



Sources of Lower Financial Decision-making Ability at Older Ages

Shachar Kariv and Dan Silverman



Sources of Lower Financial Decision-making Ability at Older Ages

Shachar Kariv

University of California-Berkeley

Dan Silverman

Arizona State University

November 2015

Michigan Retirement Research Center

University of Michigan

P.O. Box 1248

Ann Arbor, MI 48104

www.mrrc.isr.umich.edu

(734) 615-0422

Acknowledgements

The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium through the University of Michigan Retirement Research Center (5 RRC08098401-07). The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA or any agency of the Federal Government. Neither the United States Government or any agency thereof, or any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof.

Regents of the University of Michigan

Michael J. Behm, Grand Blanc; Mark J. Bernstein, Ann Arbor; Laurence B. Deitch, Bloomfield Hills; Shauna Ryder Diggs, Grosse Pointe; Denise Ilitch, Bingham Farms; Andrea Fischer Newman, Ann Arbor; Andrew C. Richner, Grosse Pointe Park; Katherine E. White, Ann Arbor; Mark S. Schlissel, *ex officio*

Sources of Lower Financial Decision-making Ability at Older Ages

Abstract

After middle age, further aging is associated with lower levels of many cognitive abilities, some of which could influence important economic decisions. Our prior research (Choi et al., 2014) shows a substantial negative relationship between age and the consistency of choices with economic rationality (decisionmaking quality). This paper investigates the sources of that negative correlation using panel data on more than 4,000 members of a panel study in the Netherlands. The analysis finds no evidence that the correlation between age and rationality is, in fact, a just a cohort, not an age effect. Similarly, there is little evidence that the correlation is due to other forms of cognitive or health declines. Rather, the findings indicate that age has an independent and negative effect on economic rationality.

Citation

Kariv, Shachar and Dan Silverman. 2015. "Sources of Lower Financial Decisionmaking Ability at Older Ages." Ann Arbor, MI. University of Michigan Retirement Research Center (MRRRC) Working Paper, WP 2015-335.
<http://www.mrrc.isr.umich.edu/publications/papers/pdf/wp335.pdf>

Authors' acknowledgements

We are particularly grateful to Syngjoo Choi and Wieland Müller for many thoughtful discussions and for their fundamental contributions to the data collection and analysis. This research is supported by a grant from the Social Security Administration via the Michigan Retirement Research Center.

1 Introduction

After middle age, further aging is associated with lower levels of many cognitive abilities. Some forms of memory, for example, are lower in older people (Rohwedder and Willis, 2010). More general measures of intelligence also tend to be lower in people beyond middle age (McCardle et al., 2002). These reduced levels of skill could have important economic consequences for older people if they influence spending, saving, insurance, or health choices. Consistent with this view, economists have found a negative relationship between age and several measures of economic decision-making ability. See, e.g., Willis (2007), Lusardi and Mitchell (2007), Heiss et al. (2010), McFadden et al. (2008), Fang et al. (2008), Argawal et al. (2009), Abaluck and Gruber (2011).

Our own research with Syngjoo Choi and Wieland Müller (Choi et al., 2014) reveals, in particular, a substantial negative relationship between age and the consistency of choices with economic rationality. The results of that study show that, in a simple economic choice experiment, people over age 55 have significantly lower levels of rationality than younger people. Specifically, older people's choices are much more likely to violate transitivity; in a simple economic environment they more often reveal a strict preference for good x over good y and also a preference for good y over good x .

Classical economic theory implies that choices that violate transitivity in this way are inconsistent with any stable objective. Because choices that violate this standard of rationality are inconsistent with any fixed goal, our approach treats them as lower quality than rational choices. This view is supported by more than normative theories of choice. We interpret rationality in the experiment as a measure of decision-making ability in part because Choi et al. (2014) also finds a large, statistically significant correlation between rationality and wealth that is robust to conditioning on correlates of preferences, constraints, information, and beliefs.

In this paper we investigate sources of the negative relationship between age and financial decision-making ability. We evaluate three hypotheses:

1. The negative relationship is not due to an underlying relationship between age and rationality, but is instead a cohort effect. Given the many changes in technology, the tasks required of people in economic experiments, especially those implemented through web-based surveys, are relatively unfamiliar to those born before widespread use of computers and the Internet. Upon reaching their mid-fifties, however, younger cohorts will not face the same difficulty and will appear just

as rational as they did in their youth.

2. The correlation between age and rationality results from normal aging processes and can be accounted for by declines in other cognitive functions. In this view, age leads to a reduction in economic decisionmaking ability but this particular form of cognitive decline is not meaningfully distinct from other manifestations of it. In this view, the correlation in Choi et al. (2014) is indicative of a more fundamental cognitive decline well-captured by standard measures of cognitive ability.
3. The average decline in financial decision-making ability with age is driven by those who have had important health problems, independent of losses in other cognitive functions. This theory of the negative correlation between rationality and age sees the decline in rationality not as an unavoidable consequence of age but as the result of illnesses and conditions that disproportionately affect older people.

The quantitative importance of each of these sources of the relationship between age and financial decision-making ability has implications for policy and research. If older people make lower quality decisions because they face unfamiliar economic environments, there is potential for them to learn and adapt, and to benefit from financial education. If the relationship is driven by cognitive decline associated with normal aging, and is not a separate phenomenon from other forms of cognitive decline with age, then there is no need to make special accommodation for older people in financial markets, and finance-related government policies. If instead there is something special about financial decision making ability in older ages, policy-makers and market participants should focus resources on assuring older people have the same effective opportunities as younger populations. Finally, if sharp declines in health drive the relationship between financial decision making quality and age, then accommodations for older populations can be better targeted at those in poor health and the consequences of health care for well-being would further underscored.

To assess the relative importance of these three sources of lower decision-making ability among older people, we analyze a replication and extension of the experiment in Choi et al. (2014), implemented with a sample about whose health and cognitive function we know much more. Specifically, we study an experiment with more than 4,000 members of the Longitudinal Internet Studies for the Social Sciences (LISS) panel, a representative sample of over 5,000 Dutch households.

Like the previous experiment, the new experiment presented each subject with 25 problems. Each problem is to choose, from a budget line, an allocation of points between accounts x and y . The subject's payoff from a choice is the points he allocated to one of the accounts, x or y , determined at random and equally likely. After completing 25 problems, one problem is selected at random to be paid. Payments are modest; on average a participant would earn €30 in addition to their normal participation payment.

Different from the survey used in Choi et al. (2014), however, the LISS contains relatively detailed health information. This information, includes self-reported health, and more objective measures ability at activities of daily living. To this we added questions that ask respondents to implement particular (target) allocations from budget sets. These questions are intended to test a respondent's ability to use a computer to navigate this choice environment. We also added classic tests of cognitive ability, Raven's Matrices. These tests ask subjects to identify, from a list, the missing piece of a patterned geometric object. Success requires subjects to identify the pattern and the geometry of the object to see what is missing.

