How Rational Are Children?*

Vinci Chow†

10th April 2014

Abstract

This study revisit the question of whether children make rational decisions, utilitizing improved techniques and a much larger subject pool than earlier studies. I find that young children are significantly less consistent than previously reported. Senior students are considerably more consistent than junior ones, with moderate evidence that the difference is due to aging. I also confirm that students' mathematics performance is significantly correlated with consistency, and that students' age, education and mathematics performance are all significantly related with their risk preference.

JEL Classification: C93, D81, G11, I21

Keywords: GARP, consistency, rationality, education, children

1 Introduction

The notion that children make lower-quality decisions than adults is often implicitly accepted, as is evident in most societies granting only a limited set of rights to the them as well as the common requirement for a guardian. However, studies on the economic rationality of children have not provided much support for this view. Harbaugh *et al.* (2001) demonstrated experimentally that choices made by students as early as in Grade 2 are almost as consistent with utility maximization as the choice of adults. Their finding provided justification for the application of economic theory to children.

Despite its appealing results, Harbaugh et al.'s experiment has major shortcomings. First, as was common among early revealed-preference studies, the

*Earlier versions of this paper was circulated under the title "Are the Choices of Children Consistent under Risk?" I thank Stefano DellaVigna, William Harbaugh, Shachar Kariv, Mike Urbancic, Eva Vivalt as well as seminar participants at University of California, Berkeley and University of Oregon for valuable comments. I also thank the CUHK student team lead by Xue Tian for the excellent research assistance they provided. This research has been supported by Research Grants Council Direct Grant 2021115.

†Department of Economics, The Chinese University of Hong Kong (email: vincichow@cuhk.edu.hk)

experiment asked its subjects to make very few choices. Simulations by Choi *et al.* (2007a) suggest that when too few choices are being made, there is a very high chance that random behavior might appear consistent.[1] Second, the resolution of the budget sets were severely limited due to the practical need of using only whole numbers. When subjects could only choose between 3-6 bundles along each budget line, their true preference might not be well-reflected by their choice. Finally, the study only had a total of 73 elementary school subjects, a sample size too small for some of the statistical analysis the authors wanted to make.

This study revisits the question of children's decision-making quality, addressing all the aforementioned issues. Utilizing improved techniques and a considerably larger subject pool, I estimate that young children are significantly less consistent than Harbaugh *et al.* have reported. I find significant improvement in consistency between junior and senior elementary school students, and I am able to confirm what the original study could only speculate about: that students' mathematics performance is significantly correlated with consistency.

Beyond updating Harbaugh *et al.*, this study advances research in two areas. First, I investigate whether decision-making quality improves with education. Although existing studies have shown a correlation between educational attainment and consistency, self-selection issue and correlation with age have hindered a clear interpretation. Exploiting the property that students within the same grade have different birthdates, I find moderate evidence that decision-making quality improves mainly with age instead of education.

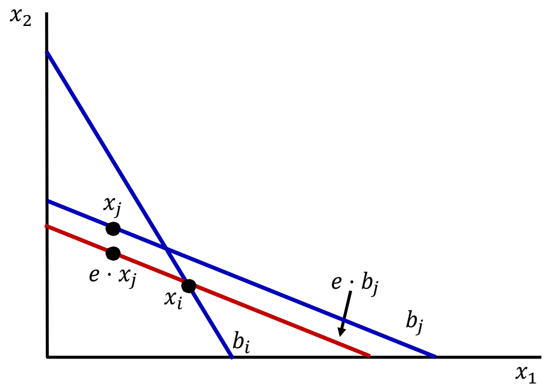
Second, I show that older students are more risk averse while more educated students are less so. While this finding is consistent with the existing literature on adults, having an age effect in this early stage of life would be difficult to account for with theories like the life-cycle hypothesis. It also suggests that the lack of an education control in earlier child development studies could have biased their results.[2]

This study is part of a growing literature that uses adherence to Generalized Axiom of Revealed Preference (GARP) as a measure of rationality. The approach has been used to study behavior ranging from household finance (Diaye *et al.* 2008, Choi *et al.* 2013) to trading behavior (Cecchi and Bulte 2013) to altruism (Andreoni and Miller 2002, Fisman *et al.* 2007), on subjects ranging from capuchin monkey (Chen *et al.* 2006) to patients with frontal lobe damage (Camille *et al.* 2011). It also contributes to a large field of research that link experimentally-measured risk preferences with demographics and cognitive

[1]Harbaugh *et al.* has seven budget sets. According to Choi *et al.* (2007a)'s simulations, with 5 budget sets there is an 80 percent chance that random behavior would result in a CCEI over 0.95, while with 10 budget sets the chance is approximately 55 percent.

[2]See Little 2006 for a review of existing research in the field of child development.

