NBER WORKING PAPER SERIES

WHO IS (MORE) RATIONAL?

Syngjoo Choi
Shachar Kariv
Wieland Müller
Dan Silverman

Working Paper 16791
http://www.nber.org/papers/w16791

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2011

We thank Douglas Gale and Raymond Fisman for detailed comments and suggestions. We are also grateful to James Andreoni, James Banks, Richard Blundell, Andrew Caplin, David Card, Thomas Crossley, Stefano DellaVigna, Guillaume Frechette, Steffen Huck, Patrick Kline, Ron Lee, John List, Nicola Persico, Imran Rasul, Joel Slemrod, Arthur van Soest and Hans-Martin von Gaudecker for helpful discussions and comments. This paper has also benefited from suggestions by the participants of seminars at several universities and conferences. Jelmer Ypma provided excellent research assistance. We thank Corrie Vis and Edwin de Vet of CentERdata (Tilburg University) for software development and technical support. We acknowledge financial support from the Economic and Social Research Council (ESRC) Grant No. RES-061-25-0348 (Choi), the Center on the Economics and Demography of Aging (CEDA) and the Coleman Fung Risk Management Research Center (OpenLink Fund) at UC Berkeley (Kariv), and the Netherlands Organisation for Scientific Research (NWO) and the Network for Studies on Pensions, Aging and Retirement (Netspar) (Mueller) The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Who Is (More) Rational?
Syngjoo Choi, Shachar Kariv, Wieland Müller, and Dan Silverman
NBER Working Paper No. 16791
February 2011
JEL No. C93,D01,D03,D12,D81

ABSTRACT

Revealed preference theory offers a criterion for decision-making quality: if decisions are high quality then there exists a utility function that the choices maximize. We conduct a large-scale field experiment that enables us to test subjects' choices for consistency with utility maximization and to combine the experimental data with a wide range of individual socioeconomic information for the subjects. There is considerable heterogeneity in subjects' consistency scores: high-income and high-education subjects display greater levels of consistency than low-income and low-education subjects, men are more consistent than women, and young subjects are more consistent than older subjects. We also find that consistency with utility maximization is strongly related to wealth: a standard deviation increase in the consistency score is associated with 15-19 percent more wealth. This result conditions on socioeconomic variables including current income, education, and family structure, and is little changed when we add controls for past income, risk tolerance and the results of a standard personality test used by psychologists.

Syngjoo Choi
Department of Economics
University College London
syngjoo.choi@ucl.ac.uk

Shachar Kariv
Department of Economics
University of California, Berkeley
Berkeley, CA 94720
kariv@berkeley.edu

Wieland Müller
Department of Economics
University of Vienna (Austria)
and Tilburg University (The Netherlands)
wieland.mueller@univie.ac.at

Dan Silverman
Department of Economics
University of Michigan
Ann Arbor MI 48109-1220
and NBER
dansilv@umich.edu

Keywords: Decision-making quality, rationality, revealed preference, risk, wealth differentials, Netherlands, experiment, CentERpanel.

1 Introduction

Traditional economic analysis assumes that choices are rational; decision makers choose their preferred alternative from the feasible set given the information available to them. In this standard view, heterogeneity in choices is attributed to heterogeneity in *preferences*, *information*, *beliefs*, or *constraints*. More recently, several strands of empirical research consider heterogeneity in choices driven, also, by differences in the *quality* of decision-making. Prominent examples of this research include Ameriks et al. (2003), Bernheim and Garrett (2003), and Agarwal et al. (2009).

Whether we treat individuals as high-quality decision makers has important consequences. If decision-making skills are poor or the costs of making an optimal decision are high, then there are potentially important wedges between the choices that some individuals actually make and the choices they would make if they had the skills or knowledge to make higher quality decisions. These wedges matter because then "revealed" preferences may not be "true" underlying preferences. In that case, positive predictions or welfare conclusions based on the revealed preferences may be misleading.

While the possibility of heterogeneity in decision-making quality has important consequences for economic analysis, definitive judgment about the quality of decision-making is generally made difficult by twin problems of *identification* and *measurement*. The identification problem is to distinguish differences in decision-making quality from unobserved differences in preferences, information, beliefs or constraints. Identification is important because welfare conclusions and thus (constrained) optimal policy will depend 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.

In this paper, we measure aspects of decision-making quality by calculating how nearly individual choice behavior in an experiment complies with economic rationality in the sense of a consistent (complete and transitive) preference ordering. Classical revealed preference theory tells us that choices are consistent with maximizing a (well-behaved) utility function if and only if they satisfy the Generalized Axiom of Revealed Preference (GARP). We take the view that if there is no utility function that choices maximize then those choices cannot be considered purposeful and, in this way, high quality. This criterion for decision-making quality is not as restrictive as might be thought. It simply requires consistent preferences over all possible alternatives, and choices that correspond to the most preferred alternative in the feasible set. Any consistent preference ordering is admissible.[1]

Because GARP offers an exact test (either choice data satisfy GARP or they do not), a variety of goodness-of-fit indices have been proposed for quantifying the extent of violation. The main index is Afriat's (1972) Critical Cost Efficiency Index (CCEI). By definition, the CCEI is bounded between zero and one. The closer it is to one, the closer the data are to satisfying GARP; and the difference between the CCEI and one can be interpreted as an *upper bound* of the fraction of income that an individual is "wasting" by making inconsistent choices.

Using revealed preference axioms, we test whether choice behavior in an experimental setting by a broad population is consistent with the utility maximization model. Our points of departure from the literature on decision-making quality, cognitive and non-cognitive skills, and economic outcomes derive from two observations:

[1] Consistency with utility maximization does not depend on preferences, and the experimental task we study makes no special demands of outside knowledge or expertise. This helps to isolate heterogeneity in decision-making quality from heterogeneity in preferences, information, beliefs or constraints (the identification problem).

[2] The CCEI (and other goodness-of-fit indices) has a coherent economic interpretation and is easily adapted for application in a variety of decision domains. The theoretical framework and portability of the measure are valuable for drawing conclusions that go beyond the particular setting of the experiment (the measurement problem).

There is a vast amount of work on the rationality of decisions, and laboratory experiments have provided key empirical guideposts for developments in this area. To connect the insights that we have gained from the experimental study of rational choice under laboratory conditions to practical questions in the broader world, we conducted an experiment using the CentERpanel, a representative sample of over 2,000 Dutch-speaking households in the Netherlands. The advantage of using the CentERpanel is the wide range of individual sociodemographic and economic information that it provides about the panel members.

In our experiment, we present subjects with 25 standard consumer decision problems that can be interpreted either as the selection of a bundle of commodities from a standard budget line or the allocation of an endowment between risky assets. These decision problems are presented using

[1]Our laboratory experiment involves decision making under risk (more below). Our criterion for decision-making quality does not, however, rely on a specific theory of risk preferences. In particular, the criterion does not require consistency with the Savage (1954) axioms. This is important because a great deal of experimental study of choice under risk involves observed violations of Expected Utility Theory (EUT). Camerer (1995) and Starmer (2000) review the experimental and theoretical work that focuses on evaluating non-EUT theories.

a graphical interface introduced by Choi et al. (2007a) and used by Choi et al. (2007b). Because the design is user-friendly, it is possible to present each subject with *many* choices in the course of a short experimental session, yielding a much larger data set than has been possible in the past. This allows us to analyze the data at the level of the individual subject rather than pooling data or assuming homogeneity across subjects. Because choices are from standard budget lines, we can use revealed preference tests to investigate the extent to which the data comply with utility maximization. Because we observe many choices over a wide range of budget lines, the data allow high-power tests of revealed preference conditions.

By combining the capacity of our experimental setup with the CentERpanel, we provide three types of analysis:

[1] We offer a purely descriptive overview of the average quality of decisions by evaluating the consistency of individual behaviors with the utility maximization model.

[2] We then move to a regression analysis of the relationship between decision-making quality and sociodemographic and economic characteristics. In this way we address the question: "who is (more) rational?"

[3] Finally, we connect the insights that we gain from the experimental study under laboratory conditions to practical questions concerning wealth differentials in the real world.

If we follow Varian's (1991) suggestion of a threshold of 0.95 for the CCEI, we find that 45.2 percent of our subjects' scores are above this threshold, and of those, 22.8 percent have no violations of GARP (a CCEI score of 1). To calibrate the CCEI scores we compare the behavior of our actual subjects to the behavior of simulated subjects whose payoffs are perturbed by small idiosyncratic preference shocks. We conclude that the choices of many subjects come close to satisfying GARP in the sense that their violations are small enough to be attributed to the effect of a "trembling hand." Nevertheless, over all subjects, the CCEI scores averaged 0.881, which implies that subjects on average waste as much as 12.0 percent of their earnings by making inefficient choices. There is also marked heterogeneity in the CCEI scores within and across the sociodemographic characteristics of our subjects. Figure 1 summarizes the mean CCEI scores and 95 percent confidence intervals across selected socioeconomic categories.[2] Alternative measures of GARP violations based on Varian (1990, 1991) and Houtman and Maks (1985) (HM) yield qualitatively similar conclusions.

[2] In contrast, Choi et al. (2007b) report that 60 of their 93 subjects (64.6 percent) had CCEI scores above the 0.95 threshold and that over all subjects the CCEI scores averaged 0.937 (see Figure 1). The subjects of Choi et al. (2007b) were undergraduate students and staff at UC Berkeley, and the experiment was conducted at the Experimental Social Science Laboratory (Xlab) at UC Berkeley. We note that the subjects of Choi et al. (2007b) were given a larger and richer menu of budget sets which provides more opportunities to violate GARP.

[Figure 1 here]

We next move to studying, more systematically, the correlations between goodness-of-fit indices and sociodemographic and economic characteristics. Using Heckman's (1979) sample-selection model, we control for the possibility of sample-selection bias. Since the recruitment of CentERpanel members to experiments is random by construction, our instrumental variable is the number of completed CentERpanel questionnaires as a fraction of the total invitations to participate in the three months preceding our experiment.

Our main findings are that high-income and high-education subjects display greater levels of consistency than lower-income and lower-education subjects. Additionally, men are more consistent than women, and young subjects tend more toward utility maximization than those who are old. The magnitudes are large, implying, for example, that low-income subjects on average waste as much as 3.3 percentage points more of their earnings by making inefficient choices relative to high-income subjects. The corresponding numbers for low-education subjects, females, and old subjects are 2.6, 2.4, and 5.1, respectively.

Beyond consistency, we can also ask whether choices can be reconciled with a utility function with some normatively appealing properties. In decision-making under uncertainty, it is natural to ask whether choices are also consistent with the *dominance principle* in the sense of Hadar and Russell (1969)–that is, the requirement that an allocation should be preferred to another, regardless of subjects' risk attitudes, if it yields unambiguously higher monetary payoff. The dominance principle is compelling and generally accepted in decision theory.[3] Overall, the choices made by subjects in our experiment show low rates of stochastic dominance violations, which decrease with education level and increase with age.

Finally, we examine whether our proposed measure of decision-making quality – the consistency of the experimental data with the utility maximization hypothesis – is useful in explaining household wealth differentials. Wealth accumulation is determined by countless individual decisions, made over time in many different environments, and involving a host of different tradeoffs concerning risk, time, and personal and social consumption. The task of explaining wealth differentials therefore provides a particularly strong test of the predictive ability of our measure of decision-making quality. The test is relatively strong because it does not just examine the power of behaviors in the laboratory to predict related choices in a particular natural decision environment. Instead, it evaluates the ability of a parsimonious measure of decision-making quality (revealed preferences in the laboratory) to predict a broad economic outcome (wealth accumulation). The advantage of this approach is

[3] As noted by Quiggin (1990) and Wakker (1993), theories of choice under uncertainty were amended to avoid violations of dominance.

that it provides a parsimonious description of the laboratory data and it increases our chance of rejecting the correlation, in Popper's sense.

We are also motivated to consider wealth accumulation by several studies that document large wealth differentials among households with similar life-time income. The extent to which these differentials can be explained either by standard observables, such as family structure or income volatility, or by standard unobservables, such as risk tolerance or intertemporal substitution, is a subject of some debate (see, Bernheim et al., 2001, Ameriks et al., 2003, and Scholz et al., 2006, for different perspectives). If heterogeneity in decision-making quality is an important determinant of wealth differentials, and if consistency with individual utility maximization in the experiment is a good proxy for decision-making quality, then the patterns of behavior that we observe in the laboratory should help explain differential patterns of wealth. Our data are particularly well-suited to such investigations, given the heterogeneity in our experimental outcomes, and the heterogeneity in our subject pool.