The main results are as follows:

1. The LISS data reveal, after late middle age, an economically substantial and statistically significant negative correlation between age and measures of economic rationality in the experiment. After conditioning on several socio-demographic characteristics, rationality scores are significantly lower for those age 63 and older.
2. There is no evidence that the negative relationship between age and economic rationality in the experiment is attributable to a cohort effect. Conditioning on the ability of participants to implement a particular choice using the experimental interface does not alter the negative correlation between rationality and age. Older people are more likely to have trouble with the interface in this way, but this ability is not strongly correlated with violations of transitivity in the experiment.
3. There is also no strong evidence that the correlation between age and economic rationality is the inevitable result of normal aging and its associated cognitive declines. Performance on the Raven's matrices task is lower among older people in the sample, but conditioning on this measure of cognitive ability does not much alter the negative correlation between rationality and age. Thus, the lower levels of economic decisionmaking ability among older people appears to be a distinct phenomenon from general cognitive decline.

4. There is no evidence that health is an important driver of the negative relationship between age and economic rationality. Older people who self-report worse health, are at an unhealthy weight, or report more difficulties with activities of daily living do not have lower decision-making abilities in the experiment.

The results of the study thus support the idea that age leads to declines in economic decisionmaking ability that are distinct from other forms of cognitive decline and are not primarily a consequence of declining health more generally.

The remainder of this paper proceeds as follows. In Section 2 we describe the experiment in greater detail and explain how we measure economic rationality in the experiment and why we interpret it as a measure of decisionmaking making quality. In Section 3 we discuss the panel data and experimental setting. Section 4 presents the basic results, demonstrating a correlation in these new data between age and economic rationality. Section 5 presents the results of the investigation of what drives the correlation, and Section 6 offers preliminary conclusions.

2 The Experiment

The experiment studied here was fielded with 4,184 participants in the LISS panel in the spring of 2014. It presented each participant with 25 problems. Each problem asks the participant to make a decision under uncertainty. A natural interpretation of a problem is that there are states of nature denoted by $s = 1, 2$ and two associated securities, each of which promises a payoff in one state and nothing in the other. The problem is to allocate an endowment of money between the two securities. In each problem the states are equally likely. Let x_s denote the demand for the security that pays off in state s and let p_s denote its price. If we normalize the individual's wealth to 1, then the budget set is $p_1x_1 + p_2x_2 = 1$ and the individual can choose any portfolio $(x_1, x_2) \geq 0$ that satisfies this constraint.

Each decision problem started with the computer selecting a budget line randomly from the set of budget lines that intersect with at least one of the axes at 50 or more points, but with no intercept exceeding 100 points. The participant's task is to use the computer's mouse or trackpad to point and click on his or her preferred portfolio. The initial location of the computer's cursor was also selected at random from a uniform distribution over the budget line. See Figure 1 for an illustration of the interface that participants faced.

During the course of the experiment, subjects were not provided information about the account that had been selected in each round. At the end of the experiment, the computer selected one round for each subject, where each round had an equal probability of being chosen, and the subject was paid the amount he had earned in that round. Payoffs were calculated in terms of points and then converted into euros. As noted in the introduction, payments were modest; on average a participant would earn €30 in addition to their normal participation payment. Participants received their payment from the LISS reimbursement system via direct deposit into a bank account.

2.1 Testing Rationality

Following Afriat’s (1967) theorem, we employ the Generalized Axiom of Revealed Preference (GARP) to test whether the set of observed price and quantity data that our experiment generated for each participant may be rationalized by a utility function.¹ The large range of budget sets faced by each participant provides a rigorous test of GARP. In particular, the changes in endowments and prices are such that budget lines cross frequently. This means that the data lead to high power tests of revealed preference conditions. See the Appendix for an evaluation of the test’s power.

To assess how nearly the data complies with GARP, we calculate both Afriat’s (1972) Critical Cost Efficiency Index (CCEI) and the Money Pump Index (MPI) due to Echenique et al. (2011). The CCEI measures the fraction by which each budget constraint must be shifted in order to remove all violations of GARP. By definition, the CCEI is between 0 and 1: indices closer to 1 mean the data are closer to perfect consistency with GARP and hence to perfect consistency with utility maximization. See the Appendix, Figure A1, for an illustration.

Echenique et al.’s (2011) MPI is related to the CCEI and is based on the classic idea that violations of GARP can be exploited as a “money pump.” The measure calculates the average or median gains available to an “arbitrageur” who could purchase a bundle that the individual declined at lower relative prices and profitably trade that bundle for another that the individual revealed he would prefer at higher relative prices. See the Appendix for a more formal discussion.

¹See the Appendix for a formal presentation of the theorem in Afriat (1967). Varian (1982, 1983) provides precise details on testing for consistency with GARP. The papers by Afriat (2012), Diewert (2012), Varian (2012) and Vermeulen (2012), published in a special volume of the *Economic Journal* on the Foundations of Revealed Preference provide an excellent overview and a discussion of some recent developments in the literature.

2.2 Interpreting Rationality as Decisionmaking Quality

Traditional economic analysis assumes that all choices are rational. In this standard view, heterogeneity in choices is attributed to heterogeneity in preferences, information, beliefs, or constraints. As noted in the introduction, a relatively recent empirical literature considers heterogeneity in choices driven, instead, by differences in the *quality* of decision-making.² In this view, there are potentially important wedges between the choices that some people actually make and the choices they would make if they had more skills or time. Indeed, in many cases, the empirical evidence strongly suggests differences in the quality of decision-making. Typically, however, definitive judgment about decision-making quality is made difficult by twin problems of *identification* and *measurement*.

The identification problem is to distinguish differences in decisionmaking quality from unobserved differences in preferences, information, beliefs or constraints. This is important because distinguishing decisionmaking quality from other sources of heterogeneity in choices is critical for policy design. Positive predictions, welfare conclusions and thus (constrained) optimal policy depend critically on the sources of any systematic differences in choices. The measurement problem is to define and implement a portable, practical, autonomous, quantifiable, and economically interpretable measure of decision-making quality.

Here we define aspects of decisionmaking quality by the compliance of choices with economic rationality. More specifically, we call choices high quality if there exists a well-defined (utility) function that the choices maximize (and choices are consistent with maximizing a utility function if and only if they satisfy GARP). We view this approach as complementing those based on one of the many tests of cognitive ability (IQ). That said, consistency with GARP offers a theoretically disciplined metric for quality of economic decisions. The measure has a well-established economic interpretation that permits, among other things, the disentangling of quality from preferences. And, classical theory tells us whether we have enough data to make it statistically useful. There is no comparable, theoretically disciplined, means using, interpreting, and evaluating an IQ test. Another advantage of the GARP test over IQ tests is that the former is easily portable to a

²This literature shows survey evidence of very low levels of basic financial knowledge, a decline in that knowledge and in more basic cognitive functioning with age, and evidence of decisions that clearly leave “money on the table.” Examples from this area of research include Willis (2007), Lusardi and Mitchell (2007), Heiss et al. (2007), McFadden et al. (2008), Fang et al. (2008), Fang et al. (2010), Argawal et al. (2009), and Abaluck and Gruber (2009) and Alshanteety et al. (2010).

variety of choice problems. We can thus make domain-specific predictions and study a comparable measure of decision-making quality across domains.