Figure 1: Violation of GARP Illustrated



This figure demonstrates a violation. x_j was cheaper than x_i when x_i was chosen, while x_i was cheaper than x_j when x_j was chosen. Intuitively, the decision maker acquired her desired bundles when they were expensive but not when they were cheap, evidence that she was making a mistake.

ability (Frederick 2005, Burks *et al.* 2009, Choi *et al.* 2013, Dohmen *et al.* 2010 and Benjamin *et al.* 2012), as well as an interdisciplinary field that seeks to understand children's risk perception (Slovic 1966, Ginsburg and Miller 1982, Hillier and Morrongiello 1998, Little 2006 and Harbaugh *et al.* 2002.).

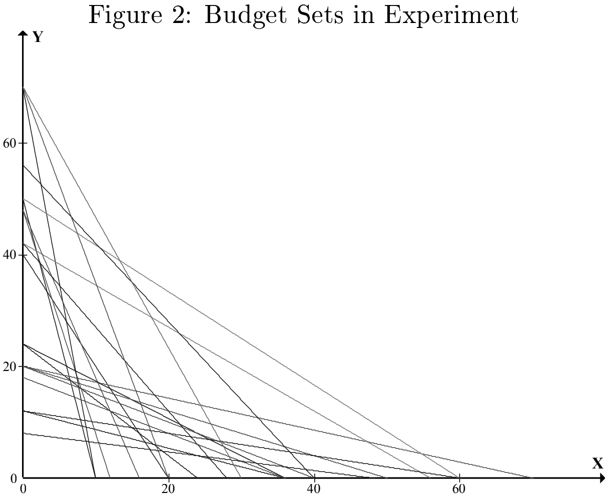
The rest of the paper is arranged as follows: section 2 explains the consistency measure, section 3 details the experimental design, section 4 and 5 discuss the findings and section 6 concludes.

2 Measuring Consistency

Following the literature, consistency is defined as adherence to GARP, which states that if x_i is revealed as preferred to x_j then $p_j \cdot x_j \leq p_j \cdot x_i$. Figure 1 demonstrates a violation. x_j was cheaper than x_i when x_i was chosen, while x_i was cheaper than x_j when x_j was chosen. Intuitively, the decision maker acquired her desired bundles when they were expensive but not when they were cheap, evidence that she was making a mistake.

The choices made by a decision marker maximizing a non-satiated utility function necessarily fulfills GARP, making adherence to GARP is an attractive measure. Furthermore, Afriat (1967) has shown that a finite set of choices that satisfy GARP can be rationalized by a utility function that is piecewise linear, continuous, increasing and concave. Adherence to GARP would provide justification for the use of utilities functions that have these normatively-appealing characteristics.

Figure 2: Budget Sets in Experiment



Because GARP is an exact test of consistency, in practice we need a way to measure the distance from perfect fulfillment of GARP. The most commonly used measure is Afriat's critical cost efficiency index (CCEI), which is defined as one minus the smallest percentage budget reduction needed to remove all violations of GARP (Afriat 1972, Varian 1991). Referring back to our example, the GARP violation can be solved by reducing the budget b_j by a ratio of $1-e$. e is CCEI.

3 Experiment Design

The experiment utilizes a 20-round budgeting task following Choi *et al.* (2007b), essentially asks subjects to allocate a budget between two Arrow-Debreu assets that pay in opposing states. I was unable to use a computerized interface as in Choi *et al.*, as the elementary school assisting the study was unable to provide sufficient computers for such an activity. Instead I use a 20-page questionnaire booklet that subjects were asked to fill out, each page representing one budgeting task.[3] The resolution of each budget set ranges from 9 to 23, averaging 13.9. Figure 2 a plot of the all budget sets.

92 Grade 2 and 88 Grade 5 students from a local elementary school participated in my experiment, conducted in the students' classrooms. All students present in the classrooms at the time of experiment were included in the experiment. The experimental task was explained to the subjects twice, once by the lead investigator in front of the whole class and once by research assistants in groups of 3-5 students. To proceed to the actual experiment, subjects needed

[3]The questionnaire and other supplementary materials can be found at: `http://ticoneva.com/econ/research/garp_children/`

to point out correctly in a trial round what would happen in each outcome of the coin toss. At the end of each session, a coin toss determined the state of the world while a dice selected a round. Each subject was rewarded by her choice in that round with M&M's chocolates. Gender, age and the subjects' most recent test scores on mathematics, language and general studies were obtained from the elementary school.