We find an economically large and statistically significant association between the CCEI and household wealth. The point estimates indicate that, conditional on measures of current income, education, basic sociodemographics and household structure, a standard deviation increase in the CCEI of the person who is primarily responsible for household financial matters is associated with 15-19 percent more household wealth. This result is little changed when we add controls for violations of the dominance principle, a summary measure of risk tolerance, and a standard conscientiousness test.[4] The point estimates indicate that the latter two measures are related to wealth in anticipated ways, but neither relationship is statistically distinguishable from zero. We find little evidence that the CCEI is capturing unobserved aspects of education or past income. We do, however, find evidence that those with higher CCEI scores are substantially more likely to own a home and they put a larger fraction of their wealth into their home. This is important because the favorable tax treatment of owner-occupied housing in the Netherlands gives home ownership important advantages over renting and over investing in other assets with similar market returns. In this way, owning a home and placing more wealth in mortgaged housing are often high quality financial decisions. Replacing the CCEI with consistency measures suggested by Varian (1990, 1991) and HM yields similar conclusions.

There are many important questions that remain to be explored using this data set. In work in progress, we use the same data to relate findings on individual-level risk attitudes from the

[4] In psychology, conscientiousness is one of the "Big Five" factors (the other four are openness, extraversion, agreeableness, and neuroticism) of personality which are commonly used to measure human personality. Barrick and Mount (1991) conclude that conscientiousness appears to be the best predictor of economic outcomes. For a recent comprehensive discussion, see Block (2010).

experimental data with economic information and sociodemographic information on individuals. We explain heterogeneity in preferences and in types of behaviors in terms of sociodemographic variables and investigate the correspondence between individual investment and savings decisions and behavior in the laboratory. This enhances our understanding of important economic decisions such as savings and portfolio allocation, but it distracts from our fundamental purpose in this paper.

The rest of the paper is organized as follows. Section 2 further discusses the notion of decision-making quality. Section 3 describes the experimental design and procedures. Section 4 organizes the experimental data. Section 5 contains the analysis of the relationship between decision-making quality and sociodemographic and economic characteristics. Section 6 discusses the correspondence between decision-making quality and wealth differentials. Section 7 describes the margins along which we extend the previous literature. Section 8 contains some concluding remarks.

2 Decision-making quality

2.1 Identification and measurement

A relatively new empirical literature considers heterogeneity in choices driven by differences in the quality of decision-making. This literature allows that the choices that some individuals actually make may be different from the choices they would make if they had the skills or knowledge to make better decisions. Research in this vein considers the possibility that, even if they have all relevant information, individuals might not have the ability to identify and make the choice that best meets their objectives.

Ameriks et al. (2003) is a prominent example. That paper provides evidence that differences in individuals' propensity to plan and budgeting behaviors, rather than more standard sources of heterogeneity, explain important variation in wealth accumulation. In another example, Bernheim and Garrett (2003) find evidence that employer-based financial education increases saving. Agarwal et al. (2009) show a U-shaped age pattern in the frequency of dominated choices regarding the use of credit, with both relatively young and old consumers more prone to error.[5]

[5] Restricting attention just to the quality of *financial* decision-making, this literature also includes, among others, Duflo and Saez (2003) who investigate the effect of financial education on saving, beyond its effect on lifetime earnings; Lusardi and Mitchell (2007) document very low levels of financial planning, financial literacy, and a positive correlation between literacy, financial planning and wealth; and Cole and Shastry (2009) emphasize the importance of education, cognitive ability and financial literacy on financial market participation. Most recently, published in the November 2010 issue of the *Economic Journal* (Volume 120, Issue 548), Banks (2010) summarizes the research on the relationships between cognitive function, financial literacy and financial outcomes at older ages; Smith et al.

The possibility that planning skills, financial education, experience or cognitive decline substantially affect the quality of decision making is important because it suggests there are circumstances when "revealed" preferences may not be "true" preferences. If so, then positive predictions and welfare conclusions based on the "revealed" preferences may be misleading.

Through the collection of uncommonly high quality data, or the exploitation of natural experiments or instrumental variables, the research in this new literature has provided convincing evidence of important differences in decision-making quality. In general, however, definitive judgement about decision-making quality is made difficult by twin problems of *identification* and *measurement*.

> **Identification** The identification problem is to distinguish differences in decision-making quality from unobserved differences in preferences, information, beliefs or constraints.

In observational data, it is usually unclear whether those with less (financial) education, or lower cognitive abilities, fewer planning skills or less financial literacy are making lower quality decisions as opposed to holding different beliefs, or having different preferences over the same outcomes, or facing different incentives and constraints. Moreover, the identification problem presupposes a measurable notion of decision-making quality. In some cases the relevant incentives are sufficiently clear and data quality is sufficiently high, so that regarding some decisions as higher quality is straightforward and uncontroversial. More generally, a measure of decision-making quality is difficult to formalize, quantify and to make practical and portable for use in a variety of choice environments. These features of a measure are especially important to the extent that decision-making quality is a trait–a characteristic of a person that affects decisions in many different contexts.

> **Measurement** The measurement problem is to define and implement a portable, practical, autonomous, quantifiable, and economically interpretable measure of decision-making quality.

2.2 The revealed preference criterion

In this paper, we measure aspects of decision-making quality by the compliance of choices with economic rationality. In his Foundations of Economic Analysis (1947), Paul Samuelson offered a natural criterion for decision-making quality based solely on observable behavior. Adopting Samuelson's approach, we will say that choices are lower quality if there is no well-defined (utility) function that the choices maximize. Classical revealed preference theory provides a direct test: choices are consistent with maximizing a utility function if and only if they satisfy the GARP.

(2010) and Banks et al. (2010) show that wealth and retirement saving patterns are associated with numerical and other cognitive abilities at middle and older ages; and Van Den Berg et al. (2010) and Jappelli (2010) explore some of the potential causes of the differences in cognitive function and financial literacy in later life.

Since GARP imposes the complete set of conditions implied by utility-maximization, the CCEI and other goodness-of-fit indices provide a stringent test of decision-making quality.

The primary contribution to method of this work is an experimental technique that allows for the collection of richer data about preferences than has previously been possible and can easily be adapted to a wide range of decision-making experiments in large-scale surveys. As a result, the entire apparatus developed here – analytical and experimental techniques – has a number of distinctive features that make it useful for evaluating the quality of economic decision-making:

- **Portable** The analytical techniques and experimental platforms are applicable to many other types of individual choice problems involving personal and social consumption. They can thus make domain-specific predictions and provide a *unified* measure of decision-making quality across domains.

- **Practical** In the real world, changes in income and relative prices are such that budget lines do not cross frequently. This means that market data typically lack power to test revealed preference conditions (Blundell et al., 2003). Our subjects were given a large and rich menu of budget sets which leads to high power tests.

- **Autonomous** Consistency with GARP is not affected by preferences and the experimental task we study makes no special demands of outside knowledge or expertise, thus helping to isolate heterogeneity in decision-making quality from heterogeneity in preferences, information, beliefs or constraints (the identification problem).

- **Quantifiable** The CCEI (and other goodness-of-fit indices) measures the *extent* of GARP violation. In contrast with hypothetical (and unincentivized) survey questions, we can understand the results in terms of economic theory, which helps us interpret (as well as design) the experiments.

- **Interpretable** The CCEI has a coherent economic interpretation. Because of our rich data set, we are able to generate fairly tight bounds on the CCEI and use these bounds to judge the effects of decision-making quality.

A last feature of the apparatus developed here is that the method and measure may be evaluated for their ability to predict important behavior in "the wild." In this paper we consider whether the CCEI as measured in the experiment can help explain heterogeneity in the wealth holdings of households with similar lifetime incomes. Using different surveys and populations, the same basic methods could be used to investigate whether heterogeneity in this measure of decision-making quality can help explain heterogeneity in consumption choices, insurance purchase, retirement decisions, or health investments among other important behaviors.

3 Experimental design

3.1 Sample

The experiment uses the CentERpanel, an online, weekly, stratified survey of a sample of over 2,000 households and 5,000 individual members. The panel is designed to be representative of the Dutch-speaking population in the Netherlands. The CentERpanel thus provides a unique opportunity to combine experimental data with sociodemographic and economic variables from the survey. The subjects in the experiment were randomly recruited from the entire CentERpanel body. The experiment was conducted *online* under the CentERdata protocol with 1,182 CentERpanel adult members, using the experimental technique introduced by Choi et al. (2007a) and used by Choi et al. (2007b).[6] The experimental method allows us to identify individual behaviors that may be related to a wide range of individual characteristics.

Table 1 below provides summary statistics of individual level characteristics. We present the data for *participants* (completed the experiment), *dropouts* (logged in and quit the experiment) and *nonparticipants* (recruited for the experiment but never logged in). We use six socioeconomic categories: gender, age, completed education, household income, occupation, and household composition. The groupings of different levels of completed education are based on the categorization of Statistics Netherlands (Centraal Bureau voor de Statistiek). The low, medium and high education levels correspond to primary or pre-vocational secondary education, pre-university secondary education or senior vocational training, and vocational college or university education, respectively. We use household monthly gross income-level categories such that the proportions of participants in each category are approximately equal. The classification of occupations is based on the categorization of Statistics Netherlands.

[Table 1 here]

[6] CentERdata is an independent research institute affiliated with the Tilburg School of Economics and Management (TiSEM) at Tilburg University in the Netherlands. CentERdata specializes in online experiments and manages the CentERpanel and several other panels. The panel members complete the questionnaires over the Internet. For more information, see http://www.centerdata.nl/en/centerpanel.

3.2 Procedures

Our experimental interface was incorporated into the CentERpanel and the experiment was hosted as part of their survey. In our experiment, we presented subjects with several decision problems under risk. Each decision problem was presented as a choice from a two-dimensional budget line. A choice of the allocation (x, y) from the budget line represents an allocation between accounts x and y (corresponding to the horizontal and vertical axes). The actual payoffs of a particular choice were determined by the allocation to the x and y accounts such that the subject received the points allocated to one of the accounts x or y, determined at random and equally likely. Choices were made using the computer mouse or the keyboard arrows to move the pointer on the computer screen to the desired point and then clicking or hitting the enter key.[7] Payoffs were calculated in terms of points and then converted into euros. Each point was worth €0.25. Subjects received their payment from the CentERpanel reimbursement system via direct deposit into a bank account.

The procedures described below are identical to those used by Choi et al. (2007b), with the exception that the experiment described here consisted of 25, rather than 50, decision problems and that there were some minor additional changes resulting from the online experimental setting.[8] 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. This variation in budget lines (prices and incomes) is essential for a thorough test of consistency (more below). The budget lines selected for each subject in different decision problems were independent of each other and of the sets selected for any of the other subjects in their decision problems. Choices were restricted to allocations on the budget constraint.[9] During

[7] The experimental method is applicable to other types of individual choice problems. Ahn et al. (2010) extended the work of Choi et al. (2007b) on risk (known probabilities) to settings with ambiguity (unknown probabilities). Fisman et al. (2007, 2010) employ a similar platform to study distributional preferences and produce very different behaviors. It is possible that presenting choice problems graphically biases choice behavior in a particular way—and that is a useful topic of experiment—but there is no evidence that this is the case, as emphasized by Choi et al. (2007b) and Fisman et al. (2007).

[8] The number of individual decisions is still higher than is usual in the literature, and the experiments provide us with a rich data set consisting of enough individual decisions over a wide range of budget lines to provide a powerful test of consistency. The revealed preference analysis presented below shows that the variation in budget lines (prices and incomes) is sufficient for a rigorous test of consistency.

[9] Like Choi et al. (2007b), we restricted choices to allocations on the budget constraint so that subjects could not dispose of payoffs. In Fisman et al. (2007), each choice involved choosing a point on a graph representing a *budget set* over possible allocations. Since most of their subjects had no violations of budget balancedness (those who did violate budget balancedness also had many GARP violations even among the subset of their choices that were on

the course of the experiment, subjects were not provided with any information about the account that had been selected in each round. At the end of the experiment, the computer selected one decision 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.

