.12	0.34***	0.32***	0.19**	0.20**
	(0.092)	(0.093)	(0.093)	(0.092)	(0.098)	(0.099)	(0.095)	(0.099)
65+	-0.26***	-0.28***	0.02	0.04	0.46***	0.50***	0.11	0.13
	(0.095)	(0.099)	(0.094)	(0.096)	(0.098)	(0.101)	(0.100)	(0.106)
Male	0.08	0.08	-0.13**	-0.12*	-0.11*	-0.10	-0.05	-0.04
	(0.067)	(0.065)	(0.062)	(0.062)	(0.066)	(0.064)	(0.067)	(0.065)
Education:								
Some College		-0.02		-0.08		0.30***		0.03
		(0.077)		(0.075)		(0.085)		(0.076)
4-year College		0.15*		-0.06		0.26***		0.16**
		(0.085)		(0.079)		(0.087)		(0.077)
Income:								
2nd Quartile		-0.03		0.08		-0.12		0.12
		(0.089)		(0.094)		(0.102)		(0.096)
3rd Quartile		0.20**		0.03		-0.09		0.07
		(0.092)		(0.090)		(0.109)		(0.091)
4th Quartile		0.22**		-0.21**		0.14		0.05
		(0.100)		(0.096)		(0.093)		(0.099)
Unreported		0.34***		-0.07		0.00		-0.25**
		(0.119)		(0.113)		(0.108)		(0.121)
Stock Investor		0.06		-0.12*		-0.01		-0.19**
		(0.074)		(0.068)		(0.071)		(0.088)
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	0.04

Note: All dependent variables are standardized. Robust standard errors are displayed in parentheses.

Table D.5: 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***
				(0.080)				(0.077)				(0.080)				(0.083)
Top Tercile				0.27***				-0.41***				0.38***				0.35***
				(0.085)				(0.076)				(0.083)				(0.076)
Education:																
Some College	0.04		0.03	-0.02	-0.16**		-0.14*	-0.08	0.35***		0.35***	0.30***	0.09		0.08	0.03
	(0.078)		(0.078)	(0.077)	(0.076)		(0.075)	(0.075)	(0.082)		(0.083)	(0.085)	(0.080)		(0.077)	(0.076)
4-year College	0.32***		0.23***	0.15*	-0.27***		-0.16**	-0.06	0.40***		0.35***	0.26***	0.22***		0.24***	0.16**
	(0.077)		(0.084)	(0.085)	(0.077)		(0.077)	(0.079)	(0.080)		(0.085)	(0.087)	(0.074)		(0.079)	(0.077)
Income:																
2nd Quartile		-0.00	-0.01	-0.03		0.04	0.06	0.08		-0.07	-0.10	-0.12		0.16	0.14	0.12
		(0.087)	(0.087)	(0.089)		(0.097)	(0.096)	(0.094)		(0.108)	(0.103)	(0.102)		(0.097)	(0.096)	(0.096)
3rd Quartile		0.26***	0.22**	0.20**		-0.02	0.01	0.03		0.00	-0.07	-0.09		0.13	0.08	0.07
		(0.094)	(0.093)	(0.092)		(0.093)	(0.092)	(0.090)		(0.112)	(0.109)	(0.109)		(0.096)	(0.094)	(0.091)
Top Quartile		0.33***	0.26**	0.22**		-0.29***	-0.25**	-0.21**		0.27***	0.17*	0.14		0.15	0.08	0.05
		(0.099)	(0.101)	(0.100)		(0.097)	(0.097)	(0.096)		(0.096)	(0.096)	(0.093)		(0.101)	(0.100)	(0.099)
Unreported		0.38***	0.36***	0.34***		-0.09	-0.08	-0.07		0.01	0.01	0.00		-0.23*	-0.24*	-0.25**
		(0.122)	(0.121)	(0.119)		(0.113)	(0.113)	(0.113)		(0.111)	(0.110)	(0.108)		(0.128)	(0.126)	(0.121)
Stock Investor		0.11	0.07	0.06		-0.17**	-0.15**	-0.12*		0.06	0.01	-0.01		-0.13	-0.17*	-0.19**
		(0.073)	(0.074)	(0.074)		(0.067)	(0.068)	(0.068)		(0.071)	(0.072)	(0.071)		(0.085)	(0.088)	(0.088)
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	0.04	0.06	0.06	0.04	0.07	0.08	0.01	0.02	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.

Table D.6: Correlations with individual characteristics are similar when removing fastest 50% of respondents on entire survey.

	Loss Aversion (λ)	Risk Aversion ($1-\rho$)	Patience (δ)	Choice Consistency (μ)
Cognitive Ability	0.22***	-0.21***	0.16***	0.17***
	(0.039)	(0.039)	(0.043)	(0.035)
IQ	0.20***	-0.18***	0.09*	0.15***
	(0.039)	(0.040)	(0.050)	(0.036)
CRT	0.18***	-0.18***	0.23***	0.14***
	(0.045)	(0.044)	(0.040)	(0.035)
Income	0.11***	-0.14***	0.10**	0.11***
	(0.045)	(0.047)	(0.045)	(0.039)
Education	0.07	-0.14***	0.15***	0.10***
	(0.042)	(0.044)	(0.043)	(0.039)
Male	0.08*	-0.12***	-0.04	0.02
	(0.043)	(0.045)	(0.051)	(0.043)
Age	-0.08*	0.04	0.18***	-0.01
	(0.047)	(0.047)	(0.057)	(0.045)
Stock Investor	0.13***	-0.13***	0.08*	-0.02
	(0.043)	(0.042)	(0.046)	(0.042)
Non-white	-0.16***	0.09*	-0.18***	-0.00
	(0.045)	(0.048)	(0.055)	(0.047)
Own Home	0.07*	-0.08*	0.11**	0.05
	(0.044)	(0.044)	(0.052)	(0.045)
Employed	0.08**	-0.04	-0.01	0.07*
	(0.043)	(0.043)	(0.051)	(0.039)
Church Attendance	0.00	0.08*	-0.01	0.06
	(0.044)	(0.045)	(0.050)	(0.045)
Marital Status	0.02	-0.00	-0.14***	-0.05
	(0.048)	(0.046)	(0.059)	(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.

The next four tables show that the correlations between DOSE and the other characteristics are also robust to removing the fastest 50% of participants on the survey (Tables D.6 and D.7) or on the DOSE module (Tables D.8 and D.9).

Table D.7: Correlations with individual characteristics are similar when removing fastest 50% of respondents 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***	0.17***	0.14***
	(0.039)	(0.047)	(0.041)	(0.048)	(0.042)	(0.046)	(0.037)	(0.041)
Male	0.07	0.11	-0.16*	-0.22**	-0.15	-0.20*	-0.03	-0.10
	(0.085)	(0.093)	(0.087)	(0.092)	(0.097)	(0.105)	(0.087)	(0.089)
Age	-0.05	-0.09*	0.01	0.02	0.20***	0.20***	0.01	-0.05
	(0.045)	(0.055)	(0.047)	(0.052)	(0.057)	(0.066)	(0.045)	(0.048)
Education		-0.02		-0.03		0.11**		0.06
		(0.046)		(0.052)		(0.051)		(0.042)
Income		0.03		-0.06		0.03		0.09**
		(0.051)		(0.050)		(0.051)		(0.043)
Stock Investor		0.21*		-0.12		-0.11		-0.21**
		(0.109)		(0.096)		(0.107)		(0.101)
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.

The relationship between expected value maximizing choices and cognitive ability discussed in Section 4.2 is robust to controlling for other socio-demographic characteristics, as showin in Table D.10. 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 5. 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

Table D.8: Correlations with individual characteristics are similar when removing fastest 50% of respondents on DOSE module.