4 Consistency

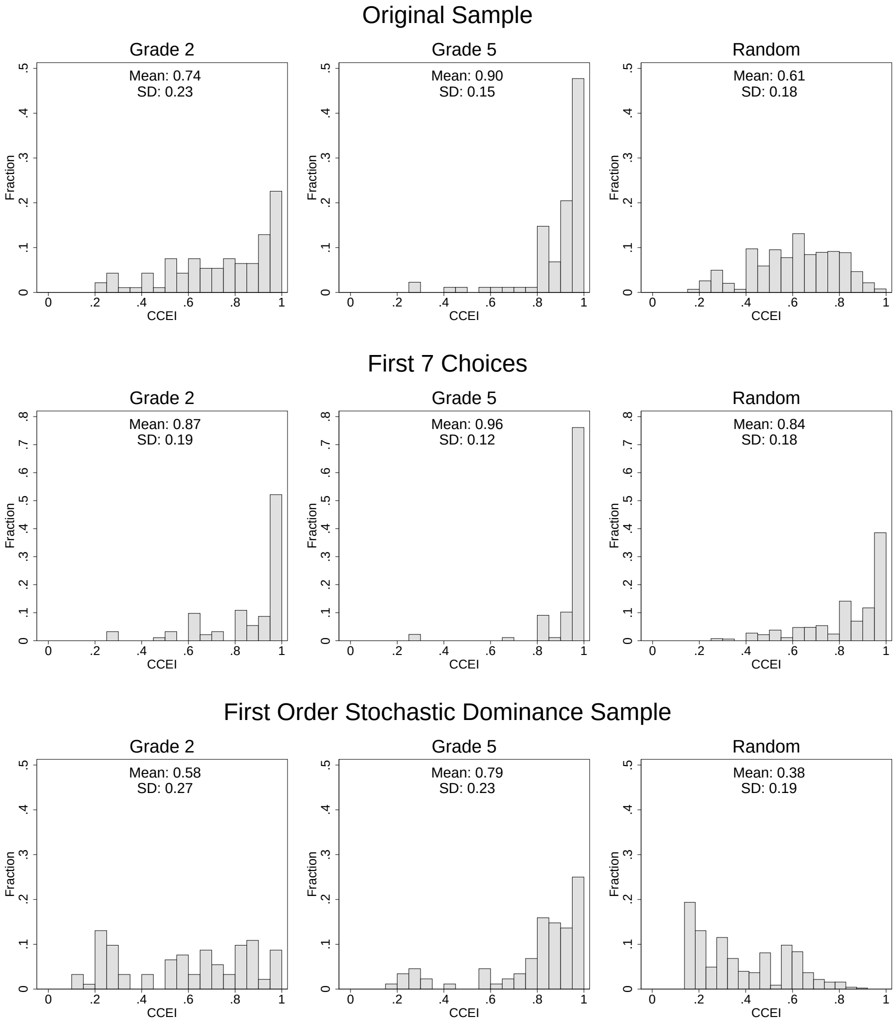
Figure 3 plots the distribution of CCEI within each grade, as well as that from a sample of 18,000 simulated subjects who made choices at random. Subjects from both grades demonstrated a degree of consistency markedly higher than what would have resulted from random chance, confirmed by Mann-Whitney tests (Grade 2 versus random: $z = 6.339$, $p = 0.000$; Grade 5 versus random: $z = 13.279$, $p = 0.000$). As a common ground for comparison, previous studies have followed Varian (1991) in taking a CCEI of 0.95 as threshold for high consistency with utility maximization. 22.6% of the subjects in Grade 2 and 45.2% of the subjects in Grade 5 passed this threshold, in contrast to the 0.8% from the simulated sample.[4]

In contrast with Harbaugh *et al.*, which reported very high CCEI across all age groups, in my experiment Grade 2 students were significantly less consistent than Grade 5 students (Mann-Whitney: $z = -5.159$, $p = 0.000$). To investigate whether this is due to experimental design or to difference in the subject pools, I recalculated each subject's CCEI using only their first seven choices, which is the number of observations Harbaugh *et al.* had for their subjects. The truncation results in a distribution that is significantly skewed towards consistency: average CCEI for Grade 2 and Grade rise by 17.6% and 6.7% respectively, reaching levels that are comparable with Harbaugh *et al.*'s reported values. Therefore the high level of consistency among children reported by Harbaugh *et al.* seems likely to be due to the small number of budget sets they employed.

Although the subjects' responses appear consistent at first glance, their preference could still be quite far from what one might consider "reasonable". For example, a subject who always select the same option—regardless of the cost—is perfectly consistent, yet few would consider such behavior normatively appeal-

[4]Simulations in previous studies have reported higher rates of passing this threshold value. Choi *et al.* (2007a) reported a rate of approximately 7 percent with 20 questions, while in Appendix 2 of their paper, Choi *et al.* (2013) indicate a rate of 2 percent with 25 questions. Choi *et al.* (2007b) reported an even higher chance of 14.3 percent when 25 questions are used. As shown in the appendix of this paper, the probability that a random choice set passes the 0.95 threshold increases with resolution of the choice set. Intuitively, a simulated agent is forced to make choices that are further apart when the resolution is low, making violations of GARP more likely.

Figure 3: CCEI Distribution



ing. To investigate whether this concern is material, a new sample is constructed by appending to the original sample the mirrored version of it, which has prices and allocations reversed across axes. Consistency under this new sample requires a preference for options that adhere to first-order stochastic dominance, ruling out, for instance, the aforementioned example as being consistent. Under this sample, CCEI dropped by 0.16 for Grade 2 and 0.11 for Grade 5, comparable to the 0.15 reported in Choi *et al.* (2013).[5] The random sample, in contrast, drops by 0.23, further proving the strategic nature of real subjects' behavior.