The resolution of the budget lines allow subjects to adjust their allocations by amounts as small as 0.2 points. At the beginning of each decision round, the experimental program dialog window went blank and the entire setup reappeared. The appearance and behavior of the pointer were set to the Windows mouse default and the pointer was automatically repositioned randomly on the budget line at the beginning of each round. Full experimental instructions, including the computer program dialog windows, are also available at Online Appendix I.[10]

4 Data description

We next provide an overview of some basic features of the individual-level data. Without loss of generality, assume the individual's income is normalized to 1. The budget set is then $p_1x_1+p_2x_2 = 1$ and the individual can choose any allocation x that satisfies this constraint. Let $\{(p^i, x^i)\}_{i=1}^{25}$ be the data generated by an individual's choices, where p^i denotes the i-th observation of the price vector and x^i denotes the associated allocation.[11]

4.1 Consistency

Following Afriat's (1967) theorem, we employ the Generalized Axiom of Revealed Preference (GARP) to test whether the finite set of observed price and quantity data that our experiment generated may be rationalized by a utility function $U(x_1, x_2)$.[12] GARP (which is a generalization of various other revealed preference tests) requires that if x^i is indirectly revealed preferred to x^j, then x^j is not *strictly* and directly revealed preferred ($p^jx^i \geq p^jx^j$) to x^i. The theory tells us that if the data satisfy GARP, then a utility function that rationalizes the observed allocations exists and, moreover, may be chosen to be *well-behaved* (piecewise linear, continuous, increasing, and

the budget constraint), we restricted choices to allocations on the budget constraint, which simplified the decision problem and made the computer program easier to use.

[10] Online Appendix I: http://emlab.berkeley.edu/~kariv/CKMS_I_A1.pdf.

[11] More precisely, the data generated by an individual's choices are $\{(\bar{x}_1^i, \bar{x}_2^i, x_1^i, x_2^i)\}_{i=1}^{25}$, where (x_1^i, x_2^i) are the coordinates of the choice made by the subject and $(\bar{x}_1^i, \bar{x}_2^i)$ are the endpoints of the budget line, so we can calculate the budget line $x_1^i/\bar{x}_1^i + x_2^i/\bar{x}_2^i = 1$ for each observation i.

[12] Varian (1982, 1983) modifies Afriat's (1967) results and describes efficient and general techniques for testing the extent to which choices satisfy GARP.

concave).[13]

Although testing conformity with GARP is conceptually straightforward, there is an obvious difficulty: GARP provides an exact test of utility maximization – either the data satisfy GARP or they do not. We assess how nearly individual choice behavior complies with GARP by using Afriat's (1972) CCEI, which measures the fraction by which each budget constraint must be shifted in order to remove *all* violations of GARP. If the CCEI is close to one, the subject is wasting very little of his earnings. Otherwise, he may be wasting quite a lot. In this sense the CCEI measures the overall "efficiency" of individual behavior.

Put precisely, for any number $0 \leq e \leq 1$, define the direct revealed preference relation

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

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. The CCEI is the value of e^* associated with the data set $\{(p^i, x^i)\}_{i=1}^{25}$. By definition, the CCEI is between zero and one—indices closer to one mean the data are closer to perfect consistency with GARP and hence to perfect consistency with utility maximization—and can be interpreted as saying that the individual is wasting as much as $1 - e^*$ of the income by making inefficient choices. Hence, the CCEI may overstate the extent of inefficiency, but the above procedure is the "least costly" adjustment for removing all violations of GARP.

We provide more details on testing for consistency with GARP and discuss the alternative indices that have been proposed by Varian (1990, 1991) and HM in online Appendix II.[14] In reporting our results, we focus on the CCEI, which offers a straightforward interpretation. In practice, all these measures yield similar conclusions. The tables based on the indices proposed by Varian (1990, 1991) and HM are presented in Online Appendix III.[15,16]

Table 2 below provides a summary of the individual-level CCEI scores. We report the statistics for all subjects, as well as the statistics by socioeconomic categories. The CCEI scores averaged 0.881 over all subjects, and ranged from 0.920 for subjects younger than 35 to 0.843 for subjects age 65 and older. There is also considerable heterogeneity within and across categories. The analysis

[13] Satisfying GARP entails only that choices are consistent with the utility maximization model. The further implication, that the choices may be rationalized by a well-behaved utility function, is a consequence of linear budget constraints. Given that the budget constraints are linear, if a rationalizing utility function exists then we cannot reject the hypothesis that it is well-behaved.

[14] Online Appendix II: http://emlab.berkeley.edu/~kariv/CKMS_I_A2.pdf.

[15] Online Appendix III: http://emlab.berkeley.edu/~kariv/CKMS_I_A3.pdf.

[16] All indices are computationally intensive for even moderately large data sets. We compute the Houtman-Maks scores using the algorithm developed by Dean and Martin (2010). (The computer program and details of the algorithms are available from the authors upon request.)

of the relationship between the differences in consistency scores and sociodemographic differences among experimental subjects is the purpose of our regression analysis below.[17]

[Table 2 here]

4.2 Power and goodness-of-fit

Revealed preference tests have an important drawback: there is no natural threshold for determining whether subjects are close enough to satisfying GARP that they can be considered utility maximizers. Varian (1991) suggests a threshold of 0.95 for the CCEI. If we follow Varian's (1991) suggestion, we find that out of the 1,182 subjects, 534 subjects (45.2 percent) have CCEI scores above this threshold and of those 269 subjects (22.8 percent) have no violations of GARP.

To generate a benchmark against which to compare our 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 0.881 whereas the mean CCEI score for a random sample of 25,000 simulated subjects is only 0.659. Moreover, more than half of actual subjects have CCEI's above 0.925, while only about five 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 (more below). The setup used in this study has the highest Bronar 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. In the interest of brevity, the analysis has been relegated to Online Appendix II.[18]

4.3 Beyond consistency

Choices can be consistent with GARP and yet fail to be reconciled with any utility function that is normatively appealing given the decision problem at hand. For example, consider choices in the experiment that always allocate all points to x_1. This behavior is consistent with maximizing

[17] To allow for small trembles resulting from the slight imprecision of subjects' handling of the mouse, our consistency results allow for a narrow confidence interval of one point (that is, for any i and $j \neq i$, if $|x^i - x^j| \leq 1$ then x^i and x^j are treated as the same portfolio).

[18] Andreoni and Harbaugh (2006) develop power indices for revealed preference tests based on CCEI and discuss the prior indices of Bronars (1987) and Famulari (1995).

the utility function $U(x_1, x_2) = x_1$ and would generate a CCEI score of 1. Such preferences are, however, hard to justify given that the commodity in each state is the same (money), and the realization of the state affects nothing except which account pays off. Moreover, for many of the budget lines that a subject faces, allocating all points to x_1 means allocating all points to the more expensive asset, a violation of monotonicity with respect to first-order stochastic dominance.

Violations of first-order stochastic dominance may reasonably be regarded as errors, regardless of risk attitudes—that is, as a failure to recognize that some allocations yield payoff distributions with unambiguously lower returns. A simple violation of dominance is illustrated in Figure 2 below. The budget line is defined by the straight line AE and the axes measure the future value of a possible allocation in each of the two states. The point B, which lies on the 45 degree line, corresponds to an allocation with a certain outcome. The individual chooses allocation x (position along AB), but could have chosen any allocation x' (position along CD) such that $F_{x'} \leq F_x$ where $F_{x'}$ and F_x are the resulting payoff distributions. If this individual only cares about the distribution of monetary payoffs, then he will be willing to pay a positive price for a portfolio yielding $F_{x'} - F_x$, which has only nonpositive payoffs (that is, for a portfolio in which each asset had an equal probability of being chosen). Notice that any decision to allocate *fewer* points to the *cheaper* asset (that is, corresponding to a position along AB) violates dominance but need not involve a violation of GARP, whereas any decision to allocate *more* points to the *cheaper* asset (that is, corresponding to a position along BE) never violates dominance.

[Figure 2 here]

If subjects identified an allocation with the resulting probability distribution over payoffs then preferences would satisfy the *reduction principle*; that is, $(x_1, x_2) \sim (x_2, x_1)$ because they generate the same payoff distribution. If preferences satisfy the reduction principle then, subject to every budget constraint, choices would allocate weakly more points to the cheaper asset. We would like to test whether preferences satisfy the reduction principle by observing choices from linear budget sets. Unfortunately, this is not possible: choices from linear budget sets determine the demand function but the demand function does *not* uniquely determine preferences (Mas-Colell, 1977; 1978). However, *symmetry* provides implications about choices from linear budget sets (that is, about demand functions) that are testable on the basis of observed choices from standard budget sets.

We identify choice behavior as *symmetric* if (x_1^*, x_2^*) is chosen subject to the budget constraint $p_1x_1 + p_2x_2 = 1$ if and only if (x_2^*, x_1^*) is chosen subject to the *mirror-image* budget constraint $p_2x_1 + p_1x_2 = 1$. That is, choice behavior responds symmetrically to inverse price ratios.[19] Clearly,

[19] The reduction principle implies that choice behavior is symmetric only when the derived demand function is single

if choice behavior is symmetric then the choice subject to every budget constraint allocates more points to the cheaper asset. Hence, symmetry imposes restrictive (if convenient) patterns on demand behavior, but it is a natural result of symmetric probabilities (each account had an equal probability of being chosen).

To test whether choice behavior is symmetric (for a given subject), we can combine the actual data from the experiment and the mirror-image data, compute the CCEI for this combined data set, and compare that number to the CCEI for the actual data.[20] By definition, the CCEI score for the combined data set consisting of 50 observations can be no bigger than the CCEI score for the actual data. Clearly, always allocating all points to one of the assets generates severe violations of GARP in the combined data set, but the subset of actual data is perfectly consistent.[21] Similarly, any decision to allocate fewer points to the cheaper asset will necessarily generate a simple violation of the Weak Axiom of Revealed Preference (WARP) involving its mirror-image decision.

Thus, we can construct a formal non-parametric test of symmetric behavior by following the strategy above: compute the CCEI for the combined data set and compare that number to the CCEI for the actual data set. The difference reflects an upper bound on the additional income that the subject is wasting by not always allocating more points to the cheaper asset. Nevertheless, the combined data set obviously provides a more stringent test of GARP so it can contain new violations of GARP even if actual choices always allocated more points to the cheaper asset.

If we again follow Varian's (1991) suggestion of a threshold of 0.95 for the CCEI, we find that in the combined data set the scores of 251 subjects (21.2 percent) are above this threshold and of those only 24 (2.0 percent) have no violations of GARP. Table 3 below reports summary statistics and percentile values of the CCEI scores for the combined data set. We report the statistics for all subjects, as well as the statistics by socioeconomic categories. The last column lists the difference between the mean CCEI's for the actual data set and for the combined data set. The CCEI scores for the combined data set averaged only 0.733 over all subjects, and ranged from 0.786 for subjects younger than 35 to 0.679 for subjects age 65 and older, representing a decrease from the CCEI scores for the actual data set of 0.148, 0.134 and 0.165, respectively. Overall, a sociodemographic category that had a lower mean actual CCEI score exhibits a larger decrease. In our econometric analysis below, we use both the CCEI scores for the actual data set and for the combined data set.

[Table 3 here]

---

valued. GARP is also compatible with multi-valued demand functions so preferences may not be strictly convex.

[20] The data generated by an individual's choices are $\left\{\left(\bar{x}_1^i, \bar{x}_2^i, x_1^i, x_2^i\right)\right\}_{i=1}^{25}$ and the mirror-image data are obtained by reversing the prices and the associated allocation for each observation $\left\{\left(\bar{x}_2^i, \bar{x}_1^i, x_2^i, x_1^i\right)\right\}_{i=1}^{25}$.

[21] Of the 1,182 subjects in the experiment, only 29 subjects (2.5 percent) almost always allocated all points to one of the assets by choosing the same endpoint of the budget line.

4.4 Risk attitudes

We summarize attitudes toward risk with a univariate measure, which we will use to capture risk aversion in the regression analysis concerning wealth differentials.[22] Because the experiment is symmetric and budget lines are drawn from a symmetric distribution, we summarize the risk attitudes of our subjects by reporting the fraction of points allocated to the cheaper asset. The only behavior consistent with infinite risk aversion is always allocating the points equally between the two assets. On the other hand, always allocating all points to the cheaper asset is the behavior that would be implied by pure local risk neutrality. In general, subjects who are less averse to risk will allocate a larger fraction of points to the cheaper asset.

Figure 3 below displays the mean fraction of points allocated to the cheaper asset and 95 percent confidence intervals across the socioeconomic categories. Note that there is considerable heterogeneity in risk attitudes across categories, which is characteristic of all these data, and that risk attitudes and CCEI scores are modestly correlated ($r^2 = 0.113$). The distributions of the fraction of points allocated to the cheaper asset are quite similar across categories: all have a mode near the midpoint of 0.5, and the distribution falls off sharply away from midpoint, with mass concentrated to the right. There have been many attempts to recover risk preferences from subjects' decisions in a variety of laboratory experiments. A significant fraction of our subjects exhibit moderate to high levels of risk preferences. As Figure 3 shows, our individual-level measures of risk aversion are higher than the measures reported in Choi et al. (2007b), but they are within the range of recent estimates from recent studies (See Choi et al., 2007b, for a discussion of these studies).

