Online Appendix 5 to "Narrow Bracketing and Dominated Choices": Regression analysis of choices on background variables

Matthew Rabin and Georg Weizsäcker

We use the variables descried in Online Appendix 4 as our main explanatory variables of interest. At least the first three of these variables (gender, age, racial/ethnic background) can safely be viewed as exogenous so that in addition to regressions on the full set of variables, we add a set of regressions where the dependent variables are regressed on each variable alone, with no other person-specific control variables.[1] In all regressions, we also include a dummy variable for the experimental treatment, because the tasks were different between the treatments.

As dependent variable, we first take a dummy that indicates whether the decisionmaker chooses a combination of lotteries that is FOS-dominated by another available combination (an A and D choice). Further, we run three additional regressions with alternative left-hand side variables: dummies that indicate whether (i) the decisionmaker makes a weakly risk averse choice in the gains domain (i.e. she chooses a sure positive outcome over a lottery that has two positive outcomes and weakly higher expected value), (ii) whether she makes a weakly risk averse choice in the losses domain,[2] and (iii) whether she makes a loss-averse choice, i.e. she is risk averse with regard to a lottery that has one positive and one negative outcome. Comparisons between these regressions will indicate whether any between-group differences in FOSD violations can be explained by differences in the revealed degrees of risk aversion between the groups. In each regression, we include only those decisionmakers who have the respective option available to generate both possible values of

[1] The following are the variables that we used in the regressions with all regressors: gender, performance in the three maths questions (4 categories, for 0-3 correct answers), dummy for self-reported attendance of a maths course at college, age, education level (4 categories), racial/ethnic background (5 categories), region of residence in the U.S. (4 categories), household size, marital status (5 categories), dummy for living in a metropolitan area, log income (20 brackets), housing status (3 categories, rent/own/do not pay for housing), employment status (9 categories).

[2] Observations of Decision 2 in Example 4 are counted for this variable, althought the high payoff of the risky option lies above 0. We decided to include this decision as the bevavior is arguably driven mostly by the risk attitudes below 0. These decisions are not counted for the next category, i.e. decisions that indicate degrees of loss aversion.

the dependent dummy variable, i.e. who face at least one relevant decision. Tables OA5.1 through OA5.6 show the resulting marginal odds ratios from logistic regressions.

Dependent var	Dominated		R. av. - gains		R. av. - losses		Loss av.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
female (odds ratio)	1.01	1.00	1.11	1.11	1.13	1.21	1.05	1.20
	(0.14)	(0.14)	(0.20)	(0.22)	(0.16)	0.19	(0.22)	(0.29)
controls	no	yes	no	yes	no	yes	no	yes
treatment dummies	yes	yes	yes	yes	yes	yes	yes	yes
Mean of dep. var	0.501	0.500	0.742	0.741	0.350	0.356	0.728	0.728
Pseudo-R^2	0.026	0.062	0.005	0.048	0.037	0.066	0.004	0.089
# of obs.	916	914	643	641	926	924	471	471

Table OA5.1: Logistic regressions on gender dummy.

Note: Robust standard error in parentheses. *: $p = 0.1$, **: $p = 0.05$, ***: $p = 0.01$.

Dependent var	Dominated			R. av. - gains		R. av. - losses		Loss av.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age (odds ratio)	1.010**	1.038*	1.008	1.004	1.006	0.992*	1.001	0.995	0.999
	(0.004)	(0.024)	(0.007)	(0.006)	(0.010)	(0.004)	(0.007)	(0.006)	(0.011)
age^2	-	0.9997	-	-	-	-	-	-	-
		(0.0002)	-	-	-	-	-	-	-
controls	no	no	yes	no	yes	no	yes	no	yes
module dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Mean of dep. var	0.501	0.501	0.500	0.742	0.741	0.350	0.356	0.728	0.728
Pseudo-R^2	0.030	0.032	0.062	0.006	0.048	0.039	0.066	0.005	0.089
# of obs.	916	916	914	643	641	926	924	471	471

Table OA5.2: Logistic regressions on age.

Note: Robust standard error in parentheses. *: $p = 0.1$, **: $p = 0.05$, ***: $p = 0.01$.

Dependent var	Dominated		R. av. - gains		R. av. - losses		Loss av.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
black (odds ratio)	0.67*	0.69	0.92	1.21	1.18	1.38	0.50**	0.60
	(0.15)	(0.17)	(0.29)	(0.43)	(0.28)	(0.36)	(0.16)	(0.23)
hispanic	0.49***	0.60**	0.49***	0.62*	0.97	0.98	0.38***	0.60
	(0.11)	(0.15)	(0.13)	(0.18)	(0.23)	(0.25)	(0.12)	(0.22)
2+ races, non-hispanic	0.94	0.93	2.13	2.01	1.14	1.25	1.06	1.52
	(0.32)	(0.33)	(1.36)	(1.30)	(0.42)	(0.46)	(0.69)	(0.92)
other, non-hispanic	0.69	0.80	0.48	0.48	1.17	1.09	1.01	1.38
	(0.26)	(0.33)	(0.23)	(0.25)	(0.47)	(0.46)	(0.82)	(1.18)
controls	no	yes	no	yes	no	yes	no	yes
module dummies	yes	yes	yes	yes	yes	yes	yes	yes
Mean of dep. var	0.501	0.500	0.742	0.741	0.350	0.356	0.728	0.728
Pseudo-R^2	0.036	0.062	0.056	0.048	0.037	0.066	0.026	0.089
# of obs.	916	914	643	641	926	924	471	471

Table OA5.3: Logistic regressions on racial/ethnic background categories.