4.1 Effects of Education and Age

Improving decision-making quality is among the oldest reasons for education. Both Choi *et al.* (2013) and Harbaugh *et al.* have shown a correlation between educational attainment and consistency, but two obstacles hinder a clear interpretation of the correlation. For Choi *et al.*'s study on adults, the issue is whether education really improves decision-making quality, or whether individuals who make better decisions simply self-select into studying longer. For Harbaugh *et al.*'s study on children, the issue is whether the observed effect is due to aging or schooling, which are often highly correlated.

I attempt to answer this question relying on the fact that while elementary education is mandatory in Hong Kong, students within the same grade can be up to a year apart in age due to difference birthdates. Table 1 lists regressions of CCEI on subjects' level of schooling and age. Level of schooling and age both show a significant positive effect of approximately 0.05 on CCEI when regressed independently. This means that students on average waste 5 percent less wealth every year. With both measures in the regression, age has a positive effect while schooling has a negative one, but the coefficients are only jointly significant ($F = 32.04$, $p = 0.000$).

The effects on the first-order stochastic dominance sample is stronger than that on the original, with approximately 7 percent less wealth being wasted every year. The 2 percent per year additional reduction in waste over the original can be interpreted as the speed at which students converge to a normatively-appealing preference. With both age and schooling in the regression, there is moderate significance for age but not schooling, suggestive evidence that age plays a larger role than schooling in the improvement of decision-making quality.

[5]The difference in difference between Grade 2 and Grade 5 is also significant ($z = 3.033$, $p = 0.002$). Plots of the choices made by each subject can be found in the online appendix.

Table 1: CCEI, Age and Schooling

	Original Sample					
	(1)		(2)		(3)	
male	0.006		0.001		0.000	
	(0.028)		(0.028)		(0.028)	
grade	0.050	***			-0.013	
	(0.009)				(0.049)	
age			0.053	***	0.065	
			(0.009)		(0.046)	
constant	0.644	***	0.350	***	0.287	
	(0.042)		(0.087)		(0.249)	
N	180		178		178	
R-sq	0.137		0.157		0.158	

	First Order Stochastic Dominance Sample					
	(4)		(5)		(6)	
male	-0.029		-0.037		-0.039	
	(0.038)		(0.038)		(0.039)	
grade	0.069	***			-0.025	
	(0.012)				(0.061)	
age			0.073	***	0.096	*
			(0.012)		(0.058)	
constant	0.456	***	0.049		-0.074	
	(0.053)		(0.111)		(0.309)	
N	180		178		178	
R-sq	0.151		0.174		0.175	

*Significant at 10%, **Significant at 5%, ***Significant at 1%. Robust standard errors are reported within parentheses.

This table reports the relationship between CCEI, age, grade and gender, estimated using ordinary least squares. The top panel uses CCEI values estimated from the original experimental data, while the second panel uses values estimated from the first-order stochastic dominance sample, which is the original sample appended with its own mirrored version.

4.2 Academic Performance

Choi *et al.* (2013) established a link between general education level and CCEI, but was unable to examine the channel through which education interacts with consistency. Harbaugh *et al.*, with a very small of subjects, investigated whether subject's performance in mathematics is correlated with CCEI and did not find a significant effect. This paper examines the relationship utilizing a much larger subject pool and test scores in three different fields of study.

I collected the subjects' most recent test scores in three fields of study: language, mathematics and general studies. Table 2 summarizes the subjects' performance in those tests. Female students performed better in language and their mathematics performance exhibits significantly more variation than the other two fields. The correlation between students' performance in different fields is also much higher in Grade 2 than in Grade 5: correlations are all in the range of 0.7 for the former, but only 0.3-0.5 for the latter. High correlation would be expected if the various fields of study are mostly affected by the same latent ability, which is more likely to be true for the more elementary curriculum of Grade 2.

Table 2: CCEI and Test Scores by Grade and Gender

	Grade 2		**Grade 5**	
	Male	**Female**	**Male**	**Female**
CCEI	0.75	0.75	0.91	0.89
	(0.21)	(0.23)	(0.12)	(0.17)
Language Score	67.64	72.08	77.40	78.83
	(15.29)	(11.71)	(8.56)	(8.28)
Mathematics Score	71.00	71.19	81.94	80.58
	(15.41)	(16.21)	(7.97)	(11.28)
General Studies Score	86.21	87.98	80.14	79.38
	(9.69)	(7.20)	(6.48)	(6.33)
N	39	53	35	53

Table 3 lists regressions of CCEI on scores standardized within each field and grade. Mathematics score in Grade 5 is significant correlated with consistency—one standard deviation increase in mathematics score is associated with a 0.058 increase in CCEI, which translates to a 5.8 percent reduction in wealth wasted. The existence of this relationship is understandable given the numerical nature of the experimental task, and Benjamin *et al.* (2013) has similarly found that subjects with better mathematics scores exhibit fewer behavioral anomalies in various other decision-making tasks.