[Figure 3 here]

5 Decision-making quality and sociodemographics

The relatively large and heterogeneous CentERpanel sample and accompanying survey data allow us to perform what is, to our knowledge, the first analysis of the correlation between sociodemographic and economic characteristics and GARP violations. Table 4 below presents the results of our individual-level econometric analysis. In column (1), we present estimates with the CCEI

[22] In work in progress with these data, we build on Choi et al. (2007b) to estimate preferences using a two-parameter utility function based on Gul (1991) – one parameter is the familiar coefficient of risk aversion and the other is a measure of loss/disappointment aversion – and we relate the individual-level estimates to individual characteristics and external choices.

scores for the actual data set using ordinary least squares.[23] The results show significant correlations. We obtain statistically significant coefficients in all sociodemographic categories, ranging in absolute values from about 0.025 to just over 0.050. These magnitudes are large, implying that sociodemographic differences can account for significant differential changes in income loss due to inconsistent choice patterns. Most notably, females, low-education, low-income, and old subjects on average waste as much as 2.4, 2.6, 3.3, and 5.1 percentage points more of their earnings, respectively, by making inefficient choices.[24] In columns (2) we repeat the estimation reported in columns (1) using the CCEI scores for the combined data set.[25] The corresponding estimates are of greater magnitude and statistically significant in the age and education categories.[26]

[Table 4 here]

Our analysis above is based on the non-randomly selected subsample of participants. The lack of observations on panel members who chose not to participate or did not complete the experiment creates a missing data problem. We correct for the possible sample selection bias in our econometric analysis below, using Heckman's (1979) method.[27] Our *exclusion restriction* rests on the number of completed CentERpanel questionnaires as a fraction of the total invitations to participate in the three months prior to our experiment entering the participation equation but not being conditionally correlated with rationality. Our identifying assumption is that this "participation ratio" influences the participation in our experiment but does not influence the laboratory outcomes of interest (Bellemare et al., 2008).

[23] To test for a potential misspecification, we used Ramsey's (1969) RESET test by adding the squared and cubed fitted values of the regression equation as additional regressors, and found no evidence of misspecification (p-value = 0.3098).

[24] Agarwal et al. (2009) document a U-shaped relationship between age and mistakes in financial decision making, suggesting that, although some cognitive abilities decline with age, experience in financial markets rises with it. We find that consistency with GARP and hence consistency with utility maximization decline dramatically over the lifecycle.

[25] We also separately examine the patterns of stochastic dominance violations in the data. Using OLS and Tobit, we obtain statistically significant coefficients only in the age and education categories. Using Heckman's (1979) method to correct for possible sample selection bias, only the age coefficients are negative and significant. Overall, subject choices generally satisfy the dominance principle. In the interest of brevity, we do not report this analysis here.

[26] Since the CCEI is a number between zero and one, we repeat the estimations reported in columns (1) and (2) using a fractional regression model (Papke and Wooldridge, 1996). The two specifications yield similar results.

[27] We also use Heckman's sample selection model to analyze the correlates of the Varian (1990, 1991) measure. For the third measure, proposed by HM, we estimate the sample selection model of Terza (1998). These results are provided in Online Appendix III.

The estimation results are reported in Table 5 below. In column (1), we omit the nonparticipants, focusing on the subsample of participants and dropouts in the data. In column (2), we repeat the estimation reported in column (1), after adding the nonparticipants. We obtain qualitatively similar results on the reduced sample and the entire sample. Finally, testing the null hypothesis that the correlation coefficient ρ is zero is equivalent to testing for sample selection. In columns (1) and (2), we find that ρ is indistinguishable from zero and thus we find no evidence of bias. We interpret these results to indicate that self-selection is not importantly driving the results. It is also noteworthy that in both specifications the coefficient on the exclusion restriction variable is positive and significant, and that many sociodemographic categories are significantly correlated with participation. In columns (3) and (4), we repeat the estimation reported in columns (1) and (2) using the CCEI scores for the combined data set and obtain similar results.

[Table 5 here]

6 Wealth differentials and decision-making quality

We next examine whether the extent to which individual decisions satisfy GARP – and are therefore consistent with the utility maximization model – is useful in explaining household wealth differentials. Conditional on income, wealth represents the accumulation of innumerable financial decisions. It is therefore natural to use wealth as a summary variable with which to evaluate the predictive ability of our proposed measure of decision-making quality. By studying the relationship between wealth and the consistency scores, we can also evaluate the role of decision-making quality in determining why households with similar life-time incomes and sociodemographic characteristics accumulate very different amounts of wealth (Bernheim et al., 2001, and Ameriks et al., 2003).[28] If heterogeneity in decision-making quality were an important determinant of heterogeneity in wealth, and if consistency in the experiment were a good proxy for financial decision-making quality, then differences across subjects should help explain differential patterns of wealth.

6.1 Wealth data

The CentERpanel collects information about wealth on an annual basis. Panel members are asked to identify a financial respondent who is "most involved with the financial administration

[28] Scholz et al. (2006) find that heterogeneity in decision-making quality is not necessary to explain much of the heterogeneity in wealth. They show that by using detailed data on household-specific earnings paths from the Health and Retirement Study (HRS), the optimal decision rules from a life cycle model can account for more than 80 percent of the cross-sectional variation in wealth. We will return to this issue in the econometric analysis (Table 8).

of the household." All panel members age 16 and older respond to questions about the assets and liabilities that they hold alone. The financial respondent also provides information about assets and liabilities that are jointly held by more than one member of the household. The inventory covers checking and saving accounts, stocks, bonds and other financial asset holdings, real estate, business assets, mortgages, loans, and extended lines of credit. The data do not include information on pension wealth.[29] Our analysis focuses on household wealth, calculated by summing net worth over household members and taking the household's average over 2008 and 2009. Summary statistics of this measure of wealth, and some of its components, are reported in Table 6 below. Table 6 summarizes the wealth of the 703 households with wealth data and a CCEI score from the household's financial respondent. The median household has a net worth of nearly €93K ($136K) and the distribution is positive-skewed. The mean household wealth of €164K ($272K) is considerably higher than the median and the highest values are more than 15 standard deviations above the mean.

[Table 6 here]

Anticipating our regression analysis, Table 7 reports some components of wealth, by the CCEI score of the household's financial respondent. More precisely, this panel gives the means of wealth components for all of the analysis sample, for those whose age-adjusted CCEI score is below the median, and for those whose age-adjusted CCEI score is above the median.[30] To reduce the importance of extreme outliers, we drop the 10 households that represent the top and bottom half of one percent of the wealth distribution and the bottom half of one percent of the CCEI distribution. The statistics reported in Table 7 show that housing wealth is a central component of wealth outside of pension systems. On average, the sample owns a house worth €184K ($324K) and holds a mortgage of €62K ($109K). Various savings accounts and stocks are the next two most important, though much smaller, components of wealth. The differences in means between households with higher and lower CCEI are never statistically significant, though some of them are economically large. For example, financial respondents with higher CCEI scores have an average household wealth that is €4,260 ($7,512) higher than those with lower CCEI scores. Restricting attention to households with positive net worth, the differences in the average log of household wealth is 0.16. These simple differences in means summarize many correlations, which we parse in

[29] Nearly all of the Dutch population is covered by the public pension system, and a large majority of workers is covered by private pensions associated with their employment. Nearly all of these employment-based plans are defined benefit, the vast majority of which pay benefits as a function of lifetime earnings. See Alessie and Kapteyn (2001) and OECD (2009) for details about the pension systems in the Netherlands.

[30] The age-adjusted CCEI is the residual from a regression of CCEI on a constant, age, age^2 and age^3.

the regression analyses below.

[Table 7 here]

6.2 Wealth regressions

To describe the relationship between consistency with utility-maximizing behavior in the laboratory and wealth differentials, our basic approach is to estimate regressions of the natural log of household wealth on sociodemographic variables (including a flexible function of age), the natural log of household income, and the CCEI score of the financial respondent in the household. The estimation results are reported in Table 8 below. The sample size drops from 703 to 566 household (80.5 percent). This decline is driven almost exclusively by 74 households (10.5 percent) with negative net worth and thus a missing dependent variable and 54 households (7.7 percent) with negative or missing household income in 2008. In addition, given our relatively small sample and the presence of extreme outliers, we also drop seven households that represent the union of the top and bottom half of one percent of the wealth distribution and the bottom half of one percent of the distribution of CCEI scores. Two additional households are dropped due to missing data on education.

[Table 8 here]

Baseline. In column (1), we present estimates from our main econometric specification using the entire analysis sample. The point estimate of 1.17 on the CCEI indicates that a standard deviation increase in the CCEI score of the household's financial respondent is associated with nearly 16 percent more household wealth. As one might expect from a relatively small sample of data on self-reported wealth, the standard error on this point estimate is fairly large. Nevertheless, we can reject a null hypothesis of no relationship at the 5 percent level (p-value=0.029) with standard errors robust to heteroskedasticity.

Life-cycle. At younger ages, those with better lifetime opportunities may have lower wealth as they borrow in order to invest or to smooth lifetime consumption. With that in mind, in column (2), we repeat the estimation reported in column (1) with the sample restricted to households with financial respondents who are at least 35 years old. We find that the point estimate on the CCEI is somewhat larger at older ages so a standard deviation increase in the CCEI score of the household's financial respondent is associated with about 19 percent more household wealth. The standard error on this point estimate is relatively large, so while we can reject a null hypothesis of no relationship with considerable confidence (p-value=0.012) we cannot reject a null hypothesis that the point estimates on the CCEI reported in columns (1) and (2) are the same.

Levels. The log specification restricts attention to those with strictly positive wealth. The specification may also cause small differences at positive but very low levels of wealth to have large effects on estimates. To evaluate the sensitivity of the results to the log specification, in column (3) we estimate the regression in levels (of wealth and income) on the sample ages 35 and older. We again see an economically large association between the CCEI and levels of wealth, though this relationship is not estimated as precisely; the coefficient on the CCEI is significant only at the 10 percent level (p-value=0.058).

Cognitive abilities and education. We interpret our CCEI scores as capturing aspects of decision-making quality. We take this view because choices that are closer to satisfying GARP can be seen as more purposeful; they reflect more consistent treatment of tradeoffs regardless of preferences, information or beliefs. An alternative view is that the CCEI captures unobserved aspects of cognitive ability or education that are correlated with financial outcomes through their correlation with preferences, information, beliefs or unobserved constraints. The CentERpanel survey does not include measures of IQ (and this is an important topic for future work), but we can assess whether *unobserved* aspects of education are driving the relationship between the CCEI and wealth if we assume that these unobserved variables are positively correlated with *observed* education levels.

If so, and if these unobserved variables are important sources of the observed correlation between consistency and wealth, then conditioning on observed education should have a substantial effect on the estimated coefficient on the CCEI. To this end, in column (4), we repeat the estimation reported in column (2) after omitting the controls for the formal education of the financial respondent. Comparing the estimates from columns (2) and (4), we see that removing the education controls has only a modest effect on the estimated coefficient on the CCEI. In this way, we find little evidence that the relationship between consistency with utility maximization in the laboratory and wealth is driven by a correlation between the CCEI and unobserved aspects of education.

Beyond consistency. Restricting attention to the sample ages 35 and older, in column (5), we repeat the estimation reported in column (2) after adding the CCEI scores for the combined data set (combining the actual data from the experiment and the mirror-image data). We find no evidence that, conditional on the CCEI score from the actual data, the CCEI score for the combined data set has an independent relationship with wealth. Adding the CCEI for the combined data set as a regressor has only a modest effect on the point estimate of the coefficient on the CCEI; and the point estimate of the conditional relationship between the CCEI for the combined data set and log household wealth is small, but imprecisely estimated. These results are consistent with the idea that the CCEI for the combined data set, while imposing natural requirements on demand behavior, merely represents a noisier measure of the aspects of decision-making quality captured by the CCEI scores for the actual data set.