	Loss Aversion (λ)	Risk Aversion ($1-\rho$)	Patience (δ)	Choice Consistency (μ)
Cognitive Ability	0.26***	-0.21***	0.18***	0.15***
	(0.037)	(0.039)	(0.041)	(0.036)
IQ	0.25***	-0.20***	0.12***	0.12***
	(0.037)	(0.040)	(0.049)	(0.036)
CRT	0.19***	-0.16***	0.22***	0.14***
	(0.044)	(0.045)	(0.040)	(0.036)
Income	0.11***	-0.12***	0.10**	0.08*
	(0.044)	(0.046)	(0.046)	(0.041)
Education	0.13***	-0.14***	0.16***	0.09*
	(0.041)	(0.039)	(0.045)	(0.047)
Male	0.08*	-0.11***	-0.04	-0.01
	(0.042)	(0.044)	(0.050)	(0.047)
Age	-0.05	0.07	0.19***	0.03
	(0.047)	(0.046)	(0.056)	(0.057)
Stock Investor	0.12***	-0.09***	0.15***	-0.03
	(0.039)	(0.037)	(0.037)	(0.040)
Non-white	-0.18***	0.12***	-0.19***	0.00
	(0.044)	(0.047)	(0.056)	(0.053)
Own Home	0.08*	-0.11***	0.13***	0.00
	(0.043)	(0.043)	(0.051)	(0.048)
Employed	0.12***	-0.03	0.02	0.04
	(0.042)	(0.041)	(0.049)	(0.043)
Church Attendance	-0.03	0.05	0.03	-0.03
	(0.043)	(0.045)	(0.049)	(0.046)
Marital Status	-0.02	0.00	-0.18***	-0.06
	(0.047)	(0.046)	(0.056)	(0.049)

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.

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 D.9: Correlations with individual characteristics are similar when removing fastest 50% of respondents 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***	-0.19***	-0.16***	0.23***	0.19***	0.16***	0.11***
	(0.038)	(0.043)	(0.039)	(0.049)	(0.040)	(0.043)	(0.043)	(0.035)
Male	0.05	0.14	-0.13	-0.14	-0.16*	-0.20**	-0.07	-0.10
	(0.083)	(0.091)	(0.084)	(0.092)	(0.091)	(0.102)	(0.098)	(0.089)
Age	-0.01	-0.06	0.04	0.01	0.22***	0.18***	0.05	-0.05
	(0.045)	(0.054)	(0.046)	(0.053)	(0.055)	(0.065)	(0.058)	(0.052)
Education		0.05		-0.08*		0.09*		0.04
		(0.045)		(0.048)		(0.053)		(0.042)
Income		0.01		-0.03		0.01		0.07*
		(0.048)		(0.050)		(0.049)		(0.039)
Stock Investor		0.12		-0.02		0.03		-0.18**
		(0.097)		(0.080)		(0.083)		(0.087)
Obs.	1012	875	1012	875	1012	875	1012	875
Adj. R^2	0.07	0.08	0.05	0.05	0.09	0.08	0.02	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.

D.2 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 D.2 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.

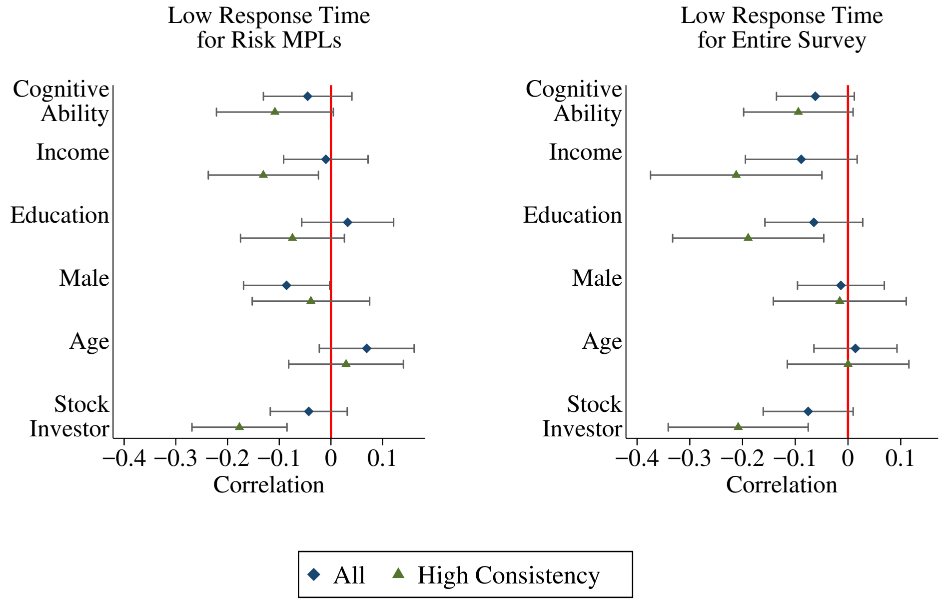
In both panels there is more evidence of correlations after restricting the sample to high consistency participants. The magnitude of the correlations in frequently higher, and several emerge as statistically significant once only high consistency participants are considered. The magnitude of the correlations is, in fact, similar to those in Figure 6, 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 D.10: 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**
	(0.023)	(0.023)	(0.020)	(0.020)	(0.032)	(0.032)	(0.031)	(0.031)
x High Cognitive Ability	0.171***	0.168***	0.114***	0.104***	-0.044	-0.063**	-0.070**	-0.066**
	(0.025)	(0.025)	(0.023)	(0.023)	(0.031)	(0.031)	(0.030)	(0.031)
EV>Sure Amount	0.264***	0.269***	0.152***	0.159***	-0.289***	-0.292***	-0.244***	-0.267***
	(0.023)	(0.022)	(0.023)	(0.026)	(0.025)	(0.025)	(0.029)	(0.033)
x Medium Cognitive Ability	0.013	0.013	-0.010	-0.009	0.040**	0.036**	0.027	0.026
	(0.020)	(0.020)	(0.019)	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)
x High Cognitive Ability	-0.088***	-0.087***	-0.107***	-0.101***	0.120***	0.104***	0.087***	0.084***
	(0.018)	(0.019)	(0.018)	(0.018)	(0.017)	(0.018)	(0.018)	(0.018)
Some College		0.007	0.010	-0.001		-0.000	0.001	-0.042
		(0.014)	(0.012)	(0.020)		(0.016)	(0.015)	(0.030)
x EV>Sure Amount				0.018				0.066*
				(0.029)				(0.035)
4-year College		0.015	0.022	0.066***		0.020	0.026	0.017
		(0.015)	(0.014)	(0.024)		(0.017)	(0.016)	(0.032)
x EV≤Sure Amount				-0.069**				0.014
				(0.031)				(0.037)
Age (Standardized)		0.035***	-0.002	-0.002		0.003	-0.012*	-0.012*
		(0.006)	(0.006)	(0.006)		(0.007)	(0.007)	(0.007)
Male		0.012	0.010	0.010		0.004	0.003	0.004
		(0.011)	(0.010)	(0.010)		(0.013)	(0.013)	(0.013)
Income: 2nd Quartile		0.032**	0.031**	0.031**		0.058***	0.057***	0.057***
		(0.016)	(0.015)	(0.015)		(0.019)	(0.018)	(0.018)
Income: 3rd Quartile		-0.003	0.000	-0.001		0.033*	0.035*	0.036*
		(0.016)	(0.015)	(0.015)		(0.019)	(0.019)	(0.019)
Income: 4th Quartile		-0.011	-0.005	-0.005		0.067***	0.068***	0.067***
		(0.017)	(0.016)	(0.016)		(0.021)	(0.020)	(0.020)
Income: Unstated		-0.006	-0.010	-0.010		0.022	0.020	0.020
		(0.020)	(0.018)	(0.018)		(0.022)	(0.021)	(0.021)
Lottery Prize ($)			0.031***	0.031***			-0.004	-0.004
			(0.002)	(0.002)			(0.004)	(0.004)
|EV - Sure Amount| ($)			-0.021***	-0.020***			-0.023***	-0.023***
			(0.004)	(0.004)			(0.007)	(0.007)
Response Time: Quartile 2			0.144***	0.143***			0.061***	0.062***
			(0.013)	(0.013)			(0.020)	(0.020)
Response Time: Quartile 3			0.220***	0.221***			0.082***	0.084***
			(0.014)	(0.014)			(0.019)	(0.019)
Response Time: Quartile 4			0.216***	0.217***			0.130***	0.131***
			(0.015)	(0.015)			(0.020)	(0.020)
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.