3 The LISS Data

The LISS is administered by CentERdata is a research institute in the Tilburg School of Economics and Management (TiSEM) in the Netherlands. CentERdata specializes in online surveys and manages several panels, each consisting of more than 2,000 households, that are representative of the Dutch-speaking population in the Netherlands. Via the Internet, CentERdata (www.centerdata.nl) can implement sophisticated experiments and collects a great deal of individual demographic and economic information from its panels.

The experiment described here was conducted with 4,160 members of the LISS panel about whom basic demographic information is non-missing. Summary statistics of this analysis sample are provided in Table 1. The average age of the sample is about 48 years with 22% age 64 and older. The sample also includes a large fraction of younger people, with 12% ages 16 to 23. Females represent a larger proportion (53%) than males. Approximately half the sample is currently employed, while 18% is retired and 12% is still in school. About 50% of the sample left school after the equivalent of an advanced high school (either academic or vocational) and just 10% has a university degree. The remainder of the sample finished some form of higher education, either academic or vocational. Personal income varies widely in the sample with a median monthly income of about €1,800 but a mean of €3,092. The top 5% of the sample has income in excess of 5,000.

4 Decisionmaking Quality: Differences by Age

Statistics of decisionmaking quality, as measured by consistency with GARP in the experiment, are presented in Tables 2a and 2b. The mean level of the CCEI score is 0.87 and the median is 0.91. The MPI score, which is also on a scale of 0 to 1 has a mean of 0.88 and a median of 0.96. Regardless of the measure, at least a quarter of the sample's choices are, in effect, perfectly consistent with GARP. This level of decisionmaking quality is similar to what was observed in the CentERpanel (Choi et al., 2014) where the average CCEI score was 0.88, the median was 0.93 and, again, at least 25% were perfectly consistent with GARP.

Our focus is on any differences in decisionmaking quality by age. Tables 2a and 2b presents simple averages by age. Table 3 presents the results of a regressions of decisionmaking quality on age and several other socio-economic characteristics. While a comparison of simple means (Tables 2a, 2b, or columns 1 and 3 of Table 3) indicates only modest differences in decisionmaking quality by age. After conditioning on gender, education, family structure, employment status, and income, however, more significant differences emerge.³ In particular, consistency with GARP is significantly lower among those age 64 and older than among younger members of the sample. Using the CCEI scores as the measure of economic rationality, the conditional measures are lower by 0.038 to 0.025, or between 27% to 18% of a standard deviation for the entire sample. When we use the MPI as the measure of rationality, the relationship with age is similar in magnitude, though the estimates are more precise. Based on the MPI, the scores of those older than 64 are 0.050 to 0.034 points lower, representing between 30% to 20% of a standard deviation for the entire sample.

These conditional correlations between age and economic rationality in the experiment are qualitatively similar to those in Choi et al. (2014). In that study, OLS estimates of the conditional correlation between age and the CCEI indicate that those older than 64 had scores 0.051 points lower than younger participants. Different from the findings in the LISS, the Choi et al (2014) findings indicate that the lower rationality scores start among people as young as the mid 50s.

5 Sources of Lower Decisionmaking Quality at Older Ages

The preceding analysis shows that older people reveal significantly lower levels of economic rationality in the experiment, conditional on several observable characteristics. In what follows, we explore alternative explanations for the lower rationality scores of older people.

³Interestingly, while many of these observable characteristics are significantly correlated with economic rationality in the experiment, the only variable that qualitatively effects the (conditional) correlation between age and rationality is an indicator for whether the participant is retired.

5.1 A Cohort Effect

One reason older people may exhibit lower rationality in the experiment is because the task requires facility with a computer interface that is relatively unfamiliar to those born before widespread use of computers and the Internet. If so, there is little reason for concern that economic rationality declines importantly in old age. Instead, the results described above would reflect a more general problem of adaptation to new technologies. Moreover, we would expect that, upon reaching their mid-sixties, younger cohorts will appear just as rational (on this task) as they did in their youth.

While this is a logical hypothesis, there are reasons to be skeptical that this is a driving force behind the relationship between age and rationality in the experiment. First, there many other sources of evidence that economic decisionmaking quality is lower among older people (e.g., Agarwal et al, 2009) that do not rely on computerized elicitation methods. Second, members of the LISS panel have been working with computer interfaces, as part of the survey, for months. Thus, even if older people generally have less facility with computers, this particular group should be relatively able.

Nevertheless, it is useful to explore to what extent struggles with the interface can explain the relationship between economic rationality in the experiment and age. We therefore added one problem at the end of the sequence of 25 which simply asked participants to choose an allocation that paid 20 points to the x account. Participants were told that of they performed this task correctly (with a 1-point margin for error), they would receive €0.50.

Just 61% of the analysis sample chose to complete the task. Of these, 84% performed the task correctly. Among those who missed the target by more than one point, the average amount by which they missed was 20 points.

The tendency to miss the target is strongly correlated with age. Those younger than 23 miss the target 11% of the time while those age 51 to 63 miss it, on average, about 18% of the time. Those older than 72 missed the target 26% of the time. As shown in Table 2, however, this greater tendency of older people to miss the target does not, however, explain their higher rates of inconsistency in the experiment.

For reference, the first two columns of Table 4 present the OLS estimates of the relationship between age and economic rationality (CCEI and MPI, respectively) in the experiment, conditional on a number of socio-demographic characteristics. This is just a reproduction of columns 2 and 4 from Table 3. As described above, the results indicate that those older

than 64 have exhibit lower levels of economic rationality. Columns 3 and 4 replicate these estimates for the those (2,374) in the analysis sample that also completed the target task. The coefficients on the dummies for age 64 to 71 and age over 72 are qualitatively unchanged in this subsample. Given the smaller sample size, however, the standard errors associated with these estimates increase, and p-values rise a bit.

Columns 5 and 6 present the specifications of interest. Here we add both an indicator for whether the participant missed the target, and a measure of the amount (in points) by which he or she missed the target. Column 5 presents the results for the CCEI and Column 6 for the MPI. There is some modest evidence that the amount by which participants missed the target is related to their level of rationality in the choice task. But in neither of these specifications does conditioning on the ability of participants to implement a particular choice using the experimental interface substantially alter the negative correlation between rationality and age.

These findings indicate that while older people are more likely to have trouble with the interface as evidenced by an inability to (or lack of willingness to) implement at particular allocation, this ability is not strongly correlated with violations of transitivity in the experiment. Thus, to the extent that difficulties with the experimental task interfered with implementation of preferred bundles, it was not evident in violations of GARP and it does not explain the lower levels of consistency among older people.

5.2 The Consequences of Normal Aging

An alternative source of the correlation between age and rationality is the cognitive decline associated with normal aging. In this view, age reduces in economic decisionmaking ability because age causes declines in cognitive abilities that influence economic choices. If so, then measures of cognitive ability should help explain the negative relationship between economic rationality and age.

To evaluate this hypothesis we asked participants to complete a 20-element Ravens' Matrix task. Each element of the task ask subjects to identify, from a list, the missing piece of a patterned geometric object. Success requires subjects to identify the pattern and the geometry of the object to see what is missing. The task is a classic test of intelligence and have been used by social scientists for decades. An important advantage of the test is that it does not require literacy, but just an understanding of what it means to "solve" each puzzle.