Risk attitudes. In column (6) we add a control for risk attitudes by including the average fraction of points the financial respondent allocated to the cheaper asset.[31] The point estimate on this measure of risk attitudes indicates that risk tolerance is negatively associated with wealth. The estimate is economically large; a standard deviation increase in the fraction placed in the cheaper asset is associated with about 10 percent less household wealth. The estimate is fairly imprecise, however; we cannot reject a null hypothesis of no relationship (p-value=0.16). As one might expect given the relatively modest unconditional correlation between the CCEI scores and risk attitudes ($r^2 = 0.113$), the inclusion of the control for risk attitudes in the experiment leaves the point estimate of the coefficient on the CCEI virtually unaffected. When we omit the control for risk attitudes, the point estimate of the coefficient on the CCEI is 1.539 (p-value=0.009). The results are qualitatively similar when, in column (7), we condition on both risk tolerance and the CCEI scores for the combined data set.

Conscientiousness. In a final assessment of the relative magnitude of the correlation between the CCEI and household wealth, we add to the list of controls an influential measure from psychology, "conscientiousness." Conscientiousness is one of the "Big Five" personality traits, derived from factor analysis of wide-ranging personality surveys (see, Block, 2010, for a recent description and assessment of the "Big Five"). The CentERpanel survey contains 10 questions related to conscientiousness, each recorded on a scale from 1 to 5. For each respondent, we sum his or her responses to these 10 questions, and then normalize the scores to generate a measure with sample mean 0 and standard deviation 1.[32]

In column (8) we repeat the estimation reported in column (7), restricting attention to the sample of subjects for whom we have valid data on conscientiousness. In column (9) we add the

[31] To avoid the influence of extreme outliers, and to place this variable on equal footing, we again omit the top and bottom half of one percent of the distribution of this risk attitude measure, a total of 10 households (1.8 percent). The results are qualitatively similar when we include these households.

[32] The respondents are asked to evaluate the accuracy of 10 statements as descriptions of themselves. The statements include: "I do chores right away," "I am accurate in my work," and "I am always well prepared," among others. When the statement indicates a lack of conscientiousness, we re-ordered the responses so that higher scores reflect greater conscientiousness. No other element of the Big Five is included in the CentERpanel survey.

measure of conscientiousness to the list of controls. The magnitude of the coefficient on conscientiousness is large. A standard deviation increase in conscientiousness is associated with about 10 percent more wealth. (In this sample, a standard deviation increase in the CCEI is associated with approximately 22 percent more wealth.) The standard error on the conscientiousness coefficient is relatively large, however, and we cannot reject a null hypothesis of no correlation. Most important, adding the control for conscientiousness has very little effect on the coefficient on the CCEI, which suggests that the conditional correlation between the CCEI and wealth is unlikely to be driven by an unmeasured correlation between the CCEI and *productive* personality traits. We emphasize the term productive to highlight that, among the Big Five factors, conscientiousness appears to be the best predictor of economic performances (see, Barrick and Mount, 1991).

Finally, the estimates based on the alternative measures of GARP violations of Varian (1990, 1991) and HM yield qualitatively similar conclusions. In the case of the HM index, the standard errors are relatively small and the opposite is true of the Varian (1990, 1991) index (see Online Appendix III).

6.3 Sources of correlation

We find an economically large and statistically significant correlation between household wealth and the financial respondent's CCEI score in the experiment. We show that this correlation is robust to the inclusion of controls for current income, education, family structure, risk attitudes in the experiment, and a widely-used personality test. With these results in mind, we now turn to investigate further the sources of this correlation. The estimation results are reported in Table 9 below. We again focus on the CCEI, but the estimates based on the alternative measures of GARP violations of Varian (1990, 1991) and HM yield qualitatively very similar results (see Online Appendix III). To reduce the importance of extreme outliers, in columns (3)-(10) we drop households whose fraction of wealth in the relevant category (checking, saving, stocks, housing) is less than -0.15 or greater than 1.15.

[Table 9 here]

Income. One hypothesis is that the correlation between consistency with utility maximization in the experiment and wealth derives from a correlation between the CCEI and unobserved (past) income. Our finding that adding controls has only a modest effect on the CCEI coefficient casts some doubt on this concern. We can evaluate it further, however, by exploiting some limited panel data on household income. The CentERpanel has been operating since 1993. However, income data for most households who responded to the 2009 survey and completed our experiment do not go back nearly that far. To strike a balance between capturing more income information and maintaining reasonable sample sizes, we go back five years and use household income information for every other year.

In column (1) we repeat the estimation reported in column (2) of Table 8, this time restricting attention to the 377 households (73 percent) for whom we have household income data from 2004 and 2006, as well as from 2008. In this smaller sample, the point estimate on the CCEI remains economically large and statistically significant (p-value=0.012). In column (2), we add controls for the natural log of household income in 2004 and 2006. As a result, the magnitude of the coefficient on the CCEI declines somewhat (by 0.115). We interpret this finding to indicate that some of the correlation between wealth and the CCEI may be attributable to a correlation between the CCEI and unobserved past income. The CentERpanel income data limit our ability to pursue this hypothesis further; we therefore view it as an important topic for future research with other data.

Portfolio. An alternative set of hypotheses would attribute the correlation between wealth and the CCEI to differences in financial decisions between those with higher and lower CCEI scores. To evaluate these hypotheses, in columns (3)-(8), we present estimates that relate the CCEI of the household's financial respondent to portfolio choices, specifically, whether the household has a checking account, a savings account, owns stocks, and the fraction of the household's wealth held in each of these assets. The results provide some evidence that, conditional on household sociodemographics, education, and current income, households with financial respondents with higher CCEI scores put less of their wealth in low-risk and low-return assets such as checking and savings accounts. The coefficients on the CCEI in columns (4) and (6) are quite modest in magnitude but statistically significant at the 10 percent level (p-values of 0.078 and 0.062, respectively). The results also provide some evidence that individuals with higher CCEI scores are somewhat more likely to participate in the stock market, though this relationship is not statistically significant.

Housing. Finally, the coefficients on the CCEI in columns (9) and (10) show economically substantial and statistically significant correlations between the CCEI and decisions regarding home ownership. Households with financial respondents with higher CCEI scores are more likely to own a home and they put a larger fraction of their household's wealth in a home. A standard deviation increase in the CCEI of the financial respondent is associated with an increase of 5.2 percentage points in the probability of owning a home (about 69 percent of households in the sample own a home). Similarly, a standard deviation increase in the CCEI is associated with an increase of 4.7 percentage points in the fraction of wealth held in housing (the average fraction of wealth held in housing in the data is approximately 54 percent). The tendency for those with higher CCEI scores to own a home and put more of their wealth in housing is especially interesting given the favorable tax treatment of owner-occupied housing in the Netherlands, which indicates both that owning a home has an important advantage over renting and that, other things equal, wealth placed in mortgaged housing pays a substantial premium.[33] This suggests that owning a home and placing more wealth in mortgaged housing are often high quality financial decisions. If so, the positive correlation between the CCEI and these decisions is what we would expect if the CCEI captured a general tendency toward higher quality financial decision-making.

7 Related literature

Our paper relates to several strands of prior research. One strand is the large and rapidly growing literature that relates measures of cognitive ability and economic literacy to important economic behaviors in "the wild." Examples include Lusardi and Mitchell (2007), Fang et al. (2008) and Banks et al. (2010). Lusardi (2008) and the papers included in the November 2010 issue of the *Economic Journal* (Volume 120, Issue 548) offer nice reviews. Like much of that literature, our paper correlates economic behaviors in "the wild" with performance on an experimental task. But our approach is distinct in a couple of ways. First, we present subjects with a standard consumer decision problem: the selection of a bundle of (contingent) commodities from a standard budget set. Secondly, the decision problems are presented on a user-friendly graphical interface that allows for the collection of a rich individual-level data set. The large amount of data generated by this design allows us to apply revealed preference theory to determine whether the observed choices are consistent with utility maximization. Overall, we view our methods and results as complementing those based on cognitive functioning, IQ and economic literacy tasks, and anticipate that future work will investigate the relationship between economic rationality, performance on these other tasks, and important economic behaviors.

By analyzing the sociodemographic correlates of economic rationality, our paper also contributes to the emerging literature on the relation of *laboratory* behaviors to cognitive abilities (see, for example, Benjamin et al., 2006, and Dohmen et al., 2010). We differ from these prior studies as

[33] In the Netherlands, assets held in owner-occupied housing are not subject to the usual presumptive capital income tax. If they were, the imputed rent would be 4 percent of housing value and that implicit income would be taxed at 30 percent. Instead, imputed rent is presumed to be very low (currently 0.55 percent of housing value), is subject to the progressive tax on labor income, and that tax is not even due unless the household claims a deduction for mortgage interest. Nominal mortgage interest is, in turn, fully deductable from taxable income. Thus, for purposes of federal taxation, housing assets underwritten by a mortgage will typically pay a negative rate of return. In this way, according to van Ewijk et al. (2007), the Netherlands offers by far the most favorable tax treatment of owner occupied housing in Western Europe.

we use the extent of consistency with utility-maximizing behavior as a single measure for "economic cognition" and investigate the correlation between consistency under laboratory conditions and sociodemographic and economic characteristics. Because the experimental task involves choice under risk, the paper also relates to the experimental literature that investigates whether the risk attitudes that arise in the laboratory are connected to attributes that subjects bring to the experiments from outside the lab. von Gaudecker et al. (forthcoming) also conducted risk experiments with CentERpanel members. Our findings in this paper are consistent with their conclusion that "while many people exhibit consistent choice patterns, some have very high error propensities." The relation of our experimental results to individual characteristics also sheds light on the external validity of our findings, which Levitt and List (2007) and Falk and Heckman (2009) point out is a critical concern for experimental studies.

Finally, related to our revealed-preference tests, but somewhat further afield, our paper also contributes to the literature that has implemented these tests on aggregate consumption data. Such real-world data do not, however, provide a particularly rigorous test of consistency because choice sets are such that budget lines do not cross frequently (see Blundell et al., 2003). Furthermore, even a high level of consistency in the individual-level decisions does not imply that aggregate data are consistent. Cox (1987), Sippel (1997), Mattei (2000), Harbaugh et al. (2001), and Andreoni and Miller (2002), among others, ask whether behavior in the laboratory is consistent with utility maximization. The Bronars (1987) test has been widely used, so it allows us to relate our results to this literature. Our study has the highest Bronars power of one (all random subjects had violations). We note that even random behavior can appear consistent if the sample size is small, as it often is in experimental studies.

8 Concluding remarks

Some people make better economic decisions than others. But it is usually hard to tell whether someone has made bad choices; he might have uncommon preferences, or face unobserved constraints, or hold (reasonable) beliefs that rationalize his decisions. Standard economic analysis takes a libertarian approach; in the absence of data that allow us to identify bad decisions, we assume that all choices are good. The libertarian approach has obvious appeal. We rightly hesitate to evaluate the quality of decisions when we do not have sufficient information to make a definitive judgement.

This study suggests an alternative path. We offered a new field experimental design—employing graphical representations of standard consumer decision problems and using a rich pool of subjects—that enables us to collect richer data than has been possible in the past. These data allow us to say some choices are better than others, in that some choices are more rational than others. Because the data are provided by a relatively large and heterogenous sample, we can thoroughly analyze the correlates of individual levels of rationality and relate rationality in this simple domain to important economic outcomes like wealth.

The conclusions of our investigation can be summarized under three headings:

- Many subjects reveal nearly perfect consistency with utility maximization in the individual-level decisions. Our non-trivial tests of consistency suggest that nearly half of our subjects exhibit behavior that appears to be almost optimizing in the sense that their choices come close to satisfying GARP according to a number of standard measures. At the same time, there is considerable heterogeneity in the consistency scores across subjects.

- Although individual behaviors are heterogeneous, subjects' consistency scores are correlated with sociodemographic and economic characteristics. Our study provides, to our knowledge, the first experimental evidence on the question "who is more rational?" The relationships between sociodemographics and levels of rationality may ultimately prove to be useful for informing the design of social programs (Manski, 2001) or (libertarian) paternalistic policies (Thaler and Sunstein, 2003).

- Differences in the experimental consistency scores help explain differential patterns of wealth across households. The magnitudes are large, implying that a standard deviation increase in the consistency score is associated with 15-19 percent more wealth. This result is little changed when we add controls for income, education, family structure, risk tolerance or measures of personality. This finding is especially notable given the substantial heterogeneity in our experimental outcomes, and the brief experimental exposure.

The experimental techniques that we have developed provide some promising tools for future work, and the results suggest a number of potential directions. One direction exploits the fact that the experimental platform and analytic techniques are applicable to many other types of individual choice problems. We have already run a pilot experiment to apply these techniques to the domain of intertemporal choice. A goal of this line of research is to generate analogous individual-level data with which to evaluate the quality of individual-level decision-making over time.