Figure D.2: Accounting for choice consistency leads to a clearer pattern of correlations even after removing very fast responses.

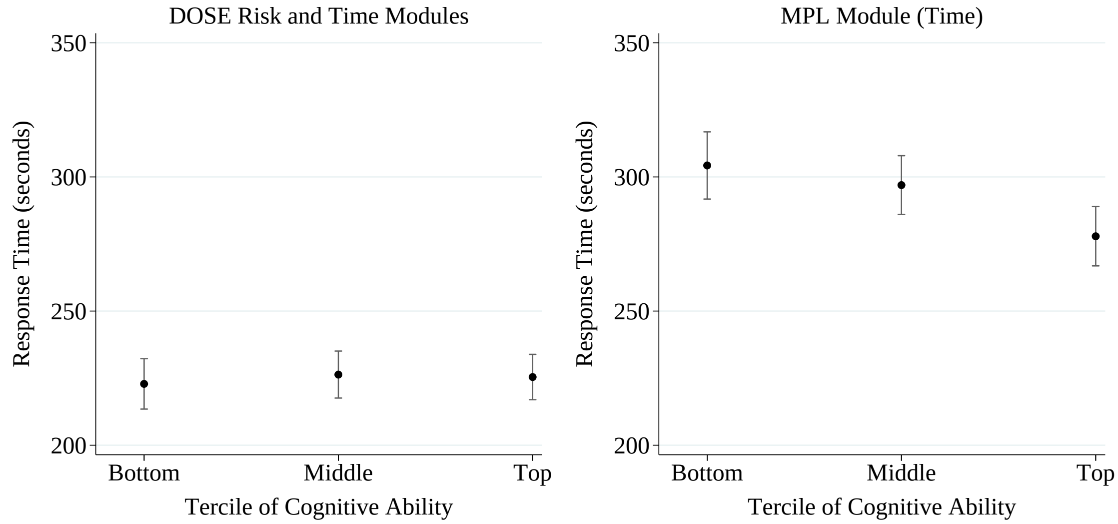


Notes: The left panel includes only participants below the median response time on the risk MPL module. The right panel includes only participants below the median response time on the entire survey. "High Consistency" refers to those with choice consistency above the median. The survey contained two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

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

Finally, note that the relationship between cognitive ability and response times for both MPL and DOSE is similar when analyzing by cognitive ability tercile (see Figure D.3), as when we do so by quartile (see Figure 8).

Figure D.3: Participants in bottom cognitive ability tercile take longer for MPL questions, but not DOSE.



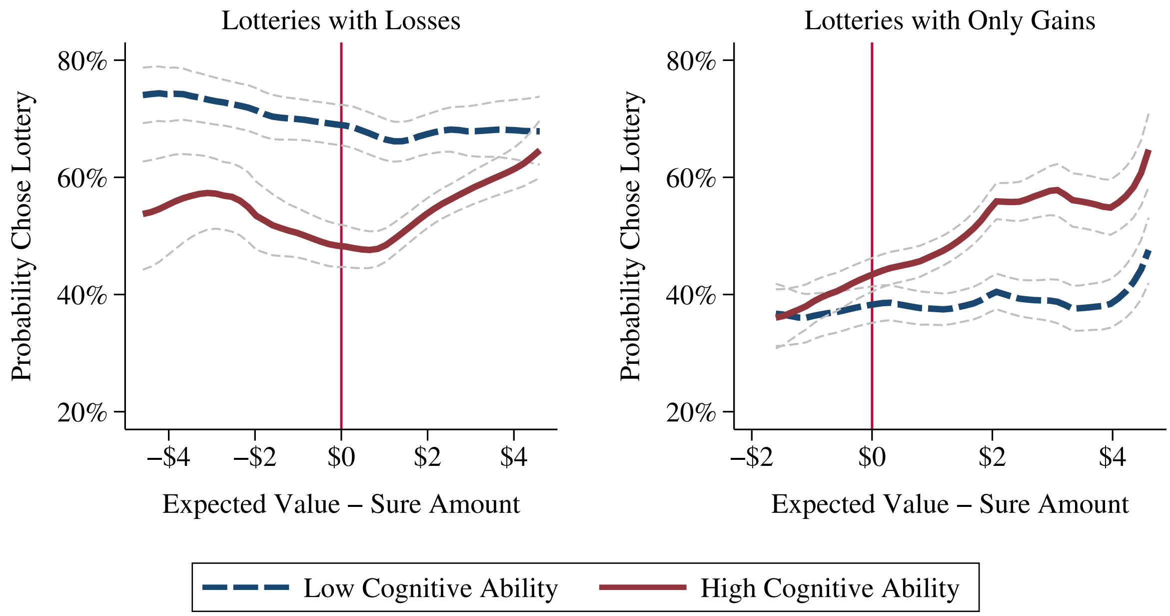
Notes: DOSE module includes 20 questions addressing both risk and time preferences. MPL module includes two MPLs assessing time preferences, which was the first MPL module on the survey. Respondents with response times of over 15 minutes for either module are excluded.

D.3 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 D.11 and D.12, 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 D.4 and Table D.13.

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 D.5)—a difference that is probably explained by differential attrition. The first and third panels of the figure are very similar to those in Figure 6: 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

Figure D.4: Pattern of individual choices is the 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.

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.

D.4 Classification of Participants by DOSE

The DOSE estimates clearly capture participants' choices, providing evidence that our results are not an artefact of functional form—see Table D.14. Here participants are classified 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

Table D.11: 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***	-0.24***	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***	0.16***
	(0.035)	(0.033)	(0.032)	(0.034)
Income	0.12***	-0.19***	0.20***	0.15***
	(0.039)	(0.041)	(0.048)	(0.040)
Education	0.14***	-0.15***	0.21***	0.17***
	(0.034)	(0.035)	(0.040)	(0.036)
Male	0.06	-0.04	0.01	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***	0.16***	0.08**
	(0.035)	(0.038)	(0.038)	(0.037)
Non-white	-0.13***	0.11***	-0.22***	-0.16***
	(0.044)	(0.041)	(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	0.01
	(0.040)	(0.040)	(0.039)	(0.039)
Marital Status	0.09**	-0.03	-0.06	-0.12***
	(0.042)	(0.042)	(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.

