Who is (More) Rational?*

Syngjoo Choi[†] UCL

Shachar Kariv[‡] UC Berkeley

Wieland Müller[§] Tilburg University

Dan Silverman[¶] University of Michigan

September 7, 2010

Abstract

We report on the results of a large-scale field experiment that enables us to test for consistency with utility-maximizing behavior and to investigate the correlation between individual behavior and demographic and economic characteristics. We conducted the experiment with the CentERpanel (a representative sample of over 2,000 Dutch

*We thank Douglas Gale and Raymond Fisman for detailed comments and suggestions. We are also grateful to James Andreoni, David Card, Gary Charness, Stefano DellaVigna, Steffen Huck, Edi Karni, Patrick Kline, Dan Levin, John List, Imran Rasul, and William Zame for helpful discussions. 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 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) (Müller).

†Department of Economics, University College London, Gower Street, London WC1E 6BT, UK (Email: syngjoo.choi@ucl.ac.uk, URL: http://www.homepages.ucl.ac.uk/~uctpsc0).

‡Department of Economics, University of California, Berkeley, 508-1 Evans Hall # 3880, Berkeley, CA 94720, USA (E-mail: kariv@berkeley.edu, URL: http://emlab.berkeley.edu/~kariv/).

§Department of Economics, Tilburg University, Postbus 90153, 5000LE Tilburg, The Netherlands (E-mail: w.mueller@uvt.nl, URL: http://center.uvt.nl/staff/muller/).

¶Department of Economics, University of Michigan, 611 Tappan St., Ann Arbor, MI 48109, USA (E-mail: dansilv@umich.edu, URL: http://www-personal.umich.edu/~dansilv/).

households), using procedures similar to those used by Choi, et al. (2007b) in a setting with risk. We find a high level of consistency in the individual-level decisions. There is also considerable heterogeneity in subjects' consistency scores. High-income and high-education subjects display greater levels of consistency. Men are more consistent than women, and young subjects are more consistent than those who are old. Additionally, young and high-education subjects also display a lower rate of violations of monotonicity with respect to first-order stochastic dominance.

JEL Classification Numbers: C93, D01, D03, D12, D81

Keywords: Rationality, revealed preferences, CentERpanel, experiment.

1 Introduction

Rational choice requires a consistent (complete and transitive) preference ordering over all possible alternatives, and choices that correspond to the most preferred alternative from the feasible set. Rationality is the most important input into any measure of welfare and enters virtually every realm of individual decision-making. It also has an important impact on economic and social policy, because policy relies on an understanding of what individuals would wish for themselves. Insofar as preferences are rational, then the techniques of economic analysis may be brought to bear on modeling and predicting behavior governed by these preferences and to make positive predictions and welfare conclusions. This highlights the importance of developing a rigorous test of rational choice, and insofar as rationality differ across individuals, of explaining individual heterogeneity in terms of demographic variables.

Within economics 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 gain from the experimental study of rational choice under laboratory conditions to practical questions in the broader world, we conducted a *field* experiment utilizing 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 socio-demographic and economic information for the panel members. This is central to understanding the relationship between demographic background and personal attributes and preferences in many policy-relevant domains. Further, the data set produced by our experimental design has several advantages over data sets from earlier experiments. Most importantly, it generates a very rich data set well-suited to studying behavior at the level of the individual subject. Combining our experimental setup's capacity with the CentERpanel generates a unique opportunity to investigate the relationship between the extent to which individual behaviors comply with rationality and demographic differences in the general population.

In our experiment, we present subjects with a sequence of 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 individual wealth between *risky* assets. These decision problems are presented using a graphical interface introduced by Choi, et al. (2007a) and exploited by Choi, et al. (2007b).[1] Because the design is very user-friendly, it is possible to present each subject with *many* choices in the course of a short experimental session, yielding much larger data set than has heretofore been possible. 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 are able to use revealed preference tests to investigate the extent to which the data comply with utility maximization. Since we observe many choices over a wide range of budget lines, our data lead to high power tests of revealed preference conditions.

Classical revealed preference theory tells us that choices in a finite collection of budget sets are consistent with maximizing a well-behaved (that is, piecewise linear, continuous, increasing, and concave) utility function if and only if they satisfy the Generalized Axiom of Revealed Preference (GARP). Since GARP offers an exact test (either the data satisfy GARP or they do not), we calculate a variety of goodness-of-fit indices to quantify the extent of violation. The main index is Afriat's (1972) Critical Cost Efficiency Index (CCEI). It measures the fraction by which each budget constraint must be relaxed in order to remove *all* violations of GARP. By definition, the CCEI is bounded between zero and one. The closer it is to one, the closer the data are to satisfying GARP. Hence, the CCEI provides a summary statistic of the overall consistency of the data. The difference between the CCEI and one can be interpreted as an *upper bound* of the fraction of wealth that an individual is 'wasting' by making inconsistent choices.