Decomposing the effects by gender, only female students show a clear improvement in CCEI with academic performance. Among Grade 2 female students language score correlates with consistency, suggesting that at this age language ability might be a good measure of cognitive ability. The estimated effects for males are insignificant (mathematics in Grade 5: $F = 0.62, p = 0.432$, language in Grade 2: $F = 0.01, p = 0.937$, general studies in Grade 2: $F = 3.67, p = 0.059$). This difference might explain why female subjects' CCEI has higher variance in both the current study (overall standard deviation of male: 0.188, female: 0.216) and Choi *et al.* (2013).

5 Risk Preference

Following Choi *et al.* (2013), I measure risk attitude by the fraction of payoffs allocated to the cheaper option. This non-parametric measure is negatively correlated with risk aversion—a risk-neutral subject would maximize payoffs obtained through the cheaper option, while an infinitely risk-averse subject would seek to equalize payoffs among the two options. In the experiment, Grade 2 subjects allocated on average 67.8 percent of their payoffs to the cheaper option, while Grade 5 students allocated 68.7 percent. The difference is not statistically significant ($t = -0.528$, $p = 0.598$).

Table 4 regresses the measure on observed attributes. Older subjects were more risk averse while more educated subjects were less so, with the two effects effectively canceling each other across cohort ($F = 0.29$, $p = 0.591$). Male subjects were also more risk-seeking.

While the literature has consistently found similar relationships, early empirical studies considered the age effect evidence in support of the life-cycle hypothesis (Morin and Suarez 1983 and Bakshi and Chen 1994). Recent experimental studies have shown, however, that the relationship is present even for small-stake gambles in a laboratory setting (Dohmen *et al.* 2010 and Choi *et al.* 2013). Table 3's grade-specific regressions show that the age effect is present within Grade 2 but not within Grade 5, a phenomenon hard to explain by the life-cycle hypothesis. One possible explanation is that children become more risk averse as they experience realizations of risk, and the effect is stronger when they have less experience. This hypothesis is supported by the recent studies of Gallagher (2013) and Song (2013), which present evidence that exposure to realizations of risk increases risk aversion, as well as by the finding in child development research that younger children identifies risk factors slower than older ones (Hillier and Morrongiello 1998).

Subjects with high mathematics score were more risk-seeking, with a 1 standard deviation increase in mathematics score associated with a 2.5 percent increase in the fraction of payoffs allocated to the cheaper option. A relationship

Table 3: CCEI and Academic Performance

	Grade 2				Grade 5			
	(1)		(2)		(3)		(4)	
language score (std.)	0.038		0.106	**	-0.017		-0.032	
	(0.036)		(0.049)		(0.018)		(0.035)	
math score (std.)	-0.045		-0.048		0.058	***	0.086	***
	(0.036)		(0.038)		(0.019)		(0.032)	
general studies score (std.)	0.035		-0.027		-0.014		-0.044	
	(0.034)		(0.046)		(0.018)		(0.029)	
male			0.021				0.004	
			(0.048)				(0.029)	
male * language			-0.109	*			0.028	
			(0.064)				(0.041)	
male * math			-0.020				-0.066	*
			(0.082)				(0.039)	
male * general studies			0.120	*			0.053	
			(0.067)				(0.037)	
age			0.044				0.023	
			0.072				0.052	
constant	0.742	***	0.402		0.898	***	0.655	
	(0.023)		(0.533)		(0.016)		(0.556)	
N	93		90		88		88	
R-sq	0.040		0.096		0.105		0.151	

*Significant at 10%, **Significant at 5%, ***Significant at 1%. Robust standard errors are reported within parentheses.
This table reports the relationship between CCEI and academic performance, estimated using ordinary least squares. All test scores are standardized within subject and grade to have a mean 0 and a standard deviation of 1.

Table 4: Risk Preference

	All		Grade 2		Grade 5	
language score (std.)	-0.016		0		-0.029	**
	(0.011)		(0.019)		(0.013)	
math score (std.)	0.025	**	-0.003		0.043	***
	(0.012)		(0.019)		(0.014)	
general studies score (std.)	0.001		0.029		-0.024	*
	(0.011)		(0.018)		(0.014)	
male	0.036	**	0.051	*	0.027	
	(0.018)		(0.026)		(0.024)	
age	-0.063	**	-0.096	**	-0.016	
	(0.028)		(0.039)		(0.039)	
grade	0.066	**				
	(0.028)					
constant	0.997	***	1.374	***	0.846	**
	(0.151)		(0.29)		(0.403)	
N	178		90		88	
R-sq	0.077		0.116		0.146	

*Significant at 10%, **Significant at 5%, ***Significant at 1%. Robust standard errors are reported within parentheses.