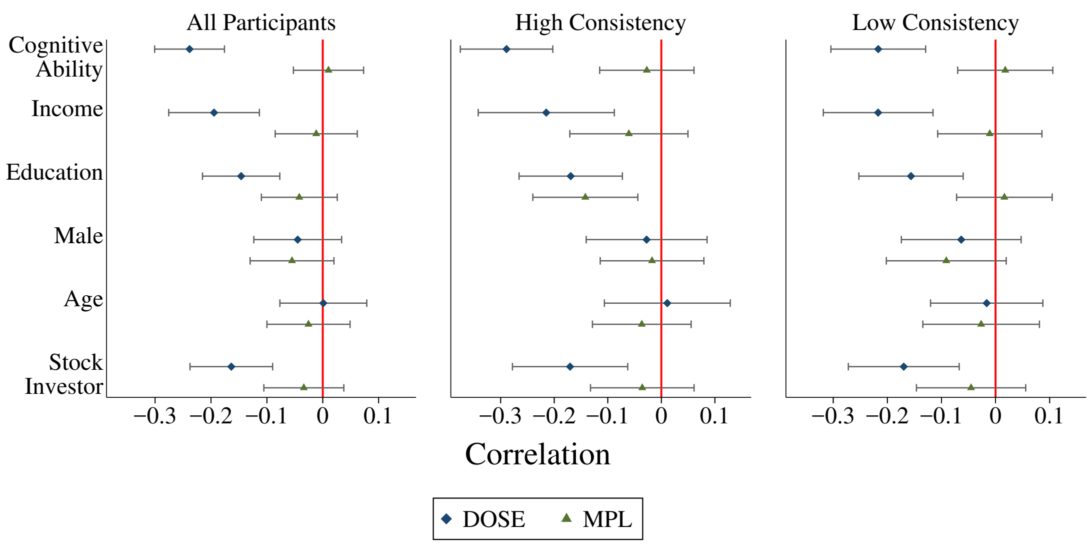
lotteries accepted varies according to the expected value (relative to a sure amount) and whether the lottery involved a loss. The pattern of behavior is as would be expected. Loss tolerant participants nearly always choose lotteries with losses, and risk loving participants

Table D.12: 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***	-0.24***	-0.18***	0.26***	0.19***	0.19***	0.15***
	(0.035)	(0.042)	(0.033)	(0.035)	(0.034)	(0.035)	(0.037)	(0.037)
Male	0.04	0.10	-0.02	-0.04	-0.04	-0.02	-0.04	-0.11
	(0.079)	(0.082)	(0.078)	(0.077)	(0.083)	(0.084)	(0.085)	(0.089)
Age	-0.09**	-0.05	-0.02	0.02	0.11**	0.04	0.09**	0.09**
	(0.040)	(0.042)	(0.038)	(0.042)	(0.046)	(0.052)	(0.042)	(0.045)
Education		0.08*		-0.03		0.08*		0.09**
		(0.040)		(0.043)		(0.043)		(0.039)
Income		0.04		-0.09**		0.08*		0.07*
		(0.041)		(0.043)		(0.049)		(0.043)
Stock Investor		-0.06		-0.21**		0.11		-0.05
		(0.080)		(0.096)		(0.098)		(0.092)
Obs.	1465	1271	1465	1271	1465	1271	1465	1271
Adj. R^2	0.07	0.08	0.06	0.08	0.07	0.09	0.04	0.06

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.

Figure D.5: 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.

nearly always choose lotteries over gains. Loss averse and risk averse participants, in contrast, are much less likely to accept such lotteries.

Table D.13: 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
	(0.026)	(0.026)	(0.024)	(0.024)	(0.038)	(0.038)	(0.037)	(0.037)
x High Cognitive Ability	0.172***	0.171***	0.127***	0.120***	0.018	0.008	0.002	0.011
	(0.031)	(0.032)	(0.029)	(0.030)	(0.040)	(0.041)	(0.040)	(0.041)
EV>Sure Amount	0.260***	0.261***	0.151***	0.156***	-0.272***	-0.273***	-0.230***	-0.246***
	(0.027)	(0.027)	(0.027)	(0.031)	(0.032)	(0.032)	(0.037)	(0.042)
x Medium Cognitive Ability	-0.035	-0.032	-0.037*	-0.036	0.036*	0.032	0.031	0.029
	(0.023)	(0.023)	(0.022)	(0.022)	(0.020)	(0.020)	(0.019)	(0.019)
x High Cognitive Ability	-0.115***	-0.109***	-0.124***	-0.120***	0.138***	0.127***	0.115***	0.110***
	(0.023)	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.022)	(0.022)
Some College		0.023	0.020	0.015		0.025	0.022	0.005
		(0.017)	(0.016)	(0.025)		(0.019)	(0.018)	(0.037)
x EV>Sure Amount				0.008				0.025
				(0.036)				(0.042)
4-year College		-0.018	-0.016	0.011		0.013	0.012	-0.018
		(0.018)	(0.017)	(0.028)		(0.021)	(0.020)	(0.040)
x EV≤Sure Amount				-0.040				0.045
				(0.037)				(0.044)
Age (Standardized)		0.034***	-0.005	-0.005		0.003	-0.008	-0.008
		(0.007)	(0.007)	(0.007)		(0.008)	(0.009)	(0.009)
Male		0.004	-0.000	-0.001		-0.003	-0.004	-0.004
		(0.013)	(0.012)	(0.012)		(0.016)	(0.015)	(0.015)
Income: 2nd Quartile		0.009	0.008	0.009		-0.013	-0.013	-0.013
		(0.020)	(0.019)	(0.019)		(0.023)	(0.022)	(0.022)
Income: 3rd Quartile		0.003	0.008	0.008		0.002	0.001	0.001
		(0.020)	(0.018)	(0.018)		(0.022)	(0.022)	(0.022)
Income: 4th Quartile		0.045**	0.048***	0.048**		0.045*	0.043*	0.043*
		(0.020)	(0.019)	(0.019)		(0.025)	(0.024)	(0.024)
Income: Unstated		0.009	0.003	0.003		-0.003	-0.004	-0.003
		(0.023)	(0.022)	(0.022)		(0.025)	(0.025)	(0.025)
Lottery Prize ($)			0.030***	0.030***			-0.002	-0.002
			(0.003)	(0.003)			(0.005)	(0.005)
|EV - Sure Amount| ($)			-0.018***	-0.018***			-0.026***	-0.026***
			(0.004)	(0.004)			(0.008)	(0.008)
Response Time: Quartile 2			0.136***	0.136***			0.049**	0.048**
			(0.015)	(0.015)			(0.023)	(0.023)
Response Time: Quartile 3			0.207***	0.208***			0.095***	0.096***
			(0.016)	(0.016)			(0.023)	(0.023)
Response Time: Quartile 4			0.198***	0.199***			0.082***	0.082***
			(0.018)	(0.018)			(0.025)	(0.025)
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.

Table D.14: 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	12%	48%	4%	42%
Loss Averse, Risk Loving	2%	51%	53%	96%
Loss Tolerant, Risk Averse	80%	98%	4%	43%
Loss Tolerant, Risk Loving	81%	100%	75%	97%

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.

E Parameter Recovery Procedure

This section provides full details of our simulation procedure, and shows that the results in Section 3.2 may underestimate the benefits of DOSE relative to other elicitation methods. The first subsection provides a detailed explanation of the procedure used to simulate DOSE, 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.

E.1 Simulation 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.

Table E.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.

DOSE Simulation We simulate a 20 question DOSE procedure for each individual, with each binary choice made probabilistically according to the logit probability (2). 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 E.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 E.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.

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 (2), 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 expected inaccuracy is given by:

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

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.

Table E.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 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. Neither value is defined in the first row because the design does not allow the right hand side to be selected.

The estimated inaccuracy of the Lottery Menu procedure is significantly higher under this alternative procedure. Drawing the consistency parameter at random (as in Table 1), the average inaccuracy for the risk aversion parameter is 94%, compared to 35% under the previous procedure. Further, with this alternative procedure the Lottery Menu estimates are inaccurate for the very inconsistent participants too—at the lowest consistency ventile, the average inaccuracy is 139% (for the highest ventile, it is 59%).

Double Multiple Price List (MPL)

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

Table E.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
$\vdots$	$\vdots$
$\vdots$	$\vdots$
50% of -\$10, 50% of \$10	\$9
50% of -\$10, 50% of \$10	\$10

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 E.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 E.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 (2)) 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:

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

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.