If we follow Varian's (1991) suggestion of a threshold of 0.95 for the CCEI, we find that 45.2 percent our subjects' scores are above this thresh-

[1] In a concurrent paper, we provide a corroboration of the earlier work of Choi, et al. (2007b) on risk preferences and explain heterogeneity in risk preferences in terms of demographic variables.

old and of those 22.8 percent have no violations of GARP. 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 many subjects are close enough to passing 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 only 0.881, which implies that subjects on average waste *as much as* nearly twelve percent of their earnings by making inefficient choices. There is also marked heterogeneity in the CCEI scores within and across the demographic characteristics of our subjects. Figure 1 summarizes the mean CCEI scores and 95 percent confidence intervals across standard socioeconomic categories.

[Figure 1 here]

We therefore move to studying the correlations between goodness-of-fit indices and demographic and economic characteristics. Our data are particularly well-suited to such investigations, given the heterogeneity in our experimental outcomes, and the heterogeneity in our subject pool. Using Heckman's (1979) sample-selection model, we rule out sample selection bias to participate in the experiment. 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. Additionally, men are more consistent than women, and young subjects lean more towards 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, the most natural question to ask about our data is whether choices are 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 risk attitudes, if it yields unambiguously higher monetary payoff. The dominance principle is compelling and generally accepted in decision theory.[2] Clearly, violations of the axiom of

[2] As noted by Quiggin (1990) and Wakker (1993), theories of choice under uncertainty

monotonicity with respect to first-order stochastic dominance may be regarded as errors. The choices made by subjects in our experiment also show low rates of stochastic dominance violations, which decrease with education level and increase with age.

The rest of the paper is organized as follows. Section 2 describes the experimental design and procedures. Section 3 evaluates the consistency of the data. Section 4 provides the econometric analysis, and Section 5 contains some concluding remarks and describes the margins along which we extend the previous literature.

2 Experimental Design

The experiment utilizes the CentERpanel, an on-line, weekly, stratified survey of a sample of over 2,000 households and 5,000 individual members. The panel is an appropriate representation of the Dutch-speaking population in the Netherlands. The CentERpanel thus provides a unique opportunity to combine experimental data with demographic and economic variables from the survey. The subjects in the experiment were randomly recruited from the entire CentERpanel body. The experiment was conducted *on-line* under the CentERdata protocol with 1,182 CentERpanel adult members, using the experimental technique introduced by Choi, et al. (2007a) and exploited by Choi, et al. (2007b).[3] The experimental methodology allows us to identify individual behaviors that may be related to a wide range of individual characteristics.

Table 1 below provides summary statistics on the individual level characteristics. We present the data for *participants* (completed the experiment), *dropouts* (logged in and quitted the experiment) and *non-participants* (recruited for the experiment but did not log in). We use five standard socioeconomic categories: age, education, income, occupation, and household composition. The groupings of different levels of education are based on the categorization of Statistics Netherlands (Centraal Bureau voor de Statistiek). The low, medium and high education levels correspond to primary education or lower secondary education, secondary education, and higher education, respectively. We use household monthly gross income-level cate-

were amended to avoid violations of dominance.

[3]CentERdata is a research institute in Tilburg School of Economics and Management (TiSEM) at Tilburg University in the Netherlands. The CentERdata specializes in online experiments and manages the CentERpanel and several other panels. The panel members complete the questionnaires on the internet from home. For more information, see http://www.centerdata.nl/en/centerpanel.

gories such that the proportions of participants in each category are approximately equal. The classification of occupations used by Statistics Netherlands is based on type of work.

[Table 1 here]

The experimental interface was incorporated into the CentERpanel and the experiment was hosted as part of their survey. In our experiment, we presented subjects with a sequence of decision problems under uncertainty. 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, y (corresponding to the usual 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 by 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.[4] The procedures described below are *identical* to those used by Choi, et al. (2007b), with the exception that the current experiment consisted of 25 instead of 50 decision problems and some minor additional changes resulting from the on-line experimental setting.[5] Payoffs were calculated in terms of points and then converted into euros. Each point was worth €0.25. Subjects received their payment via the CentERpanel reimbursement system.