A principal limitation of our current data set is sample size. While large by experimental economics standards, the sample is too small for purposes of detailed disaggregation of behavior, especially with respect to sociodemographic characteristics. Another promising direction is to conduct an additional set of experiments with more panel members. Of particular interest is the relationship between consistency of choice and age. By combining experimental and survey data on the elderly, we can better understand the reasons why, in previous studies, older people appear to be leaving "money on the table." In this new set of experiments, we will collect a much wider range of individual sociodemographic, cognitive and economic information from the panel members, focusing especially on health and medical information.

References

[1] Ahn, D., S. Choi D. Gale, and S. Kariv (2010) "Estimating Ambiguity Aversion in a Portfolio Choice Experiment." Mimeo.

[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] 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.

[5] Alessie, R. and A. Kapteyn (2001) "Savings and Pensions in The Netherlands." *Research in Economics*, *55*, pp. 61-82.

[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] Andreoni, J. and W. Harbaugh (2006) "Power Indices for Revealed Preference Tests." Mimeo.

[8] Andreoni, J. and J. Miller (2002) "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica*, *70*, pp. 737-753.

[9] Banks, J. (2010) "Cognitive Function, Financial Literacy and Financial Outcomes at Older Ages: Introduction." *Economic Journal*, *120*, pp. 357-362.

[10] Banks, J., C. O'Dea and Z. Oldfield (2010) "Cognitive Function, Numeracy and Retirement Saving Trajectories." *Economic Journal*, *120*, pp. 381-410.

[11] Barrick, M. and M. Mount (1991) "The Big Five Personality Dimensions and Job Performance: A Meta-analysis." *Personnel Psychology*, *44*, pp. 1-26.

[12] Becker, G. (1962) "Irrational Behavior and Economic Theory." *Journal of Political Economy*, *70*, pp. 1-13.

[13] Bellemare, C. S. Kröger, and A. van Soest (2008) "Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities." *Econometrica*, *76*, pp. 815-839.

[14] Benjamin, D., S. Brown and J. Shapiro (2006) "Who is "Behavioral"? Cognitive Ability and Anomalous Preferences." Mimeo.

[15] 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.

[16] Bernheim, B. D., J. Skinner, and S. Weinberg (2001) "What Accounts for the Variation in Retirement Wealth among U.S. Households?" *American Economic Review, 91*, pp. 832-857.

[17] Block, J. (2010) "The Five-factor Framing of Personality and Beyond: Some Ruminations." *Psychological Inquiry*, 21, pp. 2-25.

[18] Blundell, R., M. Browning and I. Crawford (2003) "Nonparametric engel curves and revealed preference." *Econometrica*, *71*, pp. 205-240.

[19] Bronars, S. (1987) "The power of nonparametric tests of preference maximization." *Econometrica*, *55*, pp. 693-698.

[20] Camerer, C. (1995) "Individual Decision Making." In *Handbook of Experimental Economics*, ed. John Kagel and Alvin Roth. Princeton: Princeton University Press.

[21] Charness, G., E. Karni and D. Levin (2007) "Individual and Group Decision Making under risk: An Experimental Study of Bayesian Updating and Violations of First-Order Stochastic Dominance." *Journal of Risk and Uncertainty*, *35*, pp. 129-48.

[22] 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.

[23] 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.

[24] 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.

[25] Cox, J. (1997) "On Testing the Utility Hypothesis." *Economic Journal*, *107*, pp. 1054-1078.

[26] Dean, M. and D. Martin (2010) "How Rational are your Choice Data?" Mimeo.

[27] 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.

[28] 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.

[29] Falk, A. and J. Heckman (2009) "Laboratory Experiments are a Major Source of Knowledge in Social Sciences." *Science*, *326*, pp. 535-38.

[30] Famulari, M. (1995) "A Household-Based Nonparametric Test of Demand Theory." *Review of Economics and Statistics*, *2*, pp. 371-382.

[31] Fang, H., M. Keane and D. Silverman (2008) "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy*, *116*, pp. 303-350.

[32] Fisman, R., S. Kariv and D. Markovits (2007) "Individual Preferences for Giving." *American Economic Review*, *97*, pp. 1858-1876.

[33] Fisman, R., S. Kariv and D. Markovits (2010) "Exposure to Ideology and Distributional Preferences." Mimeo.

[34] Gul, F. (1991) "A Theory of Disappointment in Decision Making under Uncertainty." *Econometrica*, *59*, pp. 667–686.

[35] Hadar, J. and W. Russell (1969) "Rules for Ordering Uncertain Prospects." *American Economic Review*, *59*, pp. 25-34.

[36] Heckman, J. (1979) "Sample Selection Bias as a Specification Error." *Econometrica*, *47*, pp. 153-161.

[37] Harbaugh, W., K. Krause and T. Berry (2001) "GARP for Kids: On the Development of Rational Choice Behavior." *American Economic Review*, *91*, pp. 1539-1545.

[38] Houtman, M. and J. Maks (1985) "Determining all Maximial Data Subsets Consistent with Revealed Preference." *Kwantitatieve Methoden*, *19*, pp. 89-104.

[39] Jappelli, T. (2010) “Economic Literacy: An International Comparison.” *Economic Journal*, *120*, pp. 429-451.

[40] Levitt, S. and J. List (2007) “What do Laboratory Experiments Tell Us About the Real World?” *Journal of Economic Perspectives*, *21*, pp. 153-174.

[41] Lusardi, A. (2008) “Households Saving Behavior: The Role of Literacy, Information and Financial Education Programs.” NBER Working Paper No. 13824.

[42] 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.

[43] Manski, C. (2001) “Designing Programs for Heterogeneous Populations: The Value of Covariate Information.” *American Economic Review*, Papers & Proceedings, *91*, pp. 103-106.

[44] Mas-Colell, A. (1977) “The Recoverability of Consumer’s Preferences from Market Demand Behavior.” *Econometrica*, *45*, pp. 1409-1430.

[45] Mas-Colell, A. (1978) “On Revealed Preferences Analysis.” *Review of Economic Studies*, *45*, pp. 121-131.

[46] Mattei, A. (2000) “Full-Scale Real Tests Of Consumer Behavior Using Experimental Data.” *Journal of Economic Behavior & Organization*, *43*, pp. 487-497.

[47] OECD (2009) “Pensions at a Glance 2009: Retirement-Income Systems in OECD Countries.” URL: www.oecd.org/els/social/pensions/PAG.

[48] Papke, L. and J. Wooldridge (1996) “Econometric Methods for Fractional Response Variables with an Application to 401 (K) Plan Participation Rates.” *Journal of Applied Econometrics*, *11*, pp. 619-632.

[49] Quiggin, J. (1990) “Stochastic Dominance in Regret Theory.” *Review of Economic Studies*, *57*, pp. 503-511.

[50] Ramsey, J. (1969) “Tests for specification errors in classical linear least-squares regression analysis.” *Journal of the Royal Statistical Society*, *31*, pp. 350-371.

[51] Samuelson, P. (1947) *Foundations of Economic Analysis*. Cambridge: Harvard University Press.

[52] Savage, L. J. (1954) *The Foundations of Statistics*. New York: Wiley.

[53] Scholz, J., A. Seshadri and S. Khitatrakun (2006) "Are Americans Saving "Optimally" for Retirement?" *Journal of Political Economy*, *114*, pp. 607-643.

[54] Sippel, R. (1997) "An Experiment on the Pure Theory of Consumer's Behavior." *Economic Journal*, *107*, pp. 1431-1444.

[55] Smith, J., J. McArdle and R. Willis (2010) "Financial Decision Making and Cognition in a Family Context." *Economic Journal*, *120*, pp. 363-380.

[56] Starmer, C. (2000) "Developments in Non-Expected Utility Theory: The Hunt for a descriptive Theory of Choice under Risk." *Journal of Economic Literature*, *38*, pp. 332-382.

[57] Terza, J. (1998) "Estimating Count Data Models with Endogenous Switching: Sample Selection and Endogenous Treatment Effects." *Journal of Econometrics*, *84*, pp. 129-154.

[58] Thaler, R. and C. Sunstein (2003) "Libertarian Paternalism." *American Economic Review*, Papers & Proceedings, *93*, pp. 175-179.

[59] Varian, H. (1982) "The Nonparametric Approach to Demand Analysis." *Econometrica*, *50*, pp. 945-972.

[60] Varian, H. (1983) "Non-Parametric Tests of Consumer Behaviour." *Review of Economic Studies*, *50*, pp. 99-110.

[61] Varian, H. (1990) "Goodness-of-Fit in Optimizing Models." *Journal of Econometrics*, *46*, pp. 125-140.

[62] Varian, H. (1991) "Goodness-of-Fit for Revealed Preference Tests." Mimeo.

[63] Van Den Berg, G., D. Deeg, M. Lindeboom and F. Portrait (2010) "The Role of Early-Life Conditions in the Cognitive Decline due to Adverse Events Later in Life." *Economic Journal*, *120*, pp. 411-428.

[64] van Ewikj, C., B. Jacobs and R. de Mooij (2007) "Welfare Effects of Fiscal Subsidies On Home Ownership in the Netherlands." *De Economist*, *155*, pp. 323-336.

[65] von Gaudecker, H-M., A. van Soest and E. Wengström (forthcoming) "Heterogeneity in Risky Choice Behaviour in a Broad Population." *American Economic Review.*

[66] Wakker, P. (1993) "Savage's Axioms Usually Imply Violation of Strict Stochastic Dominance." *Review of Economic Studies*, *60*, pp. 487-493.

Table 1. Sociodemographic information

	Participants	Dropouts	Non-participants
Female	45.43	37.89	50.00
Age			
16-34	18.53	3.16	26.14
35-49	26.14	12.11	32.13
50-64	35.62	38.42	27.58
65+	19.71	46.32	14.15
Education			
Low	33.59	42.63	30.99
Medium	29.70	22.63	31.61
High	36.72	34.74	37.40
Household monthly income			
€0-2500	22.42	34.73	21.28
€2500-3499	25.13	26.32	18.90
€3500-4999	28.85	16.32	28.93
€5000+	23.60	22.63	30.89
Occupation			
Paid work	53.13	39.47	62.91
House work	11.59	7.89	8.78
Retired	20.90	42.63	13.95
Others	14.38	10.00	14.36
Household composition			
Partnered	80.88	67.89	82.64
# of children	0.84	0.32	1.09
# of obs.	1182	190	968

Participants completed the experiment, dropouts logged in and quit the experiment, and nonparticipants were recruited for the experiment but never logged in. The low, medium and high education levels correspond to primary or pre-vocational secondary education, pre-university secondary education or senior vocational training, and vocational college or university education, respectively. The classification of occupations is based on the categorization of Statistics Netherlands.

Table 2. CCEI scores

	Mean	Std. Dev.	Percentiles 10	25	50	75	90	# of obs.
All	0.881	0.141	0.676	0.808	0.930	0.998	1.000	1182
Female	0.874	0.147	0.666	0.796	0.928	0.998	1.000	537
Age								
16-34	0.920	0.119	0.734	0.881	0.979	1.000	1.000	219
35-49	0.906	0.123	0.708	0.853	0.966	1.000	1.000	309
50-64	0.863	0.142	0.666	0.784	0.901	0.985	1.000	421
65+	0.843	0.164	0.595	0.770	0.882	0.981	1.000	233
Education								
Low	0.863	0.143	0.665	0.782	0.906	0.987	1.000	397
Medium	0.881	0.140	0.689	0.814	0.926	0.998	1.000	351
High	0.899	0.137	0.686	0.842	0.963	1.000	1.000	430
Household monthly income								
€0-2500	0.856	0.154	0.617	0.769	0.911	0.983	1.000	269
€2500-3499	0.885	0.133	0.705	0.809	0.925	0.999	1.000	302
€3500-4999	0.882	0.141	0.649	0.817	0.932	0.999	1.000	345
€5000+	0.901	0.131	0.729	0.836	0.968	1.000	1.000	266
Occupation								
Paid work	0.896	0.131	0.705	0.833	0.950	1.000	1.000	628
House work	0.873	0.151	0.649	0.795	0.937	0.999	1.000	137
Retired	0.839	0.158	0.597	0.767	0.876	0.971	1.000	247
Others	0.891	0.129	0.712	0.809	0.936	0.998	1.000	170
Household composition								
Partnered	0.878	0.142	0.673	0.802	0.927	0.998	1.000	956
Children	0.899	0.128	0.704	0.835	0.959	1.000	1.000	490

Afriat's (1972) Critical Cost Efficiency Index (CCEI) is bounded between zero and one. The closer it is to one, the closer the data are to satisfying GARP; and the difference between the CCEI and one can be interpreted as an upper bound of the fraction of income that an individual is "wasting" by making inconsistent choices.