E.2 Comparison of MPL Simulation to Survey Data

Our simulations appear to underestimate the amount of error in the MPL. To understand how this pattern compares to actual behavior we compare estimates from a previous survey to the pattern of behavior generated by simulating choices using the procedure above. By doing so we can identify whether our simulations are providing a reasonable approximation to the level of measurement error in the survey. In particular, we use the results of three MPLs collected in a separate representative U.S. survey (Chapman et al., 2018). The first two of these MPLs offered 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 offered participants a choice between fixed amounts and a 50:50 lottery between a loss of \$5 and a gain of \$5.

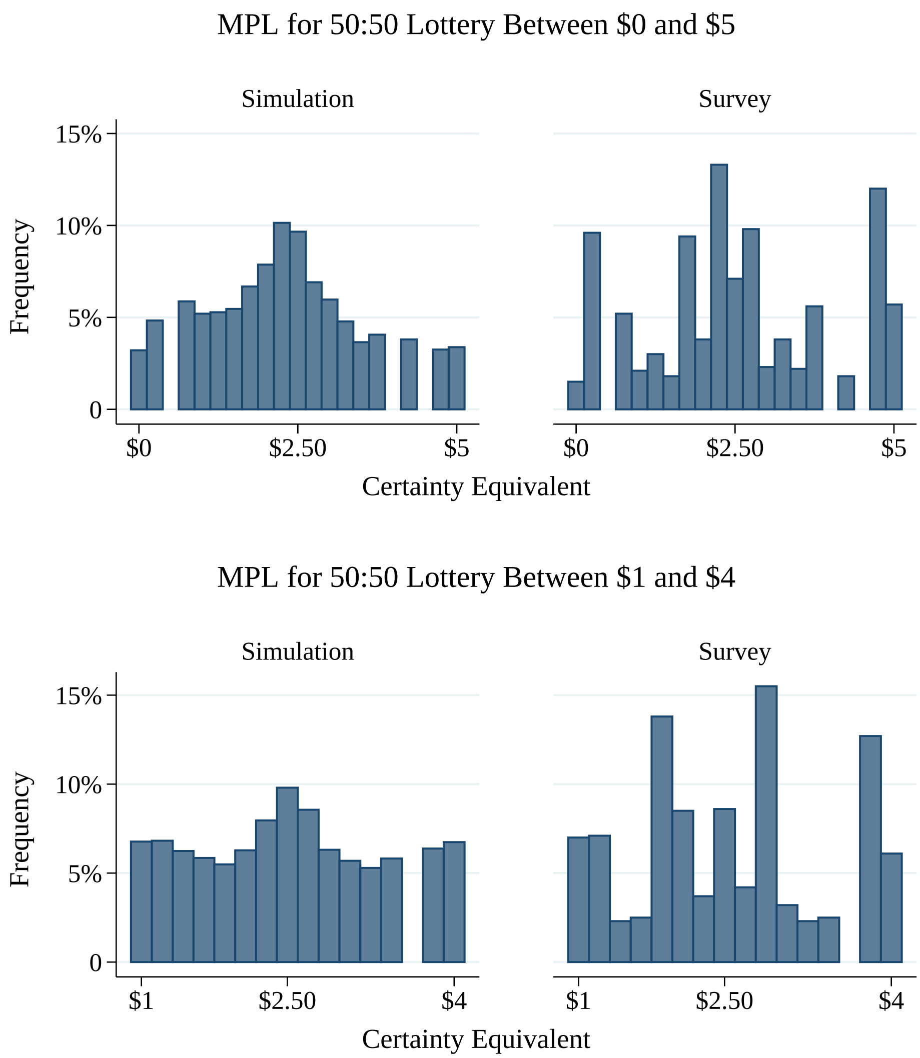
Participants in the survey were much more likely to be attracted to particularly salient rows of the MPL than our simulation would suggest, as shown in Figure E.1. A high proportion of survey participants switched in rows near the ends of the MPL (despite the fact that these values correspond to extreme parameter values).[7] This pattern suggests that the framing of the MPL affects choices and, as we did not account for framing effects in the simulation, that the simulations may miss an important source of error in the MPL.

In fact, the simulations suggest that the level of measurement error in our simulation was lower than that in the survey. In the survey, the correlation was 0.69; in the simulation it is slightly higher (0.73). Further, a significant degree of the correlation in the survey data is explained by participants repeatedly switching at the extremes at the end of MPL: 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.

Further evidence that the survey contained more measurement error is that there are also considerably fewer first order stochastically dominated choices in the simulated loss aversion

[7] 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.

Figure E.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).

MPL than in observed in reality. In our simulated data, 20% of participants have missing data. In the survey data, in comparison, we were unable to elicit loss aversion parameters for 37% of participants. Again, it appears the simulation procedure underestimates the degree of noise in the MPL method in practice.

F Robustness to Misspecification

In this section we provide detailed evidence supporting the analysis in Section 5.1. Additional simulations show that misspecifying the utility function used does not reduce the accuracy of the DOSE parameter estimates. 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. Finally, re-estimating our survey data with alternative utility functions supports the use of the utility function in (1): on average curvature over gains and losses is very similar, and the function fits the choice data better than alternative specifications.

F.1 Additional Simulation Results

To illustrate both the flexibility and robustness of DOSE, we present the results of two additional parameter recovery exercises. The first shows that misspecifying the utility function used in the question selection procedure does not lead to inaccurate parameter estimates. The second demonstrates that DOSE is able to capture meaningful information about preferences when extending the utility function to allow for risk aversion to vary across gains and losses.

F.1.1 Misspecification of the Utility Function

To test the robustness of the DOSE estimates to misspecification, we run DOSE on the same set of simulated 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—even though the question selection procedure is misspecified—the data collected can be re-estimated to elicit accurate CRRA utility parameters.

Specifically we run DOSE assuming the following exponential (CARA) utility function, as suggested by Köbberling and Wakker (2005):

$$u(x, \gamma_i, \lambda_i) = \begin{cases} \frac{1-e^{-\gamma_i x}}{\gamma_i} & \text{for } x \geq 0 \\ \lambda_i \left(\frac{e^{\gamma_i x}-1}{\gamma_i} \right) & \text{for } x < 0 \end{cases} \tag{5}$$

where λ represents loss aversion and γ captures risk aversion.

As with the main 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).[8] The joint posterior from that procedure is then used to draw simulated participants, and a 20 question DOSE procedure is simulated for each participant.

Misspecifying the utility function does not lead to a loss of accuracy, as shown in Table F.1. For loss aversion, very similar estimates are obtained even when the CARA function is incorrectly used (see the bottom panel of the table). For risk aversion, we can recover the same estimates by using the correct utility function after the data collection process. Further, 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 earlier in the paper. As such, 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.

F.1.2 Allowing for Differential Risk Aversion over Gains and Losses

In this subsection we show that DOSE can obtain reasonably accurate estimates for a utility function with differential utility curvature between gains and losses, even if no questions solely involving losses are asked.

In particular, we simulate the DOSE procedure using the same procedure as outlined in

[8] As in the main text, we implement a discretized uniform prior. For λ and μ we use the same parameter range as in the main estimation procedure. For γ we construct the range of the prior based on calculating the Coefficient of Absolute Risk Aversion for the prior range for ρ for a prize of \$1.

Table F.1: 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.

Section 3.2, but assuming a utility function with different power exponents in the gain and loss domain, as suggested by Prospect Theory (Kahneman and Tversky, 1979):

$$u(x, \rho_i^+, \rho_i^-, \lambda_i) = \begin{cases} u(x) = x^{\rho_i^+} & \text{for } x \geq 0 \\ u(x) = -\lambda_i(-x)^{\rho_i^-} & \text{for } x < 0 \end{cases} \tag{6}$$

As with the main 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).[9] The joint posterior from that procedure is then used to draw simulated participants, and a 20 question DOSE procedure is simulated for each participant.

