Online Appendix to "Who is Behavioral?"

Daniel J. Benjamin[*]

Cornell University and NBER

Sebastian A. Brown

Harvard University

Jesse M. Shapiro

University of Chicago and NBER

March 2012

1 Additional Results

Figure 1 presents histograms of the number of normative choices made in each of the key preference elicitation modules in studies 1 and 2.

Tables 1 and 2 reproduce our main tables for studies 1 and 2, respectively, using verbal instead of math scores. In the case of study 1 results are generally weaker with verbal scores than with math scores. In the case of study 2 results are somewhat stronger with verbal scores. The only qualitative difference in findings when using verbal scores is a sign change on the relationship with fairness preferences as measured by dictator-game giving.

Tables 3 and 4 show results for studies 1 and 2, respectively, comparing findings from the full sample with findings from the subset of respondents whose choices were "monotonic" in the sense that they can be characterized by a cutoff above which the respondent switched from choice "A" (the safer or immediate payoff option) to choice "B" (the riskier or delayed option). In study 1, restricting to monotonic respondents generally causes coefficients to shrink somewhat in magnitude. In study 2, it decreases one risk preference coefficient and increases the other (both remain statistically insignificant), and has little effect on time preference results.

Table 5 shows how the results from study 2 are affected by excluding cases with missing demographics. (In the models we present in the paper, we impute missing demographics at the sample mean.) For each preference measure we show three specifications. The first reproduces our baseline models with no controls. The second includes no controls but restricts to the subset of cases for which no demographics are missing. The third adds demographic controls. Excluding cases with imputed demographics leaves our qualitative conclusions unchanged.

Table 6 presents evidence on the relationship between math scores and present-bias. For each participant we compute the number of time-preference choices on which the participant behaved patiently when trading off rewards at four and five weeks but impatiently when trading off rewards between the present and one

---

[*]E-mail: daniel.benjamin@gmail.com, sebastian.brown@post.harvard.edu, jmshapir@uchicago.edu.

week. We estimate an ordered probit model relating the number of such present-biased choices to the math score. We find a negative and statistically insignificant relationship between math score and the number of present-biased choices.

Tables 7 and 8 show estimates for studies 1 and 2, respectively, showing coefficients on our main control variables and adding a specification that includes gender but not socioeconomic controls. In study 1, we find that male participants exhibit statistically insignificantly less normative behavior. The relationship between municipal income and normative behavior is statistically significant and inconsistent in direction. In study 2, we find that male participants are marginally statistically significantly more risk neutral, statistically insignificantly less patient, and statistically insignificantly more selfish. In the case of risk preference, household income and automobile wealth are statistically significant but have opposite signs. In the case of time preference, socioeconomic controls have no consistent sign or statistical significance.

Table 9 shows evidence on the importance of family background as a confound, using data on siblings from study 2. Our measure of cognitive ability is the average math GPA rather than the math score on the PSU practice exam, because PSU scores are not available for non-seniors. For each preference elicitation we show two specifications. The first uses all participants with participating siblings but does not control for sibling group fixed effects. We show this specification as a baseline because the sample and ability measure differ from that in the main specifications in the paper. The second specification in each group adds family fixed effects. These models should be interpreted cautiously because of the well-known incidental parameters problem with fixed effects in nonlinear models. In all cases the point estimate on math grade is positive in the model with sibling group fixed effects. In the case of the first risk preference elicitation, the coefficient is statistically significant. In all other cases, it is statistically insignificant.

Tables 10 and 11 present results on the relationship between preferences and elementary school GPA for studies 1 and 2, respectively. In study 1, the associations are similar to those with PSU scores, with the exception of the coefficient for risk preference in gains, which is smaller and no longer statistically significant. In study 2, the associations with elementary school GPA have the same signs but are weaker than the associations with PSU scores.

Table 12 presents an additional analysis of our cognitive load experiment. For each preference measure we compare the share of participants making entirely normative choices between those under load and those not under load. In contrast to our parametric test in the text, here we compute a Fisher exact p-value of the null hypothesis that there is no effect of cognitive load on the distribution of the dependent variable. As in our parametric test, only the effect of cognitive load on risky choices with equalized complexity is statistically significant.

Table 13 shows the relationship between math grades and treatment condition in our cognitive load study as a test of the balance of observables. The data are broken out into four groups according to the assigned questionnaire variant. A Fisher exact test does not reject the null hypothesis that the math score is balanced across experimental conditions.