Table 3. CCEI scores for the combined data set

	Mean	Std. Dev.	Percentiles 10	25	50	75	90	Δ Mean
All	0.733	0.229	0.394	0.584	0.775	0.943	0.985	0.148
Female	0.733	0.224	0.409	0.588	0.767	0.941	0.984	0.141
Age								
16-34	0.786	0.228	0.442	0.637	0.881	0.976	0.995	0.134
35-49	0.782	0.206	0.481	0.652	0.845	0.962	0.991	0.124
50-64	0.700	0.225	0.371	0.552	0.735	0.898	0.973	0.163
65+	0.679	0.242	0.334	0.489	0.703	0.902	0.968	0.165
Education								
Low	0.699	0.226	0.374	0.535	0.732	0.902	0.967	0.163
Medium	0.733	0.226	0.394	0.595	0.768	0.941	0.986	0.148
High	0.767	0.227	0.428	0.625	0.849	0.968	0.992	0.131
Household monthly income								
€0-2500	0.706	0.218	0.382	0.535	0.737	0.902	0.977	0.150
€2500-3499	0.741	0.220	0.439	0.612	0.768	0.946	0.986	0.143
€3500-4999	0.730	0.236	0.388	0.556	0.782	0.950	0.984	0.151
€5000+	0.755	0.238	0.383	0.627	0.841	0.952	0.992	0.146
Occupation								
Paid work	0.758	0.222	0.428	0.615	0.817	0.955	0.991	0.139
House work	0.719	0.233	0.380	0.548	0.765	0.928	0.986	0.154
Retired	0.675	0.231	0.334	0.502	0.698	0.872	0.964	0.164
Others	0.738	0.231	0.406	0.599	0.793	0.951	0.983	0.153
Household composition								
Partnered	0.729	0.229	0.389	0.583	0.771	0.938	0.984	0.149
Children	0.760	0.216	0.443	0.614	0.815	0.952	0.987	0.139

The combined data set is the actual data from the experiment and the mirror-image data. By definition, the CCEI score for the combined data set consisting of 50 observations can be no bigger than the CCEI score for the actual data reported in Table 2.

Table 4. The correlation between CCEI scores and subjects' individual characteristics (OLS)

	(1)	(2)
Constant	.887***	.735***
	(.022)	(.037)
Female	-.024***	-.011
	(.009)	(.015)
Age		
35-49	-.016	-.007
	(.011)	(.020)
50-64	-.052***	-.077***
	(.011)	(.020)
65+	-.051**	-.081**
	(.020)	(.032)
Education		
Medium	.009	.021
	(.011)	(.017)
High	.026**	.060***
	(.011)	(.018)
Income		
€2500-3499	.026**	.026
	(.012)	(.019)
€3500-4999	.020	.006
	(.013)	(.020)
€5000+	.033**	.017
	(.014)	(.022)
Occupation		
Paid work	.028	.030
	(.018)	(.026)
House work	.047**	.039
	(.021)	(.030)
Others	.037*	.035
	(.019)	(.030)
Household composition		
Partnered	-.026**	-.023
	(.011)	(.018)
# of children	.001	.001
	(.004)	(.007)
R^2	.068	.058
# of obs.	1182	1182

Dependent variables: (1) CCEI; (2) CCEI for the combined data set. Omitted categories: male, age under 35, low education (primary or pre-vocational secondary education), household gross monthly income under €2500, retired, and not having a partner. Standard errors in parentheses. *, **, and *** indicate 10, 5, and 1 percent significance levels, respectively.

Table 5. The correlation between CCEI scores and subjects' individual characteristics
(sample-selection)

	(1)		(2)	
	Outcome	Selection	Outcome	Selection
Constant	.888***	.544*	.891***	-2.077***
	(.022)	(.311)	(.023)	(.209)
Female	-.024***	.084	-.024***	-.031
	(.009)	(.103)	(.009)	(.068)
Age				
35-49	-.016	-.556**	-.016	-.133
	(.011)	(.230)	(.011)	(.102)
50-64	-.051***	-1.024***	-.052***	-.393***
	(.011)	(.220)	(.011)	(.102)
65+	-.050**	-1.556***	-.051**	-.824***
	(.021)	(.263)	(.020)	(.154)
Education				
Medium	.009	.191	.009	-.036
	(.011)	(.122)	(.011)	(.081)
High	.026**	.168	.026**	.006
	(.011)	(.117)	(.011)	(.084)
Income				
€2500-3499	.025**	.303**	.025**	.281***
	(.012)	(.125)	(.012)	(.094)
€3500-4999	.019	.426***	.019	.186**
	(.013)	(.141)	(.014)	(.094)
€5000+	.033**	.064	.033**	.080
	(.014)	(.147)	(.014)	(.106)
Occupation				
Paid work	.028	-.202	.029	-.040
	(.018)	(.172)	(.018)	(.131)
House work	.046**	.108	.046**	.083
	(.020)	(.200)	(.020)	(.148)
Others	.037**	.081	.037*	.110
	(.019)	(.196)	(.019)	(.147)
Household composition				
Partnered	-.026**	.262**	-.027**	.123
	(.011)	(.119)	(.011)	(.092)
# of children	.001	.145**	.001	.031
	(.004)	(.068)	(.004)	(.036)
Participation ratio		1.231***		3.387***
		(.205)		(.125)
ρ	-.028		-.047	
	(.083)		(.063)	
Log peudolikelihood	210.856		-371.973	
# of obs.	1372		2340	

Table 5
(continued)

	(3)		(4)	
	Outcome	Selection	Outcome	Selection
Constant	.759***	.545*	.757***	-2.067***
	(.043)	(.314)	(.038)	(.208)
Female	-.013	.084	-.011	-.032
	(.015)	(.104)	(.015)	(.068)
Age				
35-49	-.001	-.554**	-.009	-.135
	(.022)	(.223)	(.020)	(.101)
50-64	-.062**	-1.023***	-.079***	-.397***
	(.024)	(.212)	(.020)	(.102)
65+	-.049	-1.557***	-.078**	-.822***
	(.042)	(.258)	(.032)	(.154)
Education				
Medium	.016	.191	.021	-.036
	(.018)	(.120)	(.017)	(.081)
High	.054***	.169	.059***	.007
	(.018)	(.117)	(.018)	(.084)
Income				
€2500-3499	.017	.304**	.022	.276***
	(.021)	(.127)	(.019)	(.093)
€3500-4999	-.006	.428***	.003	.174*
	(.022)	(.138)	(.020)	(.094)
€5000+	.015	.065	.018	.075
	(.022)	(.145)	(.022)	(.106)
Occupation				
Paid work	.034	-.203	.031	-.035
	(.027)	(.173)	(.026)	(.131)
House work	.036	.109	.038	.075
	(.030)	(.205)	(.030)	(.148)
Others	.032	.081	.034	.110
	(.030)	(.193)	(.030)	(.146)
Household composition				
Partnered	-.032	.261**	-.026	.126
	(.020)	(.115)	(.018)	(.091)
# of children	-.000	.145**	.002	.028
	(.007)	(.062)	(.007)	(.036)
Participation ratio		1.230***		3.378***
		(.234)		(.125)
ρ	-.396		-.155	
			(.075)	
Log peudolikelihood			-949.787	
# of obs.	1372		2340	

Dependent variables: (1) and (2) CCEI; (3) and (4) CCEI for the combined data set. Omitted categories: male, age under 35, low education (primary or pre-vocational secondary education), household gross monthly income under €2500, retired, and not having a partner. Standard errors in parentheses. *, **, and *** indicate 10, 5, and 1 percent significance levels, respectively.

Table 6. Household wealth
(average 2008-2009 in 2008 Euros)

Mean		164,130
Std. Dev.		243,548
Max		3,984,151
Min		-180,700
Percentiles	1	-68,237
	5	-4,810
	10	0
	25	10,780
	50	92,979
	75	242,054
	90	412,494
	95	523,839
	99	955,599
# of obs.		703

Table 7. Wealth components
(average 2008-2009 in 2008 Euros by age-adjusted CCEI)

		Age-adjusted CCEI		
	All	Low	High	Δ
Household wealth	156,413	154,280	158,540	4,260
	(7,106)	(10,223)	(9,888)	(14,222)
	693	346	347	
Log household wealth	11.16	11.08	11.23	0.16
	(0.07)	(0.11)	(0.10)	(0.15)
	622	312	310	
Checking accounts	2,623	2,787	2,458	-329
	(211)	(370)	(204)	(423)
	693	346	347	
Savings accounts	21,753	20,572	22,931	2,359
	(1,387)	(1,591)	(2,271)	(2,774)
	693	346	347	
Stocks	10,766	9,151	12,376	3,226
	(1,472)	(1,569)	(2,488)	(2,943)
	693	346	347	
House value	183,538	180,528	186,539	6,011
	(6,358)	(9,430)	(8,545)	(12,724)
	693	346	347	
Debts	3,956	3,650	4,260	611
	(503)	(710)	(712)	(1,006)
	693	346	347	
Mortgages	61,581	55,083	68,061	12,978
	(3,396)	(4,776)	(4,810)	(7,365)
	693	346	347	

The age-adjusted CCEI is the residual from a regression of CCEI on a constant, age, age² and age³. Standard errors in parentheses. The last row in each category reports the number of observations.

Table 8. The relationship between households' net worth and CCEI scores

	(1)	(2)	(3)	(4)
CCEI	1.170**	1.425**	99933.2*	1.490***
	(0.535)	(0.565)	(52656.0)	(0.574)
CCEI (combined dataset)				
Risk tolerance				
Conscientiousness				
Log 2008 household income	0.623***	0.601***		0.629***
	(0.123)	(0.127)		(0.124)
2008 household income			1.74***	
			(0.3)	
Female	-0.275*	-0.228	-28223.9*	-0.258
	(0.154)	(0.164)	(15906.3)	(0.162)
Age	0.004	-0.286	-33974.7	-0.277
	(0.205)	(0.316)	(27100.3)	(0.318)
Age^2	0.002	0.006	726.5	0.006
	(0.004)	(0.005)	(471.1)	(0.005)
Age^3	0.000	0.000	-4.3	0.000
	(0.000)	(0.000)	(2.7)	(0.000)
Partnered	0.623***	0.682***	48106.5***	0.683***
	(0.173)	(0.183)	(16995.7)	(0.184)
# of children	0.125	0.103	14472.9*	0.106
	(0.086)	(0.092)	(8291.6)	(0.093)
Education				
Pre-vocational	0.242	0.267	13056.4	
	(0.459)	(0.459)	(43981.0)	
Pre-university	0.528	0.600	57288.3	
	(0.474)	(0.476)	(45189.8)	
Senior vocational training	0.407	0.403	27365.6	
	(0.465)	(0.469)	(42967.4)	
Vocational college	0.485	0.448	27964.9	
	(0.450)	(0.452)	(42704.3)	
University	0.637	0.679	73733.5	
	(0.463)	(0.470)	(48008.2)	
Constant	0.229	5.932	335793.4	5.451
	(3.554)	(5.862)	(5.0E+05)	(6.110)
R^2	0.217	0.179	0.191	0.170
# of obs.	566	517	568	517

Table 8.
(Continued)

	(5)	(6)	(7)	(8)	(9)
CCEI	1.348*	1.545***	1.563**	1.781**	1.728**
	(0.714)	(0.591)	(0.735)	(0.746)	(0.750)
CCEI (combined dataset)	0.078		-0.018	-0.091	-0.038
	(0.381)		(0.373)	(0.381)	(0.384)
Risk tolerance		-1.166	-1.165	-1.361	-1.366
		(0.828)	(0.829)	(0.838)	(0.840)
Conscientiousness					0.103
					(0.072)
Log 2008 household income	0.602***	0.595***	0.595***	0.520***	0.514***
	(0.127)	(0.128)	(0.129)	(0.121)	(0.121)
2008 household income					
Female	-0.229	-0.232	-0.232	-0.299	-0.321
	(0.164)	(0.166)	(0.166)	(0.168)	(0.169)
Age	-0.284	-0.307	-0.308	-0.310	-0.282
	(0.316)	(0.313)	(0.315)	(0.319)	(0.316)
Age^2	0.006	0.007	0.007	0.007	0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Age^3	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Partnered	0.682***	0.726***	0.725***	0.733***	0.714***
	(0.183)	(0.187)	(0.188)	(0.191)	(0.191)
# of children	0.103	0.092	0.092	0.095	0.090
	(0.093)	(0.094)	(0.095)	(0.095)	(0.095)
Education					
Pre-vocational	0.264	0.331	0.331	0.312	0.351
	(0.461)	(0.483)	(0.484)	(0.481)	(0.481)
Pre-university	0.596	0.676	0.677	0.675	0.717
	(0.486)	(0.498)	(0.498)	(0.499)	(0.496)
Senior vocational training	0.403	0.480	0.481	0.464	0.501
	(0.469)	(0.493)	(0.494)	(0.490)	(0.490)
Vocational college	0.443	0.549	0.550	0.568	0.598
	(0.452)	(0.475)	(0.480)	(0.476)	(0.475)
University	0.672	0.745	0.746	0.774	0.806
	(0.474)	(0.498)	(0.502)	(0.501)	(0.499)
Constant	5.888	6.938	6.947	7.797	7.371
	(5.879)	(5.786)	(5.812)	(5.880)	(5.841)
R^2	0.178	0.186	0.184	0.180	0.182
# of obs.	517	507	507	494	494

Dependent variable: The natural log of household wealth, except in (3) where the dependent variable is the *level* of household wealth. Sample includes households with financial respondent age 16 and older in (1) and households with financial respondent age 35 and older in (2)-(9). Otherwise, samples vary according to availability of data on regressand and regressors. The CCEI scores for the combined dataset is computed after combining the actual data from the experiment and the mirror-image data. Risk tolerance measured by the average fraction of tokens allocated to the cheaper asset. The groupings of different levels of education are based on the categorization of Statistics Netherlands. Standard errors in parentheses. *, **, and *** indicate 10, 5, and 1 percent significance levels, respectively.