DOSE extracts meaningful information about all four parameters, although with less

[9] As in the main text, we implement a discretized uniform prior, using the same prior range for both risk aversion parameters as we use for ρ.

Table F.2: DOSE estimates of the 4 parameter model are less accurate.

	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.

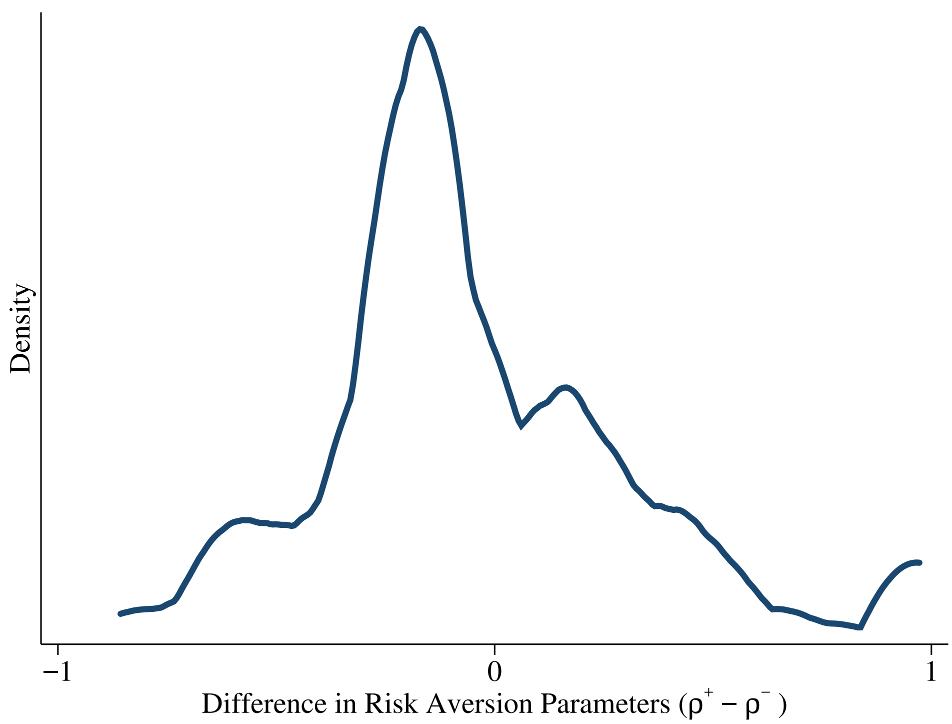
precision than in the three parameter model, as shown in Table F.2. For risk aversion over gains the accuracy of the estimates are similar to those for the 3 parameter model. For the two parameters regarding choices over losses, however, the estimates are noisier. The inaccuracy and correlation with 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%.

F.2 Risk Aversion Over Gains and Losses and Model Fit

The re-estimated survey data provides some support for our assumption of a CRRA utility with the same curvature over gains and losses in (1). On average the curvature of the risk aversion parameter is similar across the two domains. Further, our model correctly predicts more actual choices than the alternative utility functions, or assuming that participants incorporated their $10 endowment into their utility function.

Re-estimating the survey data as per (6) indicates that overall the power utility coefficients are similar in the gain and loss domains, as shown in Figure F.1. In particular the mean difference in the two parameters ($\rho^+ - \rho^- < 1$) is -0.04, and the median difference is -0.11. 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.

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 (6).

The CRRA model in (1) fits our data better than either of the two alternative utility functions, providing further evidence that the specification is appropriate. In particular, our main parameter estimates predict 88% of participants' choices correctly. The CARA model (see (5)) and the CRRA allowing for differential curvature (see (6)), in contrast, both predict less than 85% correctly. Although we should not read too much into these differences—given that they are relatively small and that the questions were selected under the assumption of a CRRA utility function—they provide some reassurance that our model is not significantly misspecified.

A similar model fitting exercise shows that participants were not incorporating the $10 initial endowment (that applied only if a choice in the DOSE module was selected for payoff) in their calculations. If participants did so, every payoff—even those that were negative (with respect to the endowments)—would appear to be a gain with respect to the amount they began the survey with (that is, zero). Thus, the only difference between questions with gains and losses (relative to the endowment) would be the size of the prizes, with those featuring losses being slightly lower, but positive (with respect to zero). To test whether such behavior could explain the choices we observe, we re-estimated our model adding $10 to each payoff, and found that this produced a much worse fit. In particular, the re-estimated parameters predict only 48% of the choices made by participants—a large decrease from the 88% predicted using our main parameter estimates. For loss tolerant participants the performance is no better: the re-estimated parameters explain 49% of choices, the main estimates explain 91%. Thus, not incorporating the initial endowment and using the loss aversion parameter significantly improves the explanatory power of the parametric model.

G Additional Analysis of Lab Data Simulations

This Appendix contains additional results from the DOSE simulations using previous laboratory data. First we plot the correlations between the DOSE estimates and the final (post-140 question) parameter estimates, and then present scatter plots displaying the evolution of the individual-level estimates as more questions are asked. The second subsection then includes the results from attempts to use Maximum Likelihood Estimation to obtain individual level estimates using the same data.

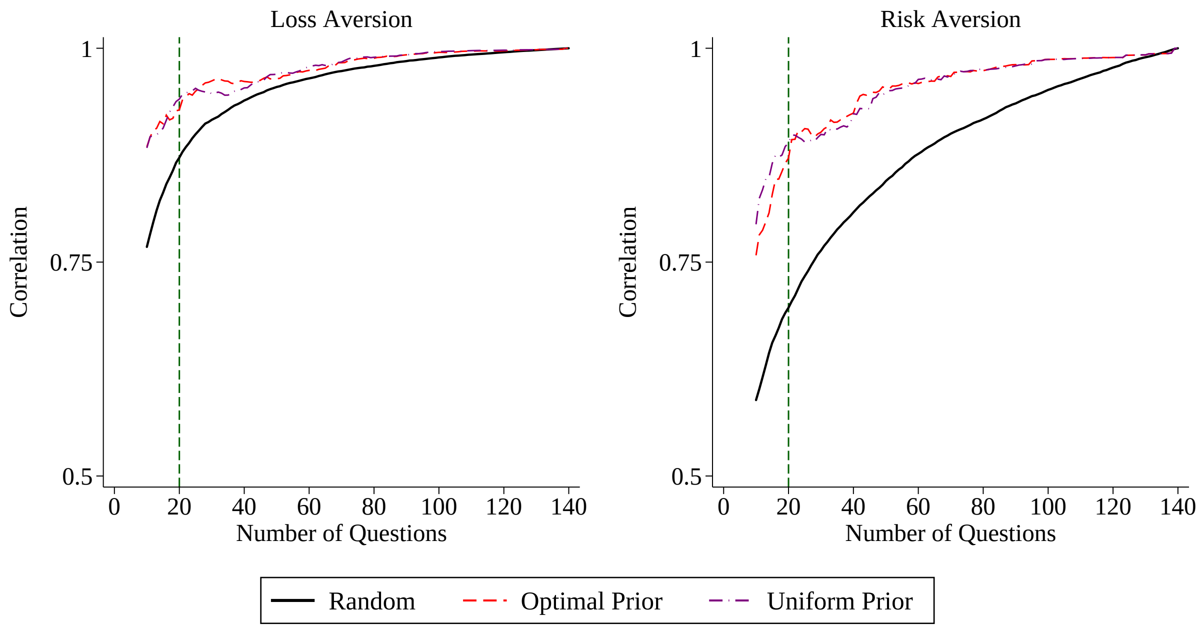
G.1 Additional Results from Bayesian Analysis

The DOSE estimates are highly correlated with the final parameter values after just a few questions, providing further evidence that the procedure elicits considerable information about preferences very quickly. As shown in Figure G.2, the correlation with the final estimate is above 0.8 for both parameters after a 20 question DOSE sequence. This correlation is much higher than obtained under the random ordering—particularly in the case of the loss aversion parameter. Further, a comparison with the correlations obtained using the "optimal prior" demonstrates again the effectiveness of using a uniform prior for question selection.