Again, not all of the analysis sample performed the task. In this case,

just 44% (1,810) completed the 20 questions. As expected, however, older people perform substantially worse on this task. Those younger than 23 solved an average of 10.7 matrices correctly. People age 51 to 63 solved 8.7, while those ages 64 and older solved an average of just 6.8 matrices correctly.

Given the sharp difference in cognitive ability by age, there is good reason to think that the Ravens' test score can explain an important amount of the correlation between age and economic rationality in the experiment. Table 5 shows, however, that this is not the case. For reference, the first two columns of Table 5 replicate the basic specification of Table 3 (columns 2 and 4). Columns 3 and 4 of Table 5 repeat these specifications but restrict attention to those who completed the Raven's matrices task. The coefficients on the older ages are smaller in this population, but still negative. Conditioning on the Raven's score (in columns 5 and 6) leaves the coefficient on the older age indicators virtually unchanged, however. The Raven's test score is positively correlated with rationality in the experiment, but not strongly so. Thus, we find no evidence that cognitive declines associated with normal aging are driving the negative relationship between economic rationality and age.

5.3 The Result of Declining Health

Finally, we consider the possibility that the lower average levels of economic rationality among older people is largely attributable to those who have had important health problems. This source of the negative correlation between rationality and age sees illnesses and conditions that disproportionately affect older people as the underlying cause of their lower decisionmaking quality.

To assess this possibility, we restrict attention to the 2,933 participants of the analysis sample about whom detailed health information is also non-missing. For reference, Table 6, columns 1 and 2 reproduce the results of the baseline specification for the entire analysis sample. Columns 3 and 4 then replicate our basic specification for this subsample. The coefficients on the indicators for older ages are very similar to those for the entire sample.

Columns 5 and 6 of Table 6 add to the basic specification a rich set of controls for health. These include indicators for self-reported health being poor or good (relative to very good or excellent), the number of activities of daily living and instrumental activities of daily living with which the participant reports important difficulty, an indicator for having a chronic ailment or disability, and a self-reported measure of body mass index (BMI). Many of these measures of health have the expected sign in the regression (results available upon request), but their inclusion has virtually no effect on

the age coefficients. Even conditional on health, older people exhibit lower levels of consistency with economic rationality in the experiment.

6 Conclusion

Social science has documented that older people have lower levels of several cognitive skills. Some of these skills are likely important for making sound spending, saving, insurance, or health choices. Indeed, our prior work (Choi et al, 2014) showed that older people exhibit lower levels of economic rationality in a simple, incentivized choice experiment. In this way, we found evidence that older people are less able (or willing) to make choices consistent with any stable objective.

In this paper, we used a new dataset to investigate the sources of the correlation between age and economic rationality. In the new data, we replicate the qualitative relationship observed in the previous work. Older people have lower levels of economic rationality and these differences are both statistically and economically significant.

The findings suggest that the decline in economic rationality with age is a true age effect, not a cohort effect. We find no evidence that the lower levels of decisionmaking quality among older people are due to their disproportionate difficulties with the computer interface used in the experiment. In this way, the findings indicate that the relatively lower levels of decisionmaking quality among older people should be expected to persist.

Similarly, we find no evidence that the decline in economic rationality in the experiment is explained by more general forms of cognitive decline. Thus, the findings indicate that the lower decisionmaking quality among older people is not a simple consequence of declines in general forms of cognitive ability. Finally, we find no evidence that the lower average levels of cognitive ability in older people is concentrated among those who have experienced important declines in physical health. Certainly older people are in worse health, but we find no evidence that this explains why they exhibit lower levels of rationality in the experiment.

Taken together the results indicate that economic decisionmaking quality declines with normal aging and this is distinct phenomenon from other forms of cognitive decline with age. The evidence thus suggests that there is something distinct about financial decision making ability that declines in older ages. If older people make different choices from younger ones or from what a normative theory would suggest, it follows that additional scrutiny is justified. The evidence presented here suggests it is less likely to be ex-

clusively a matter of special tastes among the elderly. Instead, the results presented here suggest analysts should give more credence to the possibility that diminished capacities to understand and make effective economic tradeoffs are influencing choices in older population. It follows that policy-makers and market participants that are concerned with the welfare of older people should focus more attention on assuring that products and markets are designed to that the elderly have same effective opportunities as younger populations.

References

- [1] Abaluck, J. and Jonathan Gruber (2011) “Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program.” *American Economic Review*, 101(4), 1180-1210.
- [2] Afriat, S. (1967) “The Construction of a Utility Function from Expenditure Data.” *Econometrica*, 6, pp. 67-77.
- [3] Afriat, S. (1972) “Efficiency Estimates of Production Functions.” *International Economic Review*, 8, pp. 568-598.
- [4] Afriat, S. (2012) “Afriat’s Theorem and the Index Number Problem.” *The Economic Journal*, 122, pp. 295-304.
- [5] Agarwal, S., J. Driscoll, X. Gabaix and D. Laibson (2009) “The Age of Reason: Financial Decisions over the Life-Cycle with Implications for Regulation.” *Brookings Papers on Economic Activity*, 2, pp. 51-117.
- [6] Ameriks, J., A. Caplin and J. Leahy (2003) “Wealth Accumulation and the Propensity to Plan.” *Quarterly Journal of Economics*, 118, pp. 1007-1047.
- [7] Banks, J. (2010) “Cognitive Function, Financial Literacy and Financial Outcomes at Older Ages: Introduction.” *Economic Journal*, 120, pp. 357-362.
- [8] Bernheim, D. and D. Garrett (2003) “The Effects of Financial Education in the Workplace: Evidence from a Survey of Households.” *Journal of Public Economics*, 87, pp. 1487-1519.
- [9] Bernheim, D. and A. Rangel (2008) “Choice-Theoretic Foundations for Behavioral Welfare Economics.” In *The Foundations of Positive and*

Normative Economics, ed. A. Caplin and A. Schotter. Oxford University Press.