Table 14 presents a bounding analysis of our cognitive load experiment. For each of the two prefer-

ence measures for which there were respondents who failed to answer at least one preference question, we present both our baseline specification (in which we exclude respondents with missing data) and a conservative "bounding" specification in which we assume that the missing preference is normative under load and non-normative otherwise. As expected, the estimated effect of cognitive load is weaker in the bounding specification. (Note that the table excludes the one preference measure for which we find a statistically significant effect of cognitive load, because for that measure there were no missing responses.)

Table 15 shows how the effect of our cognitive load manipulation differs with cognitive ability. We find no consistent evidence of an interaction between math grades and cognitive load. In the case of the second math quiz, we find a statistically significant negative interaction, indicating that cognitive load has a greater impact on those with greater baseline ability.

Table 16 shows how the effect of our reasoning manipulation differs with cognitive ability. This evidence suggests that the effect of additional reasoning on expressed preferences is greatest for those with the lowest measured mathematical ability. This effect is statistically significant for risk preference but not time preference.

2 Summary of Harvard Pilot

2.1 Methods

2.1.1 Participants

Participants were 60 undergraduates (virtually all Harvard students), recruited through on-campus posters and e-mail solicitations. We promised students $5 for participating in a 45-minute experiment, with the possibility to earn more "depending on your responses [in the experiment]." We allowed only non-economics majors to participate because we were concerned that economics students would be familiar with our preference elicitation procedures. In total, we held six sessions, on February 8, 21, and 22, and May 14, 15, and 16, 2004.

2.1.2 Procedure

An experimenter guided participants through a questionnaire in unison by reading instructions aloud. The questionnaire contained four preference-elicitation sections: small-stakes risk preferences, short-term time preferences, fairness preferences, and idiosyncratic consumption preferences. The order of the sections differed across sessions; this order had no effect on the results. As described below, in any given section, a random half of the participants were under cognitive load.

At the end of the questionnaire, after all the preference-elicitation questions, we asked participants for their major, year in school, gender, and highest Math and Verbal SAT I scores. 57 out of 60 participants had taken the SAT. As a check on the accuracy of the self-reported SAT scores, we included a section of SAT-like math questions near the end of the questionnaire (after participants had responded to all the

preference-elicitation sections). The correlation between performance on the SAT-like math questions and self-reported Math SAT score was 0.33, which is significantly different from 0 ($p = 0.012$).

We paid all participants $5 in cash for their participation immediately at the completion of the session. We also paid participants (by check) for their choices so that our preference-elicitation procedures were incentive-compatible. As described below, for some sections we paid participants immediately after the experiment. For other sections, we mailed checks to participants within a week of their participation.

Small-Stakes Risk Preference. Each question offered the choice between $0.50 for sure and a gamble that gave a 50% chance of winning $0 and a 50% chance of winning X, where X took the values $0.95, $1.05, $1.15, $1.25, and $1.35 in ascending order. We gave participants an example question in the instructions and the opportunity to ask questions.

For this and other sections, we gave participants 15 seconds to answer the five questions. We imposed this time limit because pretesting suggested that most participants finished within that short amount of time even when given unlimited time, and because allowing more time reduced the effectiveness of the cognitive load as a distracting task. Our intent was that participants in the cognitive load condition would attend to the cognitive load task concurrently with answering the preference questions. With more time, however, participants could more readily focus primarily on the cognitive load task, switching attention occasionally to answer a preference question.

Participants knew from the instructions that after they made a selection for each question, we would roll a die five times to determine their payment for this section. We asked a participant to roll the die to maximize our credibility. We paid participants for this section in the check that we mailed within a week of the individual's participation.

Short-Term Time Preference. After an example question and an opportunity to ask the experimenter about the instructions, participants chose between $5.00 today and X a week from now, where X was $5.05, $5.35, $5.55, $5.75, $5.95, and $6.15 in ascending order. Among the 57 participants for whom SAT scores are available, 6 did not answer all of the time preference questions (5 of whom were subjected to cognitive load, described below). These participants have been omitted from our analysis.

After giving participants 15 seconds to answer the six questions, we asked a participant to roll the die that would select the question to be implemented. The instructions explained that participants would receive a check to pay them for this section. That check would be post-dated by a week if the participant had chosen (B) for the relevant question. We gave participants their check immediately after the session.

Fairness Preference. To implement an anonymous dictator game, we informed participants that they had been randomly assigned to another participant in the same session, but that no one would ever find out who had been assigned to whom. To rule out bilateral reciprocity concerns, the experimenter made it clear that the other person had, in turn, been assigned to a different participant. As a result, each participant would affect the payoff of another participant but would not be affected by that other participant.

Participants were told they had been given $1.00, and they had the opportunity to give away $0.00, $0.25, $0.50, $0.75, or $1.00 to another participant. Participants were given 15 seconds to make a decision.