DOSE provides relatively accurate estimates of the choice consistency parameter as well as risk and loss aversion, as shown in Figure G.2. Compared to the random ordering, the DOSE estimates are closer to and more highly correlated with the post-140 question 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" (see discussion in Section 3.1). It is notable, however, that the estimated inaccuracy is significantly higher after 20 questions than for either risk or loss aversion. This difference is likely to reflect the fact that inconsistency could not be accurately identified until the procedure had asked several similar questions.

DOSE elicits information rapidly for all sets of parameter values in our simulation, as shown in Figures G.3, G.4 and G.5. These figures demonstrate the progression of the estimated value of each of the three parameters towards the final estimate after 10, 20, 50 and 100 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

Figure G.1: Optimal question selection elicits estimates highly correlated with final values.



Notes: Based on data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). Left hand panel shows Right hand panel shows the correlation between the Bayesian estimates (with uniform initial prior) obtained after each question and the final estimate, starting at question 10, under different question orders. "Optimal prior" and "Uniform prior" refer DOSE question selection using corresponding priors. "Random" orders questions randomly, averaging over 100 different random orderings.

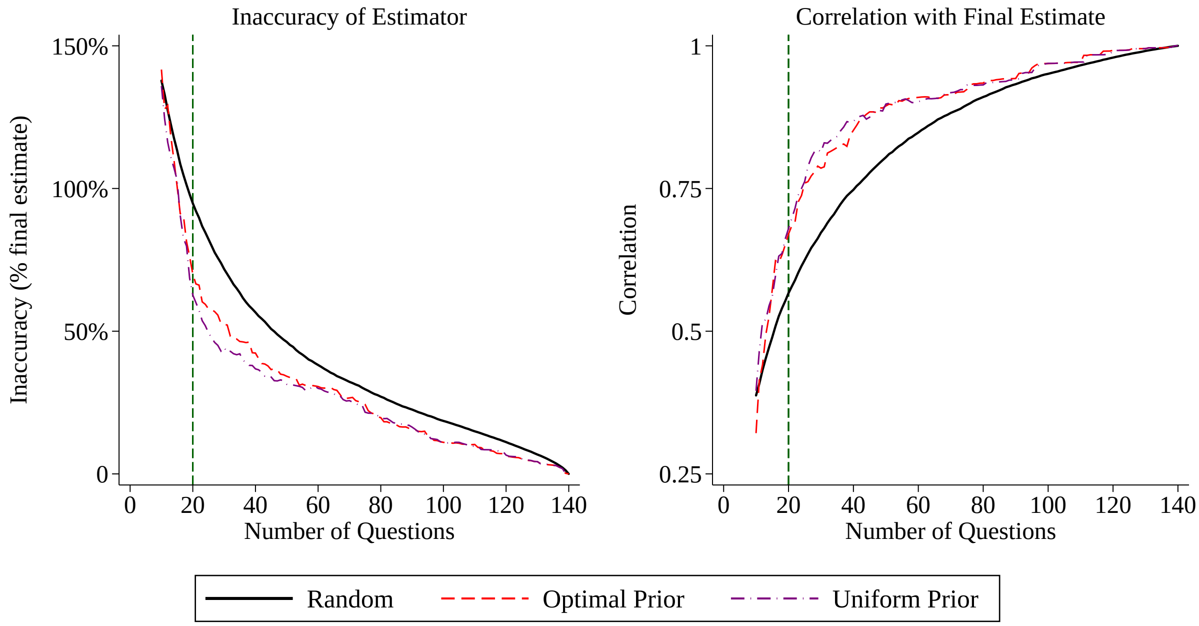
estimates. The consistency estimates take longer to converge—as discussed previously, this is likely to be a result of the fact that several similar questions have to be asked in order for the parameter to be pinned down precisely. However, there is no evidence for any of the parameters that the procedure converges faster for particular parameter values—DOSE performs well at the individual-level as well as on average.

G.2 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 participants.[10] As shown in Figure G.6, when using fewer than 40 questions (using the

[10] 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 G.2: DOSE elicits accurate estimates for the choice consistency parameter faster than the random ordering.

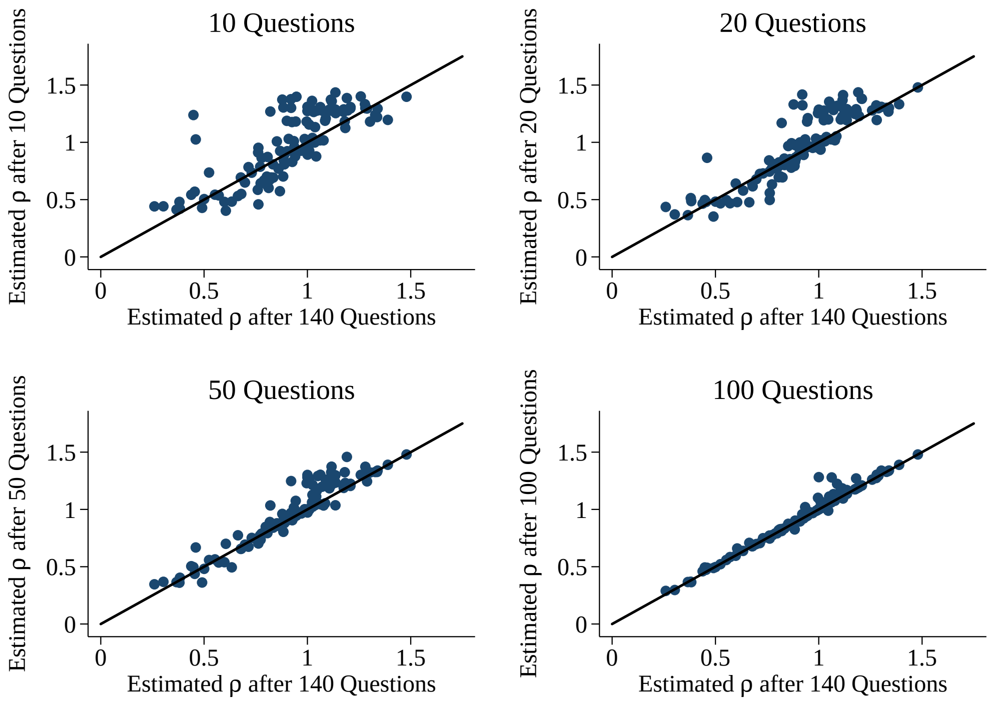


Notes: Based on data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). Left (right) hand panel shows the inaccuracy (correlation with final estimate) of Bayesian estimates (with uniform initial prior) obtained after each question, under different orders. "Optimal Prior" and "Uniform Prior"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

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 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