- [10] Bernheim, D. and A. Rangel (2009) “Beyond Revealed Preference: Theoretic Foundations for Behavioral Economics.” *Quarterly Journal of Economics*, 124, pp. 51-104.
- [11] Camerer, C., S. Issacharoff, G. Loewenstein, T. O’Donoghue and M. Rabin (2003) “Regulation for Conservatives: Behavioral Economics and the Case for Asymmetric Paternalism.” *University of Pennsylvania Law Review*, 151, pp. 1211-1254.
- [12] Choi, J., D. Laibson and B. Madrian (2004) “Plan Design and 401(k) Savings Outcomes.” *National Tax Journal*, 57, pp. 275-298.
- [13] Choi S., R. Fisman, D. Gale and S. Kariv (2007a) “Revealing Preferences Graphically: An Old Method Gets a New Tool Kit,” *American Economic Review Papers & Proceedings*, 97, pp. 153-158.
- [14] Choi S., R. Fisman, D. Gale and S. Kariv (2007b) “Consistency and Heterogeneity of Individual Behavior under Uncertainty.” *American Economic Review*, 97, pp. 1921-1938.
- [15] Choi S., S. Kariv, W. Müller and D. Silverman (2014) “Who Is (More) Rational?” *American Economic Review*, 104(6), 1518-50.
- [16] Cole, S. and G. Shastry (2009) “Smart Money: The Effect of Education, Cognitive Ability, and Financial Literacy on Financial Market Participation.” Harvard Business School Finance Working Paper No. 09-071.
- [17] Diewert, W (2012) “Afriat’s Theorem and some Extensions to Choice under Uncertainty.” *The Economic Journal*, 122, pp. 305-331.
- [18] Dohmen, T., A. Falk, D. Huffman and U. Sunde (2010) “Are Risk Aversion and Impatience Related to Cognitive Ability?” *American Economic Review*, 100, pp. 1238-1260.
- [19] Duflo, E. and E. Saez (2003) “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment.” *Quarterly Journal of Economics*, 118, pp. 815-842.
- [20] Echenique, F., S. Lee and M. Shum (2011) “The Money Pump as a Measure of Revealed Preference Violations.” *Journal of Political Economy* 119, pp. 1201-1223.

- [21] Fang, H., M. Keane and D. Silverman (2008) "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy*, 116(2), pp. 303-350.
- [22] Fisman R., S. Kariv and D. Markovits, (2007) "Individual Preferences for Giving." *American Economic Review*, 97, pp. 1858-1876.
- [23] Hadar, J. and W. Russell (1969) "Rules for Ordering Uncertain Prospects." *American Economic Review*, 59, pp. 25-34.
- [24] Heiss, F., D. McFadden and J. Winter (2010) "Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans," in David A. Wise (ed.) *Research Findings in the Economics of Aging*, NBER, University of Chicago Press, pages 413 - 481.
- [25] Houtman, M. and J. Maks (1985) "Determining all Maximial Data Subsets Consistent with Revealed Preference." *Kwantitatieve Methoden*, 19, pp. 89-104.
- [26] Levitt, S. and J. List (2007) "What do Laboratory Experiments Tell Us About the Real World?" *Journal of Economic Perspectives*, 21, pp. 153-174.
- [27] Loewenstein, G. and E. Haisley (2008) "The Economist as Therapist: Methodological Ramifications of "Light" Paternalism." In *The Foundations of Positive and Normative Economics*, ed. A. Caplin and A. Schotter. Oxford University Press.
- [28] Lusardi, A. and O. Mitchell (2007) "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics*, 54, pp. 205-224.
- [29] Madrian, B. and D. Shea (2001) "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics*, 116, pp. 1149-1525.
- [30] Manski, C. (1995) "Identification Problems in the Social Sciences." Cambridge: Harvard University Press.
- [31] McArdle, John J.; Ferrer-Caja, Emilio; Hamagami, Fumiaki; Woodcock, Richard W. (2002). *Developmental Psychology*, Vol 38(1), Jan, 115-142.

- [32] McFadden, D., J. Winter and F. Heiss (2008) "Consumer-Directed Health Care: Can Consumers Look After Themselves?" *Swiss Journal of Economics and Statistics*, 144(III), pp. 285-307.
- [33] Rohwedder, Susann, and Robert J. Willis (2010). "Mental Retirement" *Journal of Economic Perspectives*, 24(1): 119-38.
- [34] Samuelson, P. (1947) *Foundations of Economic Analysis*. Cambridge: Harvard University Press.
- [35] Smith, J., J. McArdle and R. Willis (2010) "Financial Decision Making and Cognition in a Family Context." *Economic Journal*, 120, pp. 363-380.
- [36] Thaler, R. and S. Benartzi (2004) "Save More Tomorrow: Using Behavioral Economics to Increase Employee Savings." *Journal of Political Economy*, 112, pp. 164-187.
- [37] Thaler, R. and C. Sunstein (2003) "Libertarian Paternalism." *American Economic Review*, Papers & Proceedings, 93, pp. 175-179.
- [38] Varian, H. (1982) "The Nonparametric Approach to Demand Analysis." *Econometrica*, 50, pp. 945-972.
- [39] Varian, H. (1983) "Non-Parametric Tests of Consumer Behaviour." *Review of Economic Studies*, 50, pp. 99-110.
- [40] Varian, H. (1990) "Goodness-of-Fit in Optimizing Models." *Journal of Econometrics*, 46, pp. 125-140.
- [41] Varian, H. (1991) "Goodness-of-Fit for Revealed Preference Tests." Mimeo.
- [42] Varian, H. (2012) "Revealed Preference and its Applications." *The Economic Journal*, 122, pp. 332-338.
- [43] Vermeulen, F. (2012) "Foundations of Revealed Preference: Introduction." *The Economic Journal*, 122, pp. 287-294.
- [44] Willis, R. J. (2007) "Cognitive Economics and Human Capital." Presidential Address to the Society of Labor Economists. Chicago, IL.

7 Figures and Tables

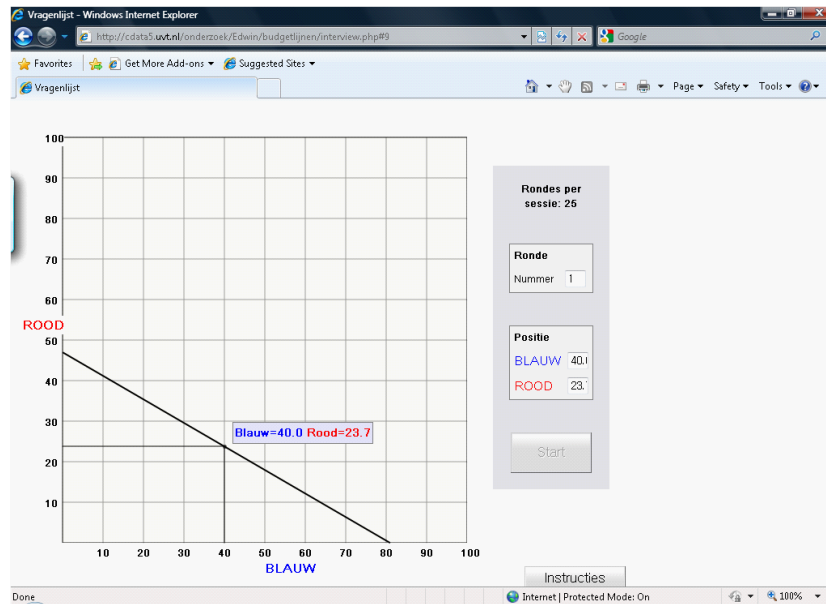


Figure 1: Example of the Experimental Interface

Table 1. Socio-demographic information

	Percent
Female	53.20
Age	
16-23	11.71
24-35	17.09
36-50	25.53
51-63	23.37
64-71	13.37
72+	8.94
Education	
Primary school	8.82
Intermediate vocational ed.	23.77
Intermediate secondary ed.	22.36
Higher vocational ed.	22.48
Higher secondary ed.	12.45
University degree	10.12
Individual gross monthly income	
€0-900	25.35
€901-1805	24.60
€1806-2835	25.12
€2835+	24.94
Occupation	
Paid work	51.54
House work	7.04
Disabled	4.04
Retired	18.03
Others	19.35
Household composition	
Partner	70.50
# of kids	0.84
# of obs.	4160