4

Idiosyncratic Preference. We asked participants five binary preference questions for which there is no normatively correct answer. We asked participants whether they preferred chocolate or coffee ice cream, red or silver-colored cars, cats or dogs as pets, Pepsi or Coke to drink, and Butterfinger or Kit Kat as a candy to eat. We informed participants that they would actually receive their preferred candy as part of their payment.

Cognitive Load Manipulation. We subjected a randomly chosen half of the participants to "cognitive load" during each section of the questionnaire. During each of the preference-elicitation sections of the questionnaire and during the SAT-like math questions, participants heard a CD of piano notes while they filled out the section. In each section, half the participants were required to remember the number of times they heard a specific sequence of musical tones. The sequence to be remembered varied across questionnaires. To incentivize participants in the cognitive load condition to pay attention to the tones, we made payment for that section contingent on correct recall of the number of repetitions of the sequence of tones.

In earlier pilot tests with Harvard students, we tried two other cognitive load procedures. We required participants to remember seven-digit numbers (as in Shiv and Fedorikhin, 1999, and Hinson, Jameson, and Whitney, 2003), and we also tried playing musical tones at a slower pace than in the actual experiment. Neither of these manipulations reliably influenced the preferences we measured.

2.2 Results

Table 17 shows the results of a probit analysis of the relationship between SAT scores and measured preferences that we conducted after we ran the pilot study and before we ran our studies in Chile. Column (1) shows that participants with above-median Math SAT scores have a 24 percentage point greater chance of behaving risk-neutrally, which is economically large and statistically significant at the 10% level. In column (2) when we include measures of both Math and Verbal SAT score, the estimated effect of an above-median Math SAT score increases and becomes significant at the 5% level, whereas we find a statistically insignificant negative effect of having an above-median Verbal SAT score. Our estimates in columns (3) and (4) of the effect of SAT scores show positive (though not quite statistically significant) effects of mathematical ability on patience and a small and insignificant negative effect of verbal ability.

As columns (5) and (6) show, we find no evidence of a relationship between Math SAT score and "selfishness" (defined as keeping 100% of the dollar available for splitting). The point estimate indicates that more cognitively able individuals are slightly less likely to behave selfishly, but this estimate is statistically indistinguishable from zero.

For four of our five idiosyncratic preference choices, we find no statistically significant relationship with Math SAT score (results not shown), although in our data, individuals with higher Math SAT scores are statistically significantly more likely to prefer cats over dogs.

Column (1) of Table 18 shows that cognitive load reduced the number of correct answers on our six-question battery of SAT-like math questions by about .7 on average, and a Mann-Whitney test rejects the null of no effect at the 10% level ($p = .080$). Columns (2)-(4) show that in this sample, cognitive load did

not have a statistically significant effect on measured risk-neutrality, expressed time preference, or dictator game behavior.

References

Hinson, J.M., T.J. Jameson, and P. Whitney. Impulsive decision making and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29:298-306, 2003.

Shiv, Baba and Alexander Fedorikhin. Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, 26(3):27-292, 1999.