This table reports the relationship between the fraction of payoff allocated to the cheaper option, students' demographics and academic performance, as estimated by ordinary least squares. All test scores are standardized within subject and grade to have a mean 0 and a standard deviation of 1.

between cognitive ability and risk preference has been noted in existing research, though none of those studies have subjects as young (see for example Frederick 2005, Burks *et al.* 2009, Dohmen *et al.* 2010 and Benjamin *et al.* 2012). The results here emphasize that age and education level are important factors: with the point estimate for Grade 2 being essentially zero, the observed effect is entirely due to the difference among Grade 5 students.

Language scores—and to a lesser extent general studies scores—are negatively correlated with risk seeking for Grade 5 students. The existing literature has very little to say on this, as most studies do not include a separate measure of verbal ability in their analysis. The exception is Frederick (2005), which reports a negative but insignificant effect for SAT verbal score. Because females tend to perform better in such tests—the correlation between performance and gender 0.1322 in this study—it is possible that part of the gender effect commonly found in existing studies is in fact a verbal ability effect.

In a companion study available online, I demonstrate further how structurally-estimated risk parameters are correlated with age, academic performance and parenting.

6 Conclusion

In conclusion, how rational are children? This study suggests that children are highly consistent in their choices, though not as much as previously reported for younger ones: not only were the Grade 2 students in the study significantly less consistent than the Grade 5 students, they were also significantly more likely to make choices that violate stochastic dominance. While the evidence is only suggestive, estimations further suggest that the observed difference is due to aging instead of education.

For Grade 5 students, who have been shown to be as highly consistent as adults in both Harbaugh *et al.* and this study, consistency is related to mathematics performance. Students good at mathematics are also less risk-averse, consistent with existing studies on adults. Studies so far have overlooked the relationship between verbal ability and risk preference, but this study finds a significantly negative correlation between the two for Grade 5 students.

Finally, this study demonstrates the importance of controlling for both age and education in child research. The combined effect of the two factors are confounded, and once separated they appear to have opposing effects on consistency and risk preference. Due to the high correlation between the two factors in developed regions, studies that do not control for both of them—including most of the child development studies closest to this study—could produce biased results.

References

1. Afriat, Sidney N. "The Construction of a Utility Function from Expenditure Data." International Economic Review. 1967, 8(10), pp. 67-77.

2. Afriat, Sidney N. "Efficiency Estimation of Production Function." International Economic Review. 1972, 13(3), pp. 568-598.

3. Andreoni, James and John Miller. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." Econometrica. Mar. 2002, 70(2), pp. 737-753.

4. Benjamin, Daniel J., Sebastian A. Brown and Jesse M. Shapiro. "Who is Behavioral? Cognitive Ability and Anomalous Preference." Journal of European Economic Association. 2013.

5. Bakshi, Gurdip S. and Zhiwu Chen. "Baby Boom, Population Aging, and Capital Markets." Journal of Business, 1994, 67(2).

6. Barksy, Robert B., F. Thomas Juster, Miles S. Kimball and Matthew D. Shapiro. "Preferences Parameters and Behavioral Heterogeneity: An Experimental Approach in the health and Retirement Study." Quarterly Journal of Economics. May 1997.

7. Burks, V. Stephen, Jeffrey P. Carpenter, Lorenz Goette and Aldo Rustichini. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." Proceedings of the National Academy of Sciences. 2009.

8. Camille, Nathalie, Cathryn A. Griffiths, Khoi Vo, Lesley K. Fellows and Joseph W. Kable. "Ventromedial Frontal Lobe Damage Disrupts Value Maximization in Humans." The Journal of Neuroscience. May 2001, 31(20), pp. 7527-7532.

9. Cecchi, Francesco and Erwin Bulte. "Does Market Experience Promote Rational Choice? Experimental Evidence from Rural Ethiopia." Economic Development and Cultural Change. Jan. 2013, 61(2), pp. 407-429.

10. Chen, M. Keith. "How Basic Are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior." Journal of Political Economy. Jun. 2006, 114(3), pp. 517-537.

11. Choi, Syngjoo, Ray Fisman, Douglas Gale and Shachar Kariv. "Revealing Preferences Graphically: An Old Method Gets a New Tool Kit." American Economic Review, Papers and Proceedings. May 2007, 97(2), pp. 153-158.

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

13. Choi, Syngjoo, Shachar Kariv, Wieland Muller and Dan Silverman. "Who is (more) Rational?" Forthcoming in American Economic Review. 2013.

14. Diaye, marc-Arthur, François Gardes and Christophe Starzec. "GARP Violation, Economic Environment Distortions and Shadow prices: Evidfence from Household Expenditure Panel Data." Annals of Econmoics and Statistics. Jun. 2008, 90, pp. 3-33.

15. Dohmen, Thomas, Armin Falk, David Huffman and Uwe Sunde. "Are Risk Aversion and Impatience Related to Cognitive Ability?" American Economic Review. Jun. 2010, 100, pp. 1238-1260.

16. Fisman, Raymondm Shachar Kariv and Daniel Markovits. "Individual Preferences for Giving." American Economics Review. Dec. 2007, 97(5), pp. 1858-1876.