Table 2A. CCEI scores

	Mean	Sd	Percentiles					# of obs.
			10	25	50	75	90	
All	0.870	0.142	0.668	0.792	0.913	0.992	1.000	4,160
Female	0.867	0.144	0.664	0.790	0.912	0.992	1.000	2,213
Age								
16-23	0.869	0.139	0.656	0.802	0.912	0.988	1.000	487
24-35	0.874	0.132	0.679	0.787	0.914	0.992	1.000	711
36-50	0.874	0.141	0.670	0.798	0.921	0.995	1.000	1,062
51-63	0.871	0.145	0.673	0.794	0.918	0.994	1.000	972
64-71	0.854	0.148	0.652	0.770	0.890	0.985	1.000	556
72+	0.867	0.144	0.663	0.793	0.911	0.990	1.000	372
Education								
Primary school	0.866	0.150	0.663	0.789	0.909	0.995	1.000	367
Intermediate vocational ed	0.870	0.138	0.669	0.794	0.910	0.991	1.000	989
Intermediate secondary ed	0.867	0.142	0.673	0.792	0.908	0.990	1.000	930
Higher vocational ed.	0.874	0.137	0.684	0.794	0.914	0.991	1.000	935
Higher secondary ed.	0.866	0.154	0.637	0.787	0.919	0.995	1.000	518
University degree	0.873	0.135	0.677	0.780	0.918	0.996	1.000	421
Individual gross monthly income								
€0-900	0.873	0.140	0.673	0.805	0.916	0.993	1.000	631
€901-1805	0.860	0.147	0.647	0.781	0.911	0.990	1.000	458
€1 806-2835	0.874	0.136	0.661	0.794	0.919	0.992	1.000	579
€2835+	0.869	0.145	0.664	0.787	0.910	0.995	1.000	768
Occupation								
Paid work	0.872	0.139	0.674	0.794	0.915	0.993	1.000	2,144
House work	0.870	0.144	0.674	0.798	0.918	0.994	1.000	293
Disabled	0.845	0.160	0.621	0.764	0.883	0.987	1.000	168
Retired	0.867	0.145	0.657	0.791	0.910	0.989	1.000	750
Others	0.870	0.140	0.667	0.788	0.916	0.993	1.000	805

Table 2B. MPI scores

	Mean	Sd	Percentiles					# of obs.
			10	25	50	75	90	
All	0.880	0.160	0.638	0.809	0.957	0.998	1.000	4,160
Female	0.877	0.163	0.636	0.807	0.955	0.998	1.000	2,213
Age								
16-23	0.884	0.155	0.640	0.818	0.957	0.997	1.000	487
24-35	0.881	0.154	0.643	0.796	0.958	0.998	1.000	711
36-50	0.887	0.155	0.647	0.821	0.963	0.999	1.000	1,062
51-63	0.881	0.162	0.641	0.815	0.959	0.999	1.000	972
64-71	0.862	0.173	0.612	0.789	0.933	0.995	1.000	556
72+	0.876	0.169	0.633	0.804	0.953	0.997	1.000	372
Education								
Primary school	0.872	0.176	0.607	0.804	0.948	0.999	1.000	367
Intermediate vocation	0.881	0.154	0.637	0.809	0.953	0.997	1.000	989
Intermediate seconda	0.878	0.161	0.650	0.807	0.950	0.998	1.000	930
Higher vocational ed.	0.883	0.159	0.639	0.813	0.961	0.998	1.000	935
Higher secondary ed.	0.879	0.168	0.609	0.814	0.958	0.999	1.000	518
University degree	0.885	0.155	0.648	0.802	0.962	0.999	1.000	421
Individual gross monthly income								
€0-900	0.888	0.152	0.653	0.818	0.964	0.998	1.000	631
€901-1805	0.872	0.165	0.607	0.790	0.952	0.997	1.000	458
€1806-2835	0.884	0.152	0.647	0.816	0.955	0.998	1.000	579
€2835+	0.881	0.160	0.632	0.803	0.959	0.999	1.000	768
Occupation								
Paid work	0.884	0.155	0.647	0.815	0.959	0.998	1.000	2,144
House work	0.874	0.164	0.634	0.798	0.951	0.998	1.000	293
Disabled	0.839	0.191	0.562	0.717	0.922	0.996	1.000	168
Retired	0.877	0.166	0.636	0.814	0.955	0.997	1.000	750
Others	0.882	0.161	0.638	0.817	0.959	0.998	1.000	805

Table 3. The correlation between rationality scores and age

	(1)	(2)	(3)	(4)
	CCEI	CCEI	MPI	MPI
Constant	.868*** (.007)	.861*** (.011)	0.885*** (0.007)	.870*** (.013)
Age				
24-35	0.004 (.008)	0.005 (0.010)	-0.005 (0.009)	-0.003 (0.011)
36-50	0.006 (.011)	0.008 (0.010)	0.001 (0.009)	0.003 (0.011)
51-63	0.002 (.011)	0.004 (0.010)	-0.005 (0.009)	-0.001 (0.011)
64-71	-0.015* (0.009)	-0.038*** (0.014)	-0.026** (0.010)	-0.050*** (0.017)
72+	-0.001 (0.010)	-0.025* (0.015)	-0.008 (0.011)	-0.034* (0.018)
Socioeconomic controls	no	yes	no	yes
R^2	0.0021	0.0079	0.0027	0.0111
# of obs.	3,910	3,910	3,910	3,910

OLS estimates with the omitted age category being those from 16-23. Socioeconomic controls include indicators for gender, employment statuses, and completed education levels, as well as levels of individual income and the number of children in the household. Standard errors, robust to heteroskedasticity, are in parentheses. *, **, *** indicate 10, 5, 1 percent significance levels, respectively.

Table 4. Evaluation a Cohort Effect:
the correlation between rationality scores, age, and performance on a target task

	(1)	(2)	(3)	(4)	(5)	(6)
	CCEI	MPI	CCEI	MPI	CCEI	MPI
Constant	.861*** (.011)	.870*** (.013)	0.865*** (0.014)	.870*** (.017)	0.868*** (0.015)	.874*** (.017)
Age						
24-35	0.005 (0.010)	-0.003 (0.011)	0.011 (0.014)	0.020 (0.015)	0.010 (0.014)	0.019 (0.015)
36-50	0.008 (0.010)	0.003 (0.011)	0.006 (0.013)	0.016 (0.014)	0.006 (0.013)	0.016 (0.014)
51-63	0.004 (0.010)	-0.001 (0.011)	0.003 (0.013)	0.009 (0.014)	0.004 (0.013)	0.010 (0.014)
64-71	-0.038*** (0.014)	-0.050*** (0.017)	-0.033* (0.019)	-0.037* (0.022)	-0.032* (0.019)	-0.036* (0.022)
72+	-0.025* (0.015)	-0.034* (0.018)	-0.029 (0.020)	-0.030 (0.023)	-0.027 (0.020)	-0.027 (0.024)
I(Missed the target)					0.002 (0.012)	0.009 (0.013)
Size of miss (# of points)					-0.001 (0.001)	-0.001* (0.001)
Socioeconomic controls	yes	yes	yes	yes	yes	yes
R^2	0.0079	0.0111	0.0125	0.0143	0.0134	0.0163
# of obs.	3,910	3,910	2,374	2,374	2,374	2,374

OLS estimates with the omitted age category being those from 16-23. Soc ioeconomic controls include indicators for gender, employment statuses, and completed education levels, as well as levels of individual income and the number of children in the household. Columns 5 and 6 include controls for whether, in a separate round, the participant failed to choose a specific target bundle and, if so, by how many points did she miss the target. Standard errors, robust to heteroskedasticity, are in parentheses. *, **, *** indicate 10, 5, 1 percent significance levels, respectively.