Figure 1: *Distribution of Main Dependent Measures: Studies 1 and 2*

Variable plotted: Number of normative (risk neutral, patient, selfish) choices

Study 1
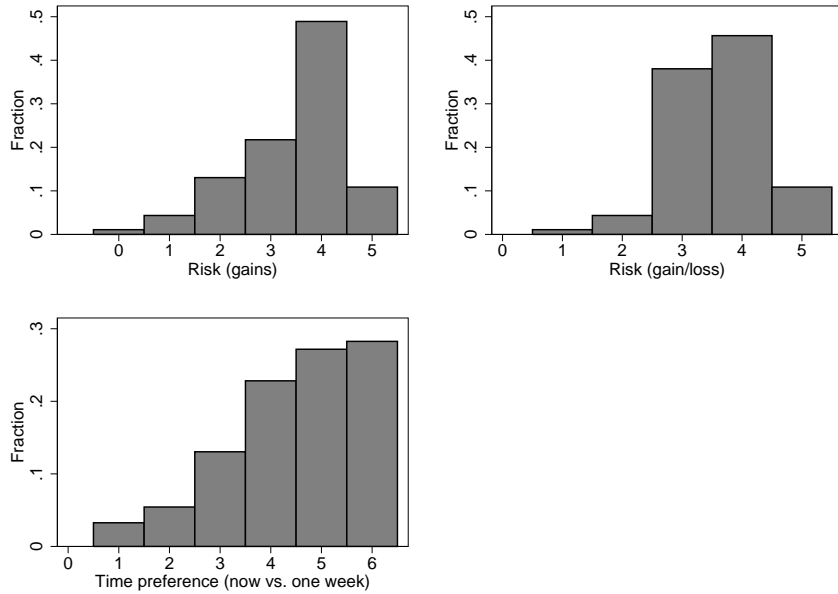
Study 2
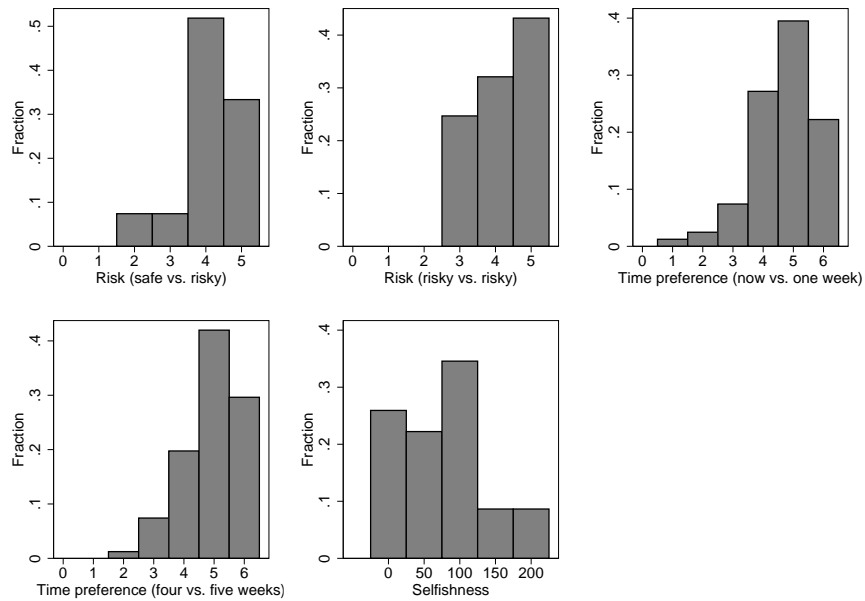
Notes: Figure plots the distribution of the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation.

Table 1: *Preferences and Verbal Scores: Study 1*

Dependent variable: Number of normative (risk neutral, patient) choices

Preference type	Risk (gains)		Risk (gain/loss)		Time preference (now vs. one week)	
	(1)	(2)	(3)	(4)	(5)	(6)
Standardized verbal score	0.2521	0.2487	0.0794	0.0736	0.2008	0.2068
	(0.1140)	(0.1143)	(0.1141)	(0.1145)	(0.1109)	(0.1113)
	[0.0453]	[0.0444]	[0.0148]	[0.0136]	[0.0667]	[0.0683]
Demographic controls?		X		X		X
Share making all normative choices	0.1087	0.1087	0.1087	0.1087	0.2826	0.2826
Pseudo-R^2	0.0195	0.0209	0.0023	0.0074	0.0113	0.0147
N	92	92	92	92	92	92

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation. Demographic controls are gender and the mean income in the participant's municipality of residence.

Table 2: *Preferences and Verbal Scores: Study 2*

Dependent variable: Number of normative (risk neutral, patient, selfish) choices

Preference type	Risk (safe vs. risky)		Risk (risky vs. risky)		Time preference (now vs. one week)		Time preference (four vs. five weeks)		Selfishness	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Standardized verbal score	0.1629	0.1339	0.1854	0.2331	0.2693	0.2812	0.1449	0.1221	0.2103	0.2616
	(0.1251)	(0.1395)	(0.1287)	(0.1432)	(0.1229)	(0.1344)	(0.1232)	(0.1348)	(0.1199)	(0.1320)
	[0.0584]	[0.0437]	[0.0718]	[0.0851]	[0.0774]	[0.0798]	[0.0496]	[0.0411]	[0.0323]	[0.0392]
Demographic controls?		X		X		X		X		X
Share making all normative choices	0.3333	0.3333	0.4321	0.4321	0.2222	0.2222	0.2963	0.2963	0.0864	0.0864
Pseudo-R^2	0.0096	0.0878	0.0120	0.0584	0.0214	0.0315	0.0066	0.0177	0.0129	0.0300
N	81	81	81	81	81	81	81	81	81	81

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation. Demographic controls are gender, father's years of schooling, mother's years of schooling, household monthly income, and the total market value of the family's automobiles.