Each decision problem started by having the computer select 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. The budget lines selected for each subject in different decision problems were independent of each other and of the sets selected for

[4] 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 experimental methodology to study distributional preferences and produce very different behaviors. It is of course possible that presenting choice problems graphically biases choice behavior in some particular way – and that is a useful topic for experiment – but there is no evidence that this is the case, as emphasized by Choi, et al. (2007b) and Fisman, et al. (2007).

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

any of the other subjects in their decision problems. Choices were restricted to allocations on the budget constraint. During the course of the experiment, subjects were not provided with any information about the account that had been selected in each round. As in Choi, et al. (2007b), 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. We refer the interested reader to Choi, et al. (2007a) for an extended description of the experimental interface.[6]

3 Testing Rationality

3.1 Consistency

The broad range of budget lines faced by each subject provides a stringent test of GARP.[7] 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 – but individual choices may involve errors. Subjects may compute incorrectly, or execute intended choices incorrectly, or err in other less obvious ways. To account for the possibility of errors, 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.[8]

Put precisely, let $\{(p^i, x^i)\}_{i=1}^{25}$ be the data generated by some individual's choices, where p^i denotes the i-th observation of the price vector and x^i

[6] Full experimental instructions, including the computer program dialog windows are also available at Online Appendix I (http://emlab.berkeley.edu/~kariv/CKM_I_A1.pdf).

[7] GARP (which is a generalization of various other revealed preference tests) is tied to utility representation through Afriat 's (1967) theorem. Varian (1982, 1983) modifies Afriat's (1967) results and describes efficient and general techniques for testing the extent to which choices satisfy GARP. Choi, et al. (2007a) provide more details on testing for consistency with GARP.

[8] Choi, et al. (2007b) (Appendix III) discuss alternative indices that have been proposed for this purpose by Varian (1990, 1991) and Houtman and Maks (1985). In reporting our results, we focus on the CCEI which offers straightforward interpretation. 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.)

denotes the associated allocation.[9] 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.

Table 2 below provides a population-level summary of the individual-level CCEI scores by reporting summary statistics and percentile values. 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 of the relationship between the differences in consistency scores and demographic differences among experimental subjects is the purpose of our econometric estimation below.[10]

[Table 2 here]

3.2 Power and goodness-of-fit

Revealed preference tests do have a major 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, but this is purely subjective. If we

[9] Without essential loss of generality, assume the individual's wealth is normalized to 1. The budget set is then $p_1 x_1 + p_2 x_2 = 1$ and the individual can choose any allocation x that satisfies this constraint. The data generated by an individual's choices are $\left\{\left(\bar{x}_1^i, \bar{x}_2^i, x_1^i, x_2^i\right)\right\}_{i=1}^{25}$, where $\left(x_1^i, x_2^i\right)$ are the coordinates of the choice made by the subject and $\left(\bar{x}_1^i, \bar{x}_2^i\right)$ 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.

[10] 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 token (that is, for any i and $j \neq i$, if $\left|x^i - x^j\right| \leq 1$ then x^i and x^j are treated as the same portfolio).

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. By comparison, Choi, et al. (2007b) report that 60 of their 93 subjects (64.5 percent) had CCEI scores above the 0.95 threshold, and of those 16 subjects (17.2 percent) did not violate GARP.[11]

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 randomizes uniformly among all allocations on each budget line as a point of comparison.[12] 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.

To provide a more informative metric of the consistency of choices, we follow Choi, et al. (2007a) who extend and generalize Bronars' (1987) test by employing a random sample of simulated subjects who maximize a utility function $u(\cdot)$ with error where the likelihood of error is assumed to be a decreasing function of its cost. In particular, we assume an idiosyncratic preference shock that has a logistic distribution

$$\Pr(x^*) = \frac{e^{\gamma \cdot u(x^*)}}{\int\limits_{x:p\cdot x=1} e^{\gamma \cdot u(x)}},$$

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

The histograms in Figure 2 below summarize the distributions of CCEI scores generated by samples of 25,000 simulated subjects who implement logarithmic von Neumann-Morgenstern utility function $\log x_1 + \log x_2$ with various levels of precision γ. The horizontal axis measures the fractions for different intervals of CCEI scores and the vertical axis measures the percentage of subjects corresponding to each interval. Each of the simulated subjects makes 25 choices from randomly generated budget lines in the same

[11] The subjects of Choi, et al. (2007b) were recruited from undergraduate classes and staff at UC Berkeley. They were given a larger menu of 50 budget lines which provides a more stringent revealed preference test (more below).