Table 5. The Role of General Cognitive Decline:
the correlation between rationality scores, age, and performance on Raven's Matrices

	(1)	(2)	(3)	(4)	(5)	(6)
	CCEI	MPI	CCEI	MPI	CCEI	MPI
Constant	.861*** (.011)	.870*** (.013)	0.851*** (0.018)	.861*** (.021)	0.848*** (0.020)	.854*** (.024)
Age						
24-35	0.005 (0.010)	-0.003 (0.011)	0.013 (0.016)	0.019 (0.018)	0.014 (0.016)	0.019 (0.018)
36-50	0.008 (0.010)	0.003 (0.011)	0.014 (0.015)	0.024 (0.016)	0.014 (0.016)	0.025 (0.017)
51-63	0.004 (0.010)	-0.001 (0.011)	0.020 (0.016)	0.025 (0.017)	0.021 (0.016)	0.027 (0.017)
64-71	-0.038*** (0.014)	-0.050*** (0.017)	-0.027 (0.023)	-0.031 (0.027)	-0.026 (0.024)	-0.029 (0.027)
72+	-0.025* (0.015)	-0.034* (0.018)	-0.016 (0.025)	-0.022 (0.029)	-0.014 (0.025)	-0.019 (0.029)
Score on Raven's Test (0-20)					0.0004 (0.0010)	0.0007 (0.0012)
Socioeconomic controls	yes	yes	yes	yes	yes	yes
R^2	0.0079	0.0111	0.0118	0.0136	0.0118	0.0138
# of obs.	3,910	3,910	1,698	1,698	1,698	1,698

OLS estimates with the omitted age category being those from 16-23. Socioeconomic controls include indicators for gender, employment statuses, and completed education levels, as well as levels of individual income and the number of children in the household. Standard errors, robust to heteroskedasticity, are in parentheses. *, **, *** indicate 10, 5, 1 percent significance levels, respectively.

Table 6. The Role of Declining Health:
the correlation between rationality scores, age, and health

	(1)	(2)	(3)	(4)	(5)	(6)
	CCEI	MPI	CCEI	MPI	CCEI	MPI
Constant	.861*** (.011)	.870*** (.013)	0.849*** (0.014)	.862*** (.016)	0.893*** (0.021)	.926*** (.024)
Age						
24-35	0.005 (0.010)	-0.003 (0.011)	0.014 (0.013)	0.004 (0.014)	0.015 (0.013)	0.005 (0.014)
36-50	0.008 (0.010)	0.003 (0.011)	0.014 (0.012)	0.007 (0.014)	0.015 (0.012)	0.008 (0.014)
51-63	0.004 (0.010)	-0.001 (0.011)	0.014 (0.012)	0.008 (0.013)	0.014 (0.012)	0.009 (0.013)
64-71	-0.038*** (0.014)	-0.050*** (0.017)	-0.034** (0.016)	-0.049** (0.020)	-0.033** (0.016)	-0.047** (0.020)
72+	-0.025* (0.015)	-0.034* (0.018)	-0.020 (0.018)	-0.035* (0.021)	-0.019 (0.018)	-0.033 (0.021)
Health controls	no	no	no	no	yes	yes
Socioeconomic controls	yes	yes	yes	yes	yes	yes
R^2	0.0079	0.0111	0.0107	0.0125	0.0164	0.0198
# of obs.	3,910	3,910	2,762	2,762	2,762	2,762

OLS estimates with the omitted age category being those from 16-23. Health controls include indicators for poor and good self-reported health (very good or excellent health category omitted), and for having a chronic ailment. Health controls also include the number of (i)adls with which the respondent has difficulty and a measure of body mass index. Socioeconomic controls include indicators for gender, employment statuses, and completed education levels, as well as levels of individual income and the number of children in the household. Standard errors, robust to heteroskedasticity, are in parentheses. *, **, *** indicate 10, 5, 1 percent significance levels, respectively.

8 Appendix: Testing for Consistency with GARP

I. Afriat's (1967) Theorem

Let $\{(p^i, x^i)\}_{i=1}^{25}$ be the data generated by a participant's choices, where p^i is the i -th observation of prices and x^i is the associated choice. An allocation x^i is *directly revealed preferred* to an allocation x^j , denoted $x^i R^D x^j$, if $p^i \cdot x^i \geq p^i \cdot x^j$. An allocation x^i is *revealed preferred* to x^j , denoted $x^i R x^j$, if there exists a sequence of allocations $\{x^k\}_{k=1}^K$ with $x^1 = x^i$ and $x^K = x^j$, such that $x^k R^D x^{k+1}$ for every $k = 1, \dots, K - 1$.

The Generalized Axiom of Revealed Preference (GARP) requires that if $x^i R x^j$ then $p^j \cdot x^j \leq p^j \cdot x^i$; that is, if x^i is revealed preferred to x^j , then x^i must cost at least as much as x^j given the prices when x^j is chosen. If the data are generated by a non-satiated utility function, then they satisfy GARP. Conversely, Afriat (1967) shows us that if a finite data set of choices satisfies GARP, then the data can be rationalized by a utility function.

Afriat's (1967) Theorem If the data set $\{(p^i, x^i)\}$ satisfies GARP, then there exists a piecewise linear, continuous, increasing, concave utility function $u(x)$ such that for each observation (p^i, x^i) $u(x) \leq u(x^i)$ for any x such that $p^i \cdot x \leq p^i \cdot x^i$.

This statement of the theorem is due to Varian (1982, 1983), who replaced the condition Afriat called *cyclical consistency* with GARP. Note that satisfying GARP entails only that choices are consistent with utility maximization. The implication that choices may be rationalized by a well-behaved utility function is a consequence of linear budget lines. When the budget constraints are linear, if any utility function can represent the choices, then a well-behaved one can too.

II. Goodness-of-fit

GARP offers an exact test. Choice data either satisfy the axiom, or they do not. It is therefore useful to measure the *extent* of GARP violations. This paper reports measures of GARP violations based on two indices: Afriat's (1972) critical cost efficiency index (CCEI), and Echinique et al.'s (2011) money pump index.