Table 3: *Restricting to Monotonic Choices: Study 1*

Dependent variable: Number of normative (risk neutral, patient) choices

Preference type	Risk (gains)		Risk (gain/loss)		Time preference (now vs. one week)	
	(1)	(2)	(3)	(4)	(5)	(6)
Standardized math score	0.3059	0.1983	0.2393	0.2044	0.2125	0.2044
	(0.1192)	(0.1295)	(0.1172)	(0.1288)	(0.1147)	(0.1161)
	[0.0535]	[0.0435]	[0.0432]	[0.0416]	[0.0703]	[0.0695]
Sample	All	Monotonic	All	Monotonic	All	Monotonic
Share making all normative choices	0.1087	0.1429	0.1087	0.1266	0.2826	0.2989
Pseudo-R^2	0.0268	0.0168	0.0200	0.0143	0.0120	0.0115
N	92	70	92	79	92	87

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation. Monotonic sample is the subset of respondents whose chose option "A" (safe/money today) until the incentive to choose option "B" (risky/money in the future) grew strong enough and then chose option "B" thereafter. The monotonic sample includes participants who chose "A" on every question as well as those who chose "B" on every question.

Table 4: *Restricting to Monotonic Choices: Study 2*

Dependent variable: Number of normative (risk neutral, patient, selfish) choices

Preference type	Risk (safe vs. risky)		Risk (risky vs. risky)		Time preference (now vs. one week)		Time preference (four vs. five weeks)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Standardized math score	0.1502	-0.0406	0.1656	0.2377	0.3887	0.3831	0.1726	0.1722
	(0.1257)	(0.1539)	(0.1297)	(0.1398)	(0.1267)	(0.1268)	(0.1232)	(0.1233)
	[0.0539]	[-0.0158]	[0.0643]	[0.0919]	[0.1074]	[0.1068]	[0.0588]	[0.0589]
Sample	All	Monotonic	All	Monotonic	All	Monotonic	All	Monotonic
Share making all normative choices	0.3333	0.4154	0.4321	0.4730	0.2222	0.2250	0.2963	0.3000
Pseudo-R^2	0.0081	0.0007	0.0095	0.0194	0.0428	0.0425	0.0094	0.0095
N	81	65	81	74	81	80	81	80

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation. Demographic controls are gender, father's years of schooling, mother's years of schooling, household monthly income, and the total market value of the family's automobiles. Monotonic sample is the subset of respondents whose chose option "A" (safe/money today) until the incentive to choose option "B" (risky/money in the future) grew strong enough and then chose option "B" thereafter. The monotonic sample includes participants who chose "A" on every question as well as those who chose "B" on every question.

Table 5: *Excluding Observations with Imputed Demographics: Study 2*

Dependent variable: Number of normative (risk neutral, patient, selfish) choices

Preference type	Risk (safe vs. risky)			Risk (risky vs. risky)			Time preference (now vs. one week)			Time preference (four vs. five weeks)			Selfishness		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Standardized math score	0.1502	0.1712	0.0657	0.1656	0.2168	0.1835	0.3887	0.3346	0.4020	0.1726	0.2320	0.1981	-0.1298	-0.0851	-0.1397
	(0.1257)	(0.1419)	(0.1557)	(0.1297)	(0.1494)	(0.1662)	(0.1267)	(0.1416)	(0.1539)	(0.1232)	(0.1401)	(0.1501)	(0.1228)	(0.1366)	(0.1481)
	[0.0539]	[0.0646]	[0.0223]	[0.0643]	[0.0838]	[0.0633]	[0.1074]	[0.0768]	[0.0886]	[0.0588]	[0.0688]	[0.0558]	[-0.0204]	[-0.0157]	[-0.0237]
Sample with all demographics?		X	X		X	X		X	X		X	X		X	X
Demographic controls?			X			X			X			X			X
Share making all normative choices	0.3333	0.3929	0.3929	0.4321	0.4643	0.4643	0.2222	0.1607	0.1607	0.2963	0.2321	0.2321	0.0864	0.1071	0.1071
Pseudo-R^2	0.0081	0.0116	0.1022	0.0095	0.0183	0.1052	0.0428	0.0390	0.0697	0.0094	0.0191	0.0536	0.0047	0.0023	0.0500
N	81	56	56	81	56	56	81	56	56	81	56	56	81	56	56

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation.

Table 6: *Evidence on Present-Bias: Study 2*
Dependent variable: Number of present-biased choices

	(1)	(2)
Standardized average math grade	-0.2027	-0.2364
	(0.1486)	(0.1539)
	[-0.0724]	[-0.0826]
Demographic controls?		X
Share making all normative choices	0.6667	0.6667
Pseudo-R^2	0.0135	0.0287
N	81	81

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of no present-biased choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of present-biased choices made, where a present-biased choice means making a patient choice in the "four vs. five weeks" elicitation but an impatient choice in the "now vs. one week" elicitation with identical rewards. See "Methods and Data" section for details of preference elicitation. Demographic controls are gender, father's years of schooling, mother's years of schooling, household monthly income, and the total market value of the family's automobiles.