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



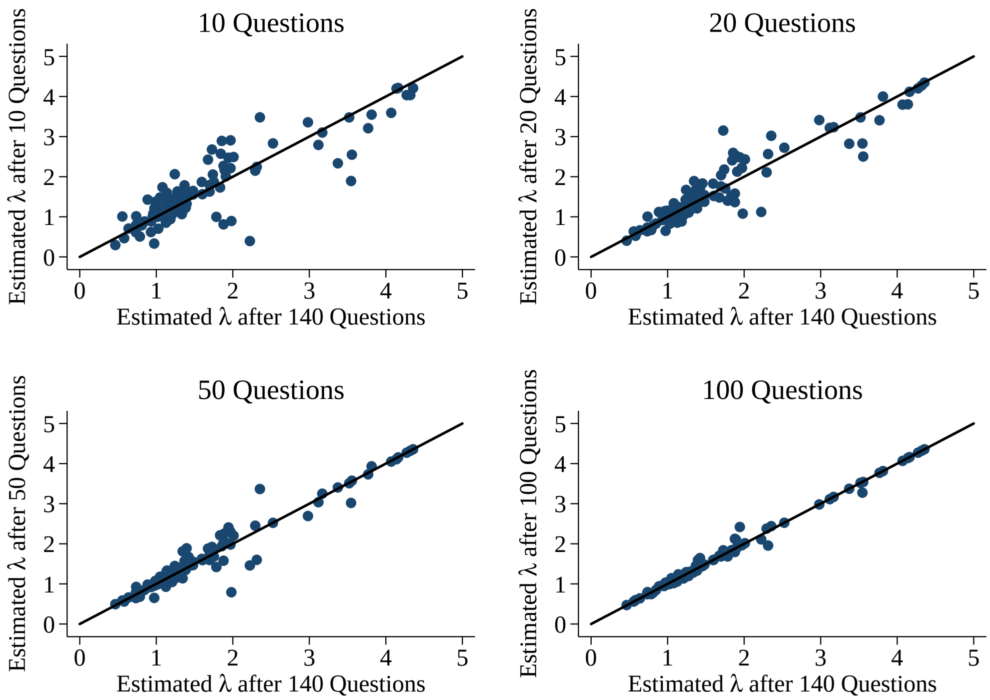
Notes: The figure is based on authors' analysis of data from Sokol-Hessner et al. (2009) and Frydman et al. (2011). Each panel plots the DOSE estimate (using a uniform prior) of the exponent from the utility function (1) against the Bayesian estimate after 140 questions.

line plots in Figure G.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.[11] In addition, the Bayesian estimates are generally more accurate than the MLE estimates that do exist even where no MLE estimate can be obtained at all.[12] Not only can the Bayesian procedure obtain an estimate in those circumstances, those estimates contain

[11] 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.

[12] Note that 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.

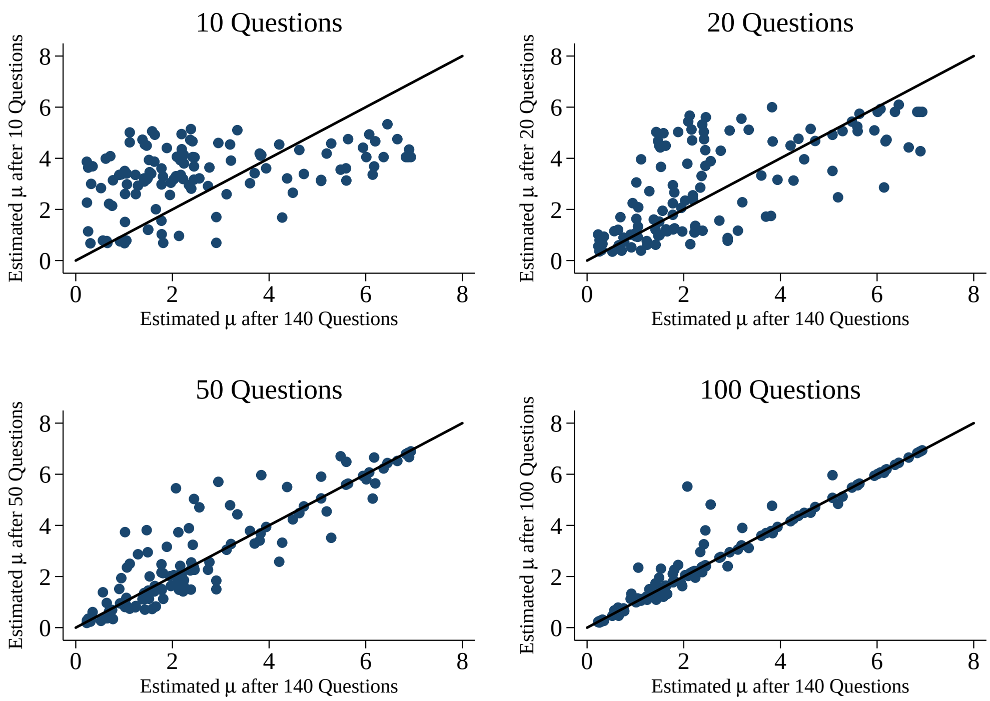
Figure G.4: Correlations between final estimates of the loss aversion parameter and the estimates after selected rounds.



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 (1) against the Bayesian estimate after 140 questions.

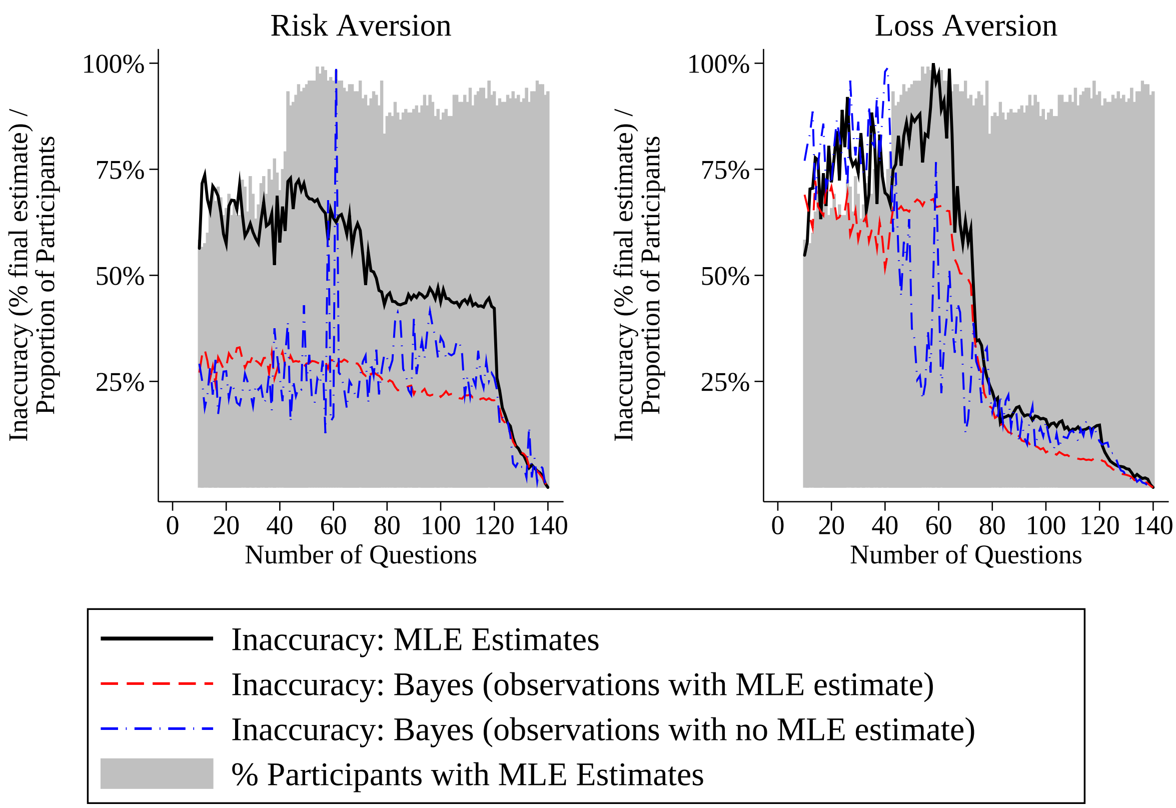
valuable information.

Figure G.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 (2) against the Bayesian estimate after 140 questions.

Figure G.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).