Afriat (1972) The CCEI measures the amount by which each budget constraint must be adjusted in order to remove all violations of GARP. For any number $0 \leq e \leq 1$, define the direct revealed preference relation $R^D(e)$ as

$$x^i R^D(e) x^j \iff e p^i \cdot x^i \geq p^i \cdot x^j,$$

and define $R(e)$ to be the transitive closure of $R^D(e)$. Let e^* be the largest value of e such that the relation $R(e)$ satisfies GARP. Afriat’s CCEI is the value of e^* associated with the data set $\{(p^i, x^i)\}$. Figure A1 illustrates the construction of the CCEI for a simple violation of GARP involving two allocations, x^1 and x^2 . Allocation x^1 is revealed preferred to x^2 because $p^1 \cdot x^1 > p^1 \cdot x^2$, yet x^1 is cheaper than x^2 at the prices at which x^2 is purchased, $p^2 \cdot x^1 < p^2 \cdot x^2$. Here we have a violation of the Weak Axiom of Revealed Preference (WARP) since $x^1 R^D x^2$ and $x^2 R^D x^1$. If we shifted the budget line through x^2 as shown ($A/B < C/D$) the violation would be removed so the CCEI score associated with this violation of GARP is A/B .

The CCEI is bounded between zero and one and the closer it is to one, the smaller the perturbation of the budget lines required to remove all violations and thus the closer the data are to satisfying GARP. Although the CCEI provides a summary statistic of the overall consistency of the data with GARP, it does not give any information about which of the observations (p^i, x^i) are causing the most severe violations. A single large violation may lead to a small value of the index while a large number of small violations may result in a much larger efficiency index.

Echenique et al. (2011) Echenique et al. also provide a goodness-of-fit for consistency with GARP. Their measure derives from the important idea that violations GARP can be exploited as a “money pump.” Calculation of their Money Pump Index for a simple violation of GARP is also illustrated in Figure A1. An “arbitrageur” who purchased allocation x^1 at prices p^2 and allocation x^2 at prices p^1 could trade x^1 with the individual at prices p^1 and x^2 at prices p^2 , and earn a profit

$$mpi = p^1(x^1 - x^2) + p^2(x^2 - x^1) = C/D + A/B.$$

Echenique et al. (2011) use the mean and median mpi ’s across all violations of GARP (cyclic sequences of allocations). The reasons for the discrepancies between the CCEI and MPI are discussed in Echenique et al. (2011).

III. The power of the GARP tests

Revealed preference tests have a drawback: there is no clear metric for judging whether individuals are “close” to satisfying GARP. To generate a benchmark against which to compare these CCEI scores, we use the test designed by Bronars (1987), which builds on Becker (1962) and employs the choices of a hypothetical subject who chooses randomly among all allocations on each budget line as a point of comparison. The mean CCEI score across all subjects in our experiment is 0.87 whereas the mean CCEI score for a

random sample of 25,000 simulated subjects is 0.659. More than half of actual subjects have CCEI's above 0.913, while only about seven percent of simulated subjects have CCEI's that high.

The Bronars' (1987) test has often been applied to experimental data, so using it situates our results in a literature. The setup used in this study has the highest Bronars power of one (all random subjects had violations). Our results show that the experiment is sufficiently powerful to exclude the possibility that consistency is the accidental result of random behavior. To provide a more informative metric of the consistency of choices, we follow Choi et al. (2007a) who extend and generalize the Bronars (1987) test by employing a random sample of simulated subjects who maximize a utility function $U(\cdot)$ with error where the likelihood of error is assumed to be a decreasing function of its cost. In particular, we assume an idiosyncratic preference shock that has a logistic distribution. This implies the probability of choosing the allocation x^* satisfies

$$\Pr(x^*) = \frac{e^{\gamma U(x^*)}}{\int_{x:p \cdot x=1} e^{\gamma U(x)},}$$

where the parameter γ reflects sensitivity to differences in utility. The choice of allocation becomes purely random as γ goes to zero (Bronars test), whereas the probability of the allocation yielding the highest utility increases as γ increases.

The histograms in Figure A2, replicated from Choi et al (2014), below summarize the distributions of CCEI scores generated by samples of 25,000 simulated subjects who implement the logarithmic von Neumann-Morgenstern utility function $\log x_1 + \log x_2$ with various levels of precision γ . The horizontal axis measures the fractions for different intervals of CCEI scores and the vertical axis measures the percentage of subjects corresponding to each interval. Each of the simulated subjects makes 25 choices from randomly generated budget lines in the same way as the human subjects do. The number above each bar of the histogram represents the percentage of actual subjects corresponding to each interval. The histograms show the extent to which subjects did worse than choosing consistently and the extent to which they did better than choosing randomly. The histograms thus demonstrate that if utility maximization is not in fact the correct model, then our experiment is sufficiently powerful to detect it.

Additional references

1. Becker, G. (1962) "Irrational Behavior and Economic Theory." *Journal of Political Economy*, 70, pp. 1-13
2. Bronars, S. (1987) "The power of nonparametric tests of preference maximization." *Econometrica*, 55, pp. 693-698.
3. Varian, H. (1982) "The Nonparametric Approach to Demand Analysis." *Econometrica*, 50, pp. 945-972.
4. Varian, H. (1983) "Non-Parametric Tests of Consumer Behaviour." *Review of Economic Studies*, 50, pp. 99-110.
5. Varian, H. (1990) "Goodness-of-Fit in Optimizing Models." *Journal of Econometrics*, 46, pp. 125-140.
6. Varian, H. (1991) "Goodness-of-Fit for Revealed Preference Tests." Mimeo.

8.1 Appendix Figures

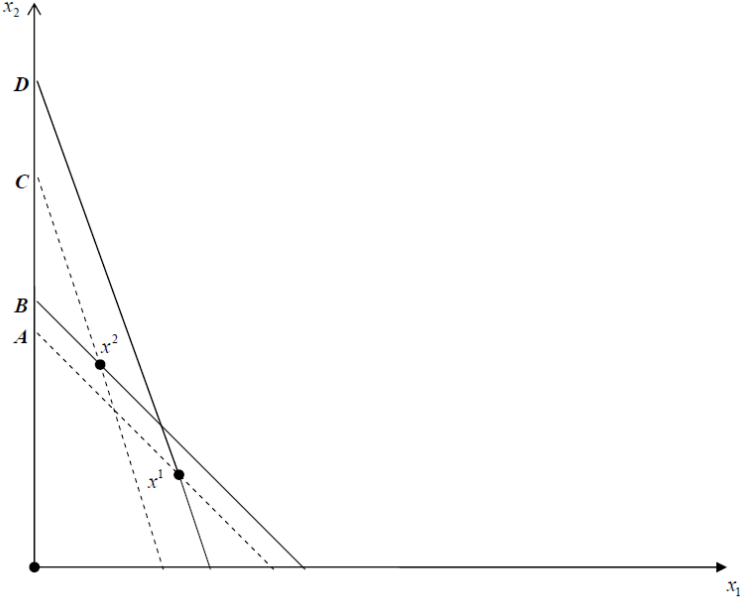


Figure A1: Calculation of the CCEI for a Simple Violation of the Weak Axiom of Revealed Preference

Figure A2. The distributions of CCEI scores of simulated subjects