Table 7: *Additional Evidence on Demographic Controls: Study 1*

Dependent variable: Number of normative (risk neutral, patient) choices

Preference type	Risk (gains)			Risk (gain/loss)			Time preference (now vs. one week)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Standardized math score	0.3059	0.3250	0.3361	0.2393	0.2623	0.2486	0.2125	0.2850	0.2797
	(0.1192)	(0.1306)	(0.1330)	(0.1172)	(0.1278)	(0.1295)	(0.1147)	(0.1250)	(0.1268)
Male		-0.0919	-0.0824		-0.1162	-0.1332		-0.3706	-0.3775
		(0.2536)	(0.2546)		(0.2562)	(0.2576)		(0.2483)	(0.2499)
Average income in municipality			-0.0550			0.0820			0.0298
(standardized)			(0.1193)			(0.1216)			(0.1189)
Pseudo-R^2	0.0268	0.0273	0.0282	0.0200	0.0209	0.0231	0.0120	0.0196	0.0198
N	92	92	92	92	92	92	92	92	92

Notes: Results are from ordered probit models, with standard errors in parentheses. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation.

Table 8: *Additional Evidence on Demographic Controls: Study 2*

Dependent variable: Number of normative (risk neutral, patient, selfish) choices

Preference type	Risk (safe vs. risky)			Risk (risky vs. risky)			Time preference (now vs. one week)			Time preference (four vs. five weeks)			Selfishness		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Standardized math score	0.1502	0.1083	0.0684	0.1656	0.1558	0.1464	0.3887	0.4191	0.4059	0.1726	0.1822	0.1524	-0.1298	-0.1529	-0.1454
	(0.1257)	(0.1287)	(0.1333)	(0.1297)	(0.1320)	(0.1365)	(0.1267)	(0.1295)	(0.1323)	(0.1232)	(0.1251)	(0.1277)	(0.1228)	(0.1248)	(0.1278)
Male		0.5024	0.4660		0.1108	0.0473		-0.2943	-0.3186		-0.1059	-0.0891		0.2480	0.3103
		(0.2597)	(0.2713)		(0.2585)	(0.2690)		(0.2449)	(0.2549)		(0.2448)	(0.2552)		(0.2421)	(0.2518)
Household income (standardized)			-0.4196			-0.4072			-0.1363			-0.2037			0.1176
			(0.1772)			(0.1876)			(0.1588)			(0.1610)			(0.1580)
Father's years of schooling			0.0917			0.0057			0.0422			0.0130			-0.0665
			(0.0693)			(0.0687)			(0.0647)			(0.0655)			(0.0650)
Mother's years of schooling			-0.0478			-0.0527			0.0071			0.0517			0.0208
			(0.0773)			(0.0764)			(0.0719)			(0.0731)			(0.0711)
Household automobile wealth (standardized)			0.3793			0.2832			0.0258			0.1461			0.0683
			(0.1722)			(0.1762)			(0.1526)			(0.1563)			(0.1522)
Pseudo-R^2	0.0081	0.0295	0.0841	0.0095	0.0106	0.0497	0.0428	0.0492	0.0548	0.0094	0.0103	0.0206	0.0047	0.0091	0.0190
N	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81

Notes: Results are from ordered probit models, with standard errors in parentheses. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation.

Table 9: *Using Siblings to Control for Family Background*

Dependent variable: Number of normative (risk neutral, patient, selfish) choices

Preference type	Risk (safe vs. risky)		Risk (risky vs. risky)		Time preference (now vs. one week)		Time preference (four vs. five weeks)		Selfishness	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Standardized average math grade	0.1135	1.6181	-0.3440	0.0502	0.1272	0.0926	0.0152	0.4267	0.2522	0.0990
	(0.1850)	(0.4675)	(0.2307)	(0.3906)	(0.1859)	(0.2971)	(0.1830)	(0.2951)	(0.1953)	(0.3589)
	[0.0433]	[0.3017]	[-0.1302]	[0.0130]	[0.0415]	[0.0157]	[0.0057]	[0.1088]	[0.0481]	[0.0116]
Family fixed effects?		X		X		X		X		X
Share making all normative choices	0.3953	0.3953	0.5581	0.5581	0.2791	0.2791	0.3721	0.3721	0.1163	0.1163
Pseudo-R^2	0.0051	0.3435	0.0234	0.2174	0.0177	0.2878	0.0010	0.2408	0.0151	0.3528
N	43	43	43	43	43	43	43	43	43	43

Notes: Data are from study 2. Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation. All specifications control for gender. Standardized average math grade is the average math grade across all years the participant was enrolled at the school, standardized by the sample standard deviation.