Note: Omitted category is white, non-hispanic. Robust standard error in parentheses.

*: $p = 0.1$, **: $p = 0.05$, ***: $p = 0.01$.

Dependent var	Dominated		R. av. - gains		R. av. - losses		Loss av.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log income (odds ratio)	0.99	0.84*	1.14	0.93	1.19**	1.33***	1.24	1.05
	(0.07)	(0.08)	(0.11)	(0.11)	(0.09)	0.14	(0.16)	(0.19)
controls	no	yes	no	yes	no	yes	no	yes
module dummies	yes	yes	yes	yes	yes	yes	yes	yes
Mean of dep. var	0.501	0.500	0.742	0.741	0.350	0.356	0.728	0.728
Pseudo-R^2	0.026	0.062	0.007	0.048	0.040	0.066	0.010	0.089
# of obs.	916	914	643	641	926	924	471	471

Table OA5.4: Logistic regressions on log income.

Note: Robust standard error in parentheses. *: $p = 0.1$, **: $p = 0.05$, ***: $p = 0.01$.

Dependent var	Dominated				R. av. - gains			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
math correct (odds ratio)	0.80	-	-	0.76	1.11	-	-	1.11
	(0.17)	-	-	(0.16)	(0.31)	-	-	(0.31)
college math course	-	1.33*	-	1.34	-	0.93	-	0.84
	-	(0.21)	-	(0.27)	-	(0.19)	-	(0.21)
bachelor's degree	-	-	1.19	1.04	-	-	1.08	1.18
	-	-	(0.20)	(0.23)	-	-	(0.24)	(0.32)
controls	yes	yes	yes	yes	yes	yes	yes	yes
Mean of dep. var	0.508	0.508	0.508	0.508	0.747	0.747	0.747	0.747
Pseudo-R^2	0.056	0.057	0.056	0.059	0.036	0.036	0.036	0.037
# of obs.	884	884	884	884	621	621	621	621

Table OA5.5: Logistic regressions of FOSD choices and risk averse choices in gains domain on math-skills related variables. Note: Robust standard error in parentheses.
*: $p = 0.1$, **: $p = 0.05$, ***: $p = 0.01$.

Dependent var	R. av. - losses				Loss av.			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
math correct (odds ratio)	1.45*	-	-	1.52**	0.82	-	-	0.84
	(0.30)	-	-	(0.32)	(0.29)	-	-	(0.32)
college math course	-	0.84	-	0.82	-	1.21	-	1.61
	-	(0.14)	-	(0.17)	-	(0.33)	-	(0.53)
bachelor's degree	-	-	0.90	0.95	-	-	0.85	0.64
	-	-	(0.16)	(0.21)	-	-	(0.24)	(0.23)
controls	yes	yes	yes	yes	yes	yes	yes	yes
Mean of dep. var	0.344	0.344	0.344	0.344	0.735	0.735	0.735	0.735
Pseudo-R^2	0.064	0.062	0.062	0.066	0.081	0.081	0.081	0.085
# of obs.	893	893	893	893	453	453	453	453

Table OA5.6: Logistic regressions of risk averse choices around zero and in losses domain on education and math-skills related variables. Note: Robust standard error in parentheses.
*: $p = 0.1$, **: $p = 0.05$, ***: $p = 0.01$.

The tables show that between most subgroups of respondents, the differences in FOSD violations are insignificant. In particular, Table OA5.1 indicates no effect at all of the decisionmaker's gender, not even on the revealed degrees of risk aversion in any of the three subsets of lottery outcomes that we consider. Before turning to the other explanatory variables of interest, notice another indication of the lack of explanatory power of any background variable — the Pseudo-R^2 value lies below 0.1 for all regressions.

Age has a mildly adverse effect on behavior (see Table OA5.2), in the sense that older participants make weakly more FOS-dominated choices. But the significance of the effect goes away when squared age is included, or when other controls are included, so we regard it as an non-robust effect.

Nonwhites makes significantly fewer FOS-dominated choices, and in particular the hispanic population in the sample has a much lower frequency of FOSD violations. Table OA5.3 shows that their frequency is at less than 50 percent of the frequency of the omitted category (white, non-hispanic). Furthermore, columns (3), (5) and (7) give a more detailed account of these differences, indicating that hispanics exhibit a similar behavior when the offered lotteries that lie in the negative domain, but that they are much less risk averse around zero and in the domain of gains. This corresponds closely to the comparison between the estimated preferences in Figure OA4.3 in Online Appendix 4, where the two groups showed strong differences for payoffs around zero and higher.

The respondent's income is only weakly correlated with the frequency of FOSD violations, as Table OA5.4 shows. Columns (5) and (6) indicate that higher-income participants behave less risk-seeking in the losses domain than lower-income participants (again, in close correspondence to the model estimates – see Figure OA4.4 in Online Appendix 4), but this difference does not carry over to a significant and robust reduction in FOS-dominated choices.

Perhaps the most surprising result is that none of the variables that may proxy for analytical skills yields a significant reduction in the number of FOSD violations. Neither the ability to answer our three mathematics questions correctly, nor their general educational background and their mathematics-related background are found to be significantly correlated with the number of A and D choices — see Tables OA5.5 and OA5..6. Figures OA4.5 through OA5.7 show that this finding is consistent with the absence of differences between the respondents' risk attitudes.