17. Frederick, Shane. "Cognitive Reflection and Decision Making." Journal of Economic Perspectives. Fall 2005, 19(4), pp. 25-42.

18. Gallagher, Justin. "Learning about an Infrequent Event: Evidence from Flood Insuraqnce Take-up in the US." Forthcoming in American Economic Journal: Applied Economics. 2013.

19. Ginsbury, Harvey J. and Shirley M. Miller. "Sex Differences in Children's Risk-Taking Behavior." Child Development. Apr. 1982, 53(2), pp. 426-428.

20. Gul, Frank. "A Theory of Disappointment Aversion." Econometrica, 59(3), pp. 667-86.

21. Harbaugh, William T., Kate Kraus and Tim Berry. "GARP for Kids: On the Development of Rational Choice Behavior." American Economic Review. Dec 2001, 91(5), pp. 1539-1545.

22. Harbaugh, William T., Kate Kraus and Vesterlund. "Small and Large Probability Gains and Losses." Experimental Economics. 2002, 5, pp. 53-84.

23. Hillier, M. Loretta and Barbara A. Morrongiello. "Age and Gender Differences in School-Age Children's Appraisals of Injury Risk." Journal of Pediatric Psychology. 1998, 23(4), pp. 229-238.

24. Holt, Charles A. and Susan K. Laury. “Risk Aversion and Incentive Effects.” American Economic Review. Dec. 2992, 92(5), pp. 1644-1655.

25. Little, Helen. “Children’s Risk-Taking Behavior: Implications for Early Childhood Policy and Practice.” International Journal of Early Years Education. 2006, 14(2), pp. 141-154.

26. Morin, Roger-A. and A. Fernandez Suarez. “Risk Aversion Revisited.” The Journal of Finance. Sept 1983, 38(4), pp.1201-1216.

27. Slovic, Paul. “Risk-Taking in Children: Age and Sex Differences.” Child Development. Mar. 1966, 37(1), pp. 169-176.

28. Song, Changcheng. “Financial Illiteracy and Pension Contributions: A Field Experiment on Compound Interest in China.” Working Paper. 2013.

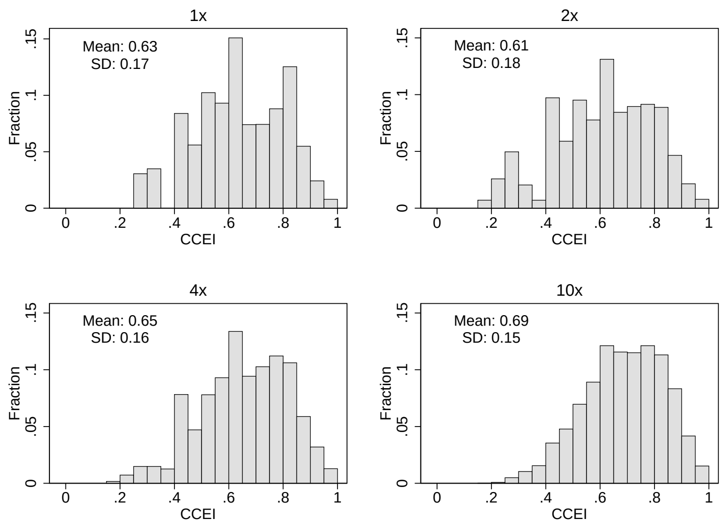
29. Varian, Hal R. “Goodness-of-Fit” for Revealed Preference Tests. Unpublished. 1991.

Appendices

A Resolution and Consistency

To investigate the effect budget-line resolution have on the level of consistency obtainable randomly, I increase resolution of the budget sets used in the experiment by a constant fraction, then regenerate 18,000 new simulated subjects. Figure 4 demonstrates the CCEI distributions for 1x, 2x, 4x and 10x the original resolution. The randomly generated choices clearly cluster towards being more consistent as resolution increases (Mann-Whitney: $1x$ vs. $2x$ $z = 12.32$, $2x$ vs. $4x$ $z = -23.05$, $4x$ vs. $10x$ $z = -18.55$, all $p = 0.000$). OLS regression suggests that doubling resolution would result in CCEI increasing by 0.0074 on average ($t = 44.62$, $p = 0.000$).

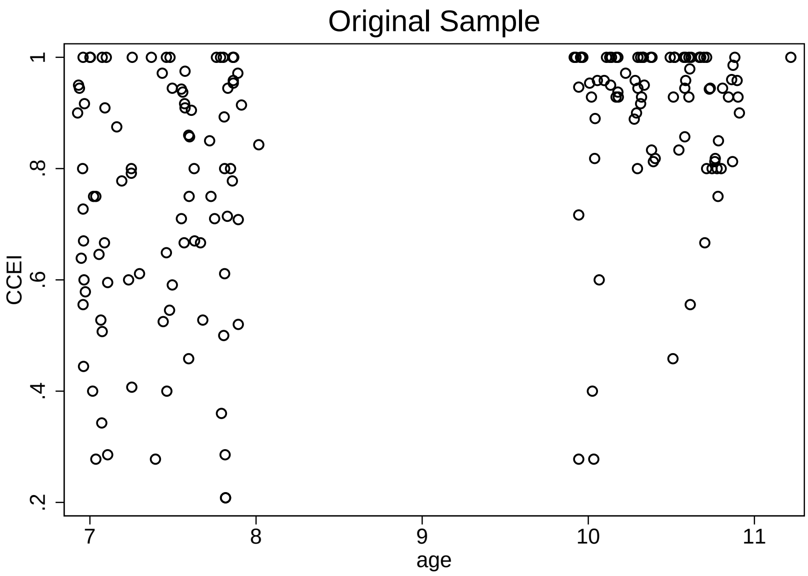
Figure 4: Resolution of Budget Sets and CCEI Distribution