Table 10: *Preferences and Elementary School Performance: Study 1*

Dependent variable: Number of normative (risk neutral, patient) choices

Preference type	Risk (gains)	Risk (gain/loss)	Time preference (now vs. one week)
	(1)	(2)	(3)
Standardized math GPA	0.1653	0.2784	0.2431
(elementary school)	(0.1183)	(0.1217)	(0.1158)
	[0.0296]	[0.0496]	[0.0801]
Share making all normative choices	0.1059	0.1059	0.2824
Pseudo-R^2	0.0082	0.0269	0.0163
N	85	85	85

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation. Elementary school is defined as grades 1-6.

Table 11: *Preferences and Elementary School Performance: Study 2*

Dependent variable: Number of normative (risk neutral, patient, selfish) choices

Preference type	Risk (safe vs. risky)	Risk (risky vs. risky)	Time preference (now vs. one week)	Time preference (four vs. five weeks)	Selfishness
	(1)	(2)	(3)	(4)	(5)
Standardized math GPA	0.0388	0.0587	0.0041	0.0960	-0.0600
(elementary school)	(0.1320)	(0.1327)	(0.1276)	(0.1286)	(0.1282)
	[0.0133]	[0.0231]	[0.0012]	[0.0336]	[-0.0089]
Share making all normative choices	0.2933	0.4400	0.2267	0.3067	0.0800
Pseudo-R^2	0.0005	0.0012	0.0000	0.0029	0.0010
N	75	75	75	75	75

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation. Elementary school is defined as grades 1-6.

Table 12: *Fisher Exact Tests for Cognitive Load: Study 3*

Dependent variable: Dummy for perfectly normative (risk-neutral, patient, selfish, mathematically correct) decision-making

Preference type	Risk neutral		Patient		Selfish	Math quiz	
	Safe vs. Risky	Risky vs. Risky	Now vs. One week	Four vs. Five weeks		One	Two
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Not under load	0.3077	0.5882	0.2857	0.3333	0.0000	0.1333	0.3182
Under load	0.1429	0.2000	0.1333	0.3158	0.1176	0.0909	0.2000
Fisher exact *p*-value	0.387	0.021	0.424	1.000	0.204	1.000	0.481
No. of observations	34	37	36	37	37	37	37
No. not under load	13	17	21	18	20	15	22
No. under load	21	20	15	19	17	22	15

Notes: Cognitive load indicates that the participant was asked to remember a seven-digit number while performing the task. In all cases, dependent measure is a dummy for whether the participant behaved in all choices in a manner consistent with the maximization of un-discounted expected value. For each of the two math quizzes, the "math quiz" dependent measure is a dummy for answering all questions correctly. See "Methods and Data" section for details of preference elicitation and math quizzes.

Table 13: *Balance of Cognitive Ability by Treatment Condition: Study 3*

Treatment condition	Number with average math grade:		
(Questionnaire format)	Below median	At or above median	Total
1	6	4	10
2	4	4	8
3	7	5	12
4	2	5	7
Fisher exact *p*-value	0.636		

Notes: Table shows number of participants with below- or above-median average math GPA for each of the four questionnaire variants we used in the study. Fisher exact *p*-value is for the null hypothesis that the likelihood of having an above-median math grade is independent of questionnaire format.

Table 14: *Missing Value Bounds for Cognitive Load: Study 3*

Dependent variable: Number of normative (risk neutral, patient, selfish) choices

Preference type	Risk		Time preference	
	(safe vs. risky)		(now vs. one week)	
	(1)	(2)	(3)	(4)
Cognitive Load	-0.3684	-0.0731	-0.0364	0.0107
	(0.3785)	(0.3545)	(0.3543)	(0.3503)
	[-0.1035]	[-0.0198]	[-0.0108]	[0.0031]
Specification	Excluded	Bounds	Excluded	Bounds
Share making all normative choices	0.2059	0.1892	0.2222	0.2162
Pseudo-R^2	0.0102	0.0004	0.0001	0.0000
N	34	37	36	37

Notes: Results are from ordered probit models, with standard errors in parentheses and estimated marginal effect on the probability of all normative choices (evaluated at the sample mean of the independent variables) in brackets. Dependent variable is the number of normative choices made in the given preference elicitation. In the "excluded" specification, participants with missing data on one or more preferences are excluded from the sample. In the "bounds" specification, missing data is imputed to be normative under load and non-normative otherwise. See "Methods and Data" section for details of preference elicitation.