Table 9. The sources of the relationship between households' net worth and CCEI scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Log household wealth		Have checking	Fraction in checking	Have saving	Fraction in saving
CCEI	1.907**	1.792**	0.029	-0.100*	-0.052	-0.179*
	(0.751)	(0.705)	(0.032)	(0.057)	(0.051)	(0.096)
Log household income						
2008	0.686***	0.250	0.001	-0.030**	0.004	-0.062***
	(0.133)	(0.184)	(0.002)	(0.013)	(0.009)	(0.021)
2006		0.467*				
		(0.262)				
2004		0.312*				
		(0.176)				
Female	-0.144	-0.037	0.007	0.019	0.008	0.017
	(0.188)	(0.184)	(0.005)	(0.018)	(0.018)	(0.029)
Age	-0.315	-0.241	-0.005	0.029	-0.007	0.001
	(0.318)	(0.319)	(0.010)	(0.044)	(0.042)	(0.061)
Age^2	0.007	0.006	0.000	-0.001	0.000	0.000
	(0.005)	(0.005)	(0.000)	(0.001)	(0.001)	(0.001)
Age^3	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Partnered	0.790***	0.740***	-0.005	-0.033	0.016	-0.058*
	(0.222)	(0.222)	(0.004)	(0.021)	(0.022)	(0.033)
# of children	0.111	0.123	0.000	-0.004	-0.025*	-0.043***
	(0.107)	(0.101)	(0.001)	(0.009)	(0.014)	(0.013)
Education						
Pre-vocational	0.622	0.658	-0.007	0.002	0.036	0.003
	(0.613)	(0.626)	(0.007)	(0.063)	(0.068)	(0.085)
Pre-university	0.883	0.889	-0.017	-0.021	-0.004	-0.073
	(0.626)	(0.641)	(0.018)	(0.063)	(0.075)	(0.087)
Senior vocational training	0.899	0.956	0.005	0.000	0.043	-0.046
	(0.616)	(0.632)	(0.004)	(0.063)	(0.069)	(0.086)
Vocational college	1.011*	0.954	0.002	-0.005	0.014	-0.037
	(0.601)	(0.617)	(0.002)	(0.062)	(0.069)	(0.083)
University	1.182*	1.027	0.003	0.011	0.019	-0.076
	(0.610)	(0.625)	(0.003)	(0.064)	(0.071)	(0.086)
Constant	4.282	-0.415	1.036***	0.045	1.117	1.420
	(5.947)	(6.032)	(0.183)	(0.750)	(0.779)	(1.197)
R^2	0.236	0.269	0.024	0.051	0.020	0.104
# of obs.	377	377	512	512	502	502

Table 9.
(Continued)

	(7) Have stocks	(8) Fraction in stocks	(9) Have a house	(10) Fraction in house
CCEI	0.160	-0.003	0.371**	0.336***
	(0.160)	(0.050)	(0.150)	(0.129)
Log household income				
2008	0.154***	0.011	0.138***	0.098***
	(0.032)	(0.010)	(0.027)	(0.023)
2006				
2004				
Female	-0.005	0.005	-0.021	-0.039
	(0.046)	(0.012)	(0.044)	(0.039)
Age	0.089	0.011	-0.001	-0.030
	(0.088)	(0.021)	(0.081)	(0.065)
Age²	-0.001	0.000	0.000	0.001
	(0.002)	(0.000)	(0.001)	(0.001)
Age³	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Partnered	0.003	-0.008	0.211***	0.136***
	(0.049)	(0.014)	(0.051)	(0.044)
# of children	0.003	0.000	0.051**	0.069***
	(0.026)	(0.007)	(0.020)	(0.019)
Education				
Pre-vocational	-0.074	-0.042	0.030	0.067
	(0.128)	(0.049)	(0.123)	(0.109)
Pre-university	0.036	-0.011	0.090	0.071
	(0.138)	(0.052)	(0.129)	(0.114)
Senior vocational training	-0.050	-0.029	0.064	0.064
	(0.131)	(0.048)	(0.124)	(0.111)
Vocational college	0.024	-0.023	0.078	0.044
	(0.128)	(0.049)	(0.122)	(0.108)
University	0.194	0.011	0.089	0.051
	(0.136)	(0.051)	(0.127)	(0.113)
Constant	-3.326**	-0.287	-1.508	-0.777
	(1.678)	(0.388)	(1.592)	(1.289)
R^2	0.106	0.029	0.174	0.146
# of obs.	514	514	479	479

To reduce the importance of extreme outliers, in columns (3)-(10) we drop households whose fraction of wealth in the relevant category (checking, saving, stocks, housing) is less than -0.15 or greater than 1.15. The groupings of different levels of education are based on the categorization of Statistics Netherlands. Standard errors in parentheses. *, **, and *** indicate 10, 5, and 1 percent significance levels, respectively.

Figure 1. Mean CCEI scores

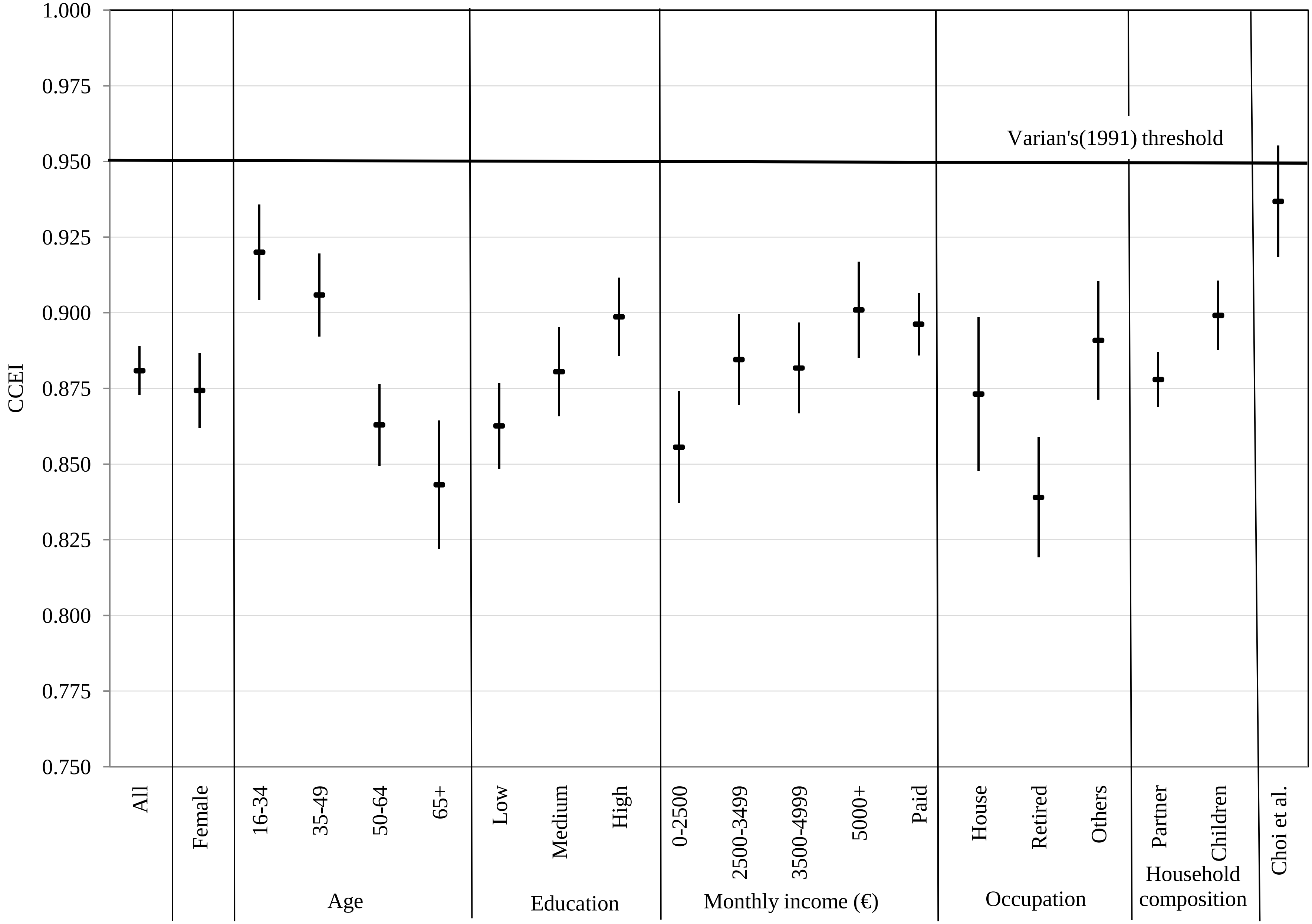
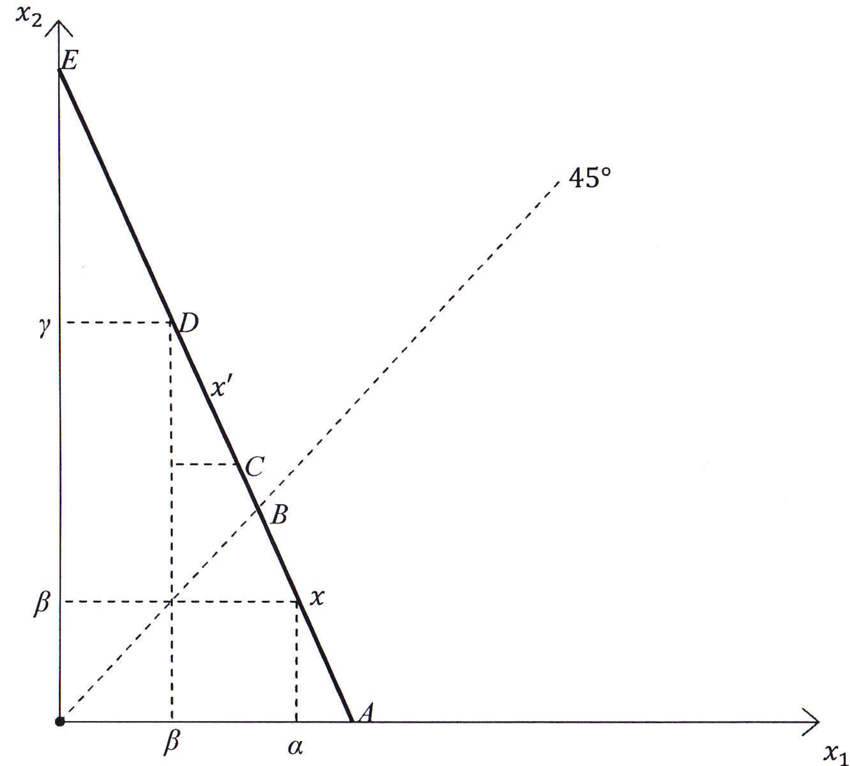


Figure 2. A violation of stochastic dominance



The individual can choose any allocation x' (position along CD) but prefers allocation x (position along AB) such that $F_{x'} < F_x$ where $F_{x'}$ and F_x are the resulting payoff distributions.

Figure 3. The average fraction of tokens allocated to the cheaper asset