This figure plots the distribution of CCEI simulated under 1x, 2x, 4x and 10x the resolution used in the actual experiment. 18,000 simulated subjects were generated for each resolution.

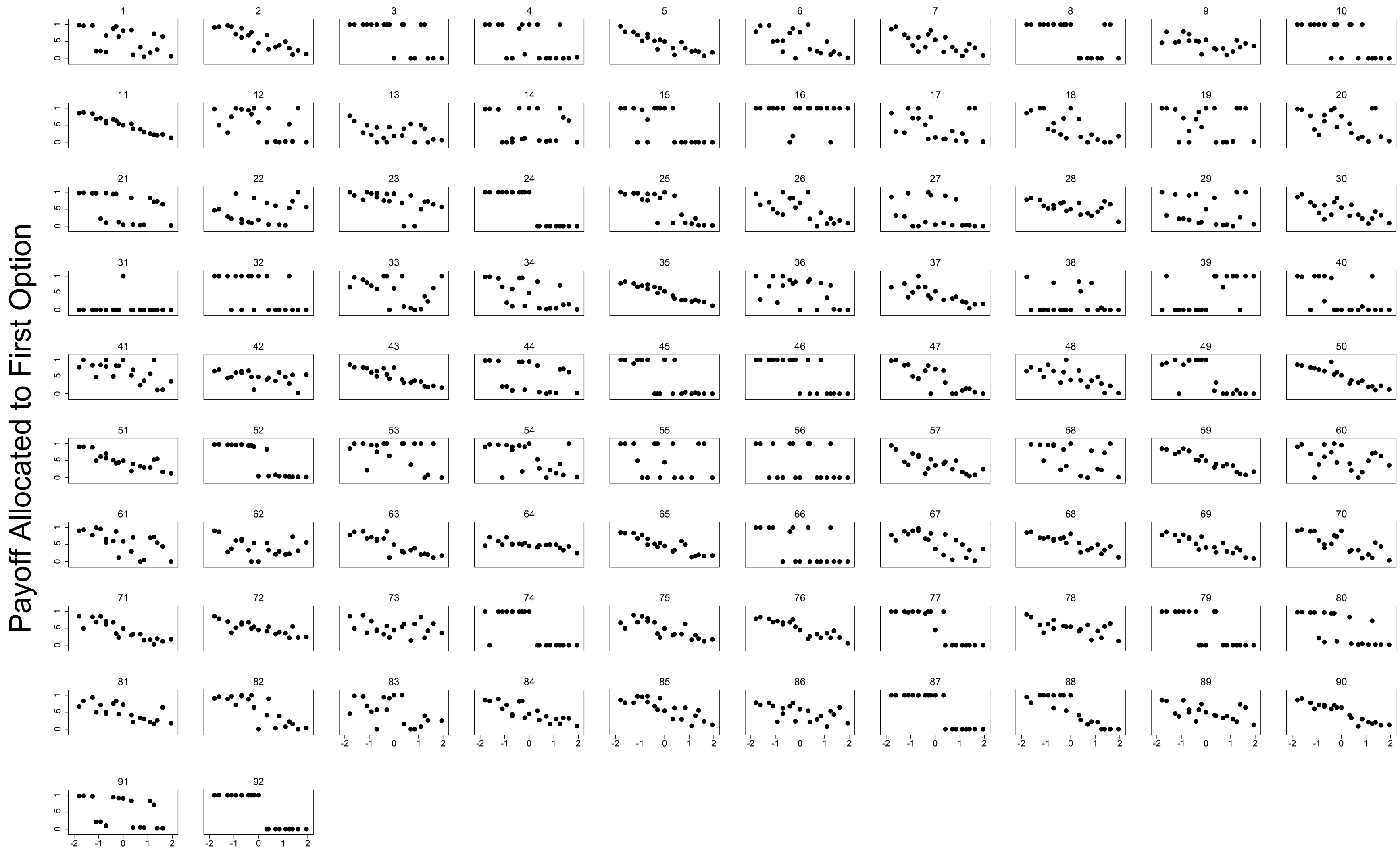
B Addition Tables and Diagrams - To be Published Online

Figure 5: CCEI Scatter Plot



Individual Choice Plots

Grade 2



Individual Choice Plots

Grade 5