Table 15: *How Effects of Cognitive Load Vary With Cognitive Ability: Study 3*

Dependent variable: Number of normative (risk neutral, patient, selfish, mathematically correct) choices

Preference type	Risk (safe vs. risky)	Risk (risky vs. risky)	Time preference (now vs. one week)	Time preference (four vs. five weeks)	Selfishness	Math quiz 1	Math quiz 2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cognitive load	1.1831	0.5060	0.2897	0.0702	0.0914	1.0314	1.6057
	(0.4059)	(0.3357)	(0.2483)	(0.2623)	(0.2341)	(0.3189)	(0.3515)
Standardized math GPA	-0.9753	-1.0182	0.0833	0.1865	0.0479	-0.1858	-0.0023
(high school)	(0.4306)	(0.3919)	(0.3758)	(0.3619)	(0.3521)	(0.3834)	(0.3869)
Cognitive load X	-0.2314	-0.4950	-0.1543	0.0899	0.1447	-0.4297	-0.9753
Standardized math GPA	(0.4579)	(0.4123)	(0.3733)	(0.3613)	(0.3578)	(0.3902)	(0.4331)
Pseudo-R^2	0.2127	0.1005	0.0139	0.0074	0.0092	0.1451	0.2342
N	34	37	36	37	37	37	37

Notes: Results are from ordered probit models, with standard errors in parentheses. Dependent variable is the number of normative choices made in the given preference elicitation. See "Methods and Data" section for details of preference elicitation and math quizzes.

Table 16: *How Effects of Reasoning Vary With Cognitive Ability: Study 3*

Dependent variable	*Difference* between reasoning and non-reasoning task in number of choices consistent with...	
	Risk-neutrality (Safe vs. Risky)	Patience (Now vs. One Week)
	(1)	(2)
Standardized math GPA	-0.5099	-0.1054
(high school)	(0.1571)	(0.2100)
Cognitive load in non-reasoning task	0.5049	-0.2150
	(0.3064)	(0.4220)
Constant	0.2858	0.8258
	(0.2367)	(0.2655)
R^2	0.2694	0.0128
No. of observations	34	34

Notes: Results are from OLS regressions, with standard errors in parentheses. Reasoning task indicates that the participant was asked to think about and express the reasons for her choice. See "Methods and Data" section for details of preference elicitation.

Table 17: *Preferences and Cognitive Ability: Harvard Undergraduates*

Dependent variable	Risk neutral (dummy)		Patient (dummy)		Selfish (dummy)	
	(1)	(2)	(3)	(4)	(5)	(6)
Math SAT $\geq$ median	0.2446	0.2951	0.2549	0.2615	-0.1635	-0.1295
	(0.1222)	(0.1317)	(0.1522)	(0.1569)	(0.1421)	(0.1495)
Verbal SAT $\geq$ median		-0.1333		-0.0253		-0.1244
		(0.1335)		(0.1499)		(0.1468)
Male	0.1512	0.1279	-0.2136	-0.2158	-0.3548	-0.3781
	(0.1395)	(0.1407)	(0.1582)	(0.1586)	(0.1378)	(0.1387)
Year in school (1-4)	0.0710	0.0771	0.0817	0.0812	-0.1493	-0.1583
	(0.0715)	(0.0720)	(0.0847)	(0.0848)	(0.0827)	(0.0846)
Mean of dependent variable	0.2982	0.2982	0.5294	0.5294	0.4912	0.4912
Pseudo-R^2	0.0809	0.0952	0.0635	0.0639	0.1090	0.1180
N	57	57	51	51	57	57

Notes: Results are from probit models. Coefficients are estimated marginal effects evaluated at the sample mean of the independent variables, with standard errors in parentheses. In all cases, dependent measure is a dummy for whether the participant behaved in all choices in a manner consistent with the maximization of un-discounted expected value. See "Procedure" section of this appendix for details of preference elicitation. Data on SAT scores are from participants' self-reports.

Table 18: *The Effect of Cognitive Load: Harvard Undergraduates*

	Math score (0-6) (1)	Risk neutral (dummy) (2)	Patient (dummy) (3)	Selfish (dummy) (4)
Mean for condition:				
No cognitive load	3.10	0.2424	0.5000	0.3704
Cognitive load	2.40	0.3704	0.5185	0.5455
Test	Mann-Whitney	Fisher exact	Fisher exact	Fisher exact
p-value	0.080	0.397	1.000	0.203
N	60	60	53	60

Notes: Cognitive load indicates that the participant was asked to keep track of the number of times they heard a specific sequence of musical tones while performing the task. The "math score" dependent measure is the number of math quiz questions answered correctly out of six. In other cases, dependent measure is a dummy for whether the participant behaved in all choices in a manner consistent with the maximization of un-discounted expected value. See "Procedure" section of this appendix for details of preference elicitation and math quizzes.