[12] The power of Bronars' (1987) test is defined to be the probability that a random subject violates GARP. Our setup has the highest Bronars value of one.

way as the human subjects do. The number above each bar of the histogram represents the percentage of actual subjects corresponding to each interval. The histograms in Figure 2 suggest that our experiment is sufficiently powerful to detect whether utility maximization is in fact the correct model.[13]

[Figure 2 here]

3.3 Dominance

Violations of monotonicity with respect to first-order stochastic dominance are regarded as errors, regardless of risk attitudes – failure to recognize that some allocations yield payoff distributions with unambiguously lower returns. A simple violation of dominance is illustrated in Figure 3 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 lottery yielding $F_{x'} - F_x$, which has only non-positive payoffs (each account had an equal probability of being chosen).[14] Notice that any decision to allocate *less* tokens to the *cheaper* account (position along AB) violates dominance but need not involve a violation of GARP, whereas any decision to allocate *more* tokens to the *cheaper* account (position along BE) never violates dominance.

[Figure 3 here]

We use expected payoff calculations to assess how nearly individual choice behavior complies with dominance. Suppose that we observe an individual choosing allocation x at prices p where $F_{x'} \leq F_x$ for some x' such that $p \cdot x' = 1$. The *degree* to which allocation x violates dominance can be measured by its expected return as a fraction of the *maximal* expected return that could be achieved by choosing an allocation x'. The construction of this violation index is illustrated in Figure 3 above. The point D corresponds

[13] 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).

[14] More precisely, we can identify an allocation with the resulting probability distribution over payoffs if preferences satisfy the *reduction principle*; that is, $(x_1, x_2) \sim (x_2, x_1)$ because they generate the same payoff distribution.

to the allocation x' with the highest expected return, yielding the largest upward probabilistic shift (referring to Figure 3, the outcome "α tokens" is shifted up to "γ tokens" and the outcome "β tokens" in unchanged). This suggests the following approach. For each observation (p^i, x^i), if no feasible allocation dominates the chosen allocation, then it has the highest value of one. Otherwise, it has a value less than one; specifically $(\alpha + \beta)/(\gamma + \beta)$ as illustrated in Figure 3. Since a single number is desired, for each subject, we average this violation index across all decision problems. We refer to this number as the FOSD score. Table 3 below reports summary statistics and percentile values. We again report the statistics for all subjects, as well as the statistics by socioeconomic categories.

[Table 3 here]

Over all subjects, the FOSD scores averaged 0.959. Out of the 1,182 subjects, 1,057 subjects, (89.4 percent) have FOSD scores above 0.90, and of those 839 subjects (70.1 percent) have scores above 0.95. The mean FOSD score for a random sample of 25,000 simulated subjects is 0.920, but only 73.5 percent and 18.6 percent of the random subjects' FOSD scores were above the 0.90 and 0.95 thresholds, respectively. There is also some heterogeneity in the FOSD scores within and across categories, which is another focus of our econometric estimation below. Furthermore, when we report our estimates of the CCEI correlates, we will also distinguish between the estimates for subjects with FOSD scores above 0.90 and 0.95. Finally, Figure 4 below presents a scatterplot of the CCEI and FOSD scores. To facilitate presentation of the data, we omit 25 subjects (2.1 percent) with CCEI scores below 0.5. Note that there is considerable heterogeneity in both indices, and that their values are positively correlated ($r^2 = 0.446$). Nevertheless, the data are clearly concentrated in the upper right corner, so the majority of subjects exhibit high CCEI and FOSD scores. These facts strongly suggest that most subjects in our experiment had no difficulties understanding the decision problem or using the computer program. In fact, only 29 (2.5 percent) have no violations of GARP and FOSD scores below the 0.90 threshold. All of these are subjects almost always chose the same endpoint. These choices probably indicate that it was too costly or too difficult for these subjects to maximize the true preferences, rather than a genuine expression of preferences.

[Figure 4 here]

4 Econometric Analysis

4.1 Consistency

We next attempt to correlate the differences in consistency that arise in the experiment and demographic differences among experimental subjects. The personal data on panel members create the opportunity to analyze the correlation between demographic and economic characteristics and GARP violations. Table 4 below presents the results of our individual-level econometric analysis.[15] In column (1), we present estimates using ordinary least squares (OLS). In column (2), we repeat the estimation reported in column (1) using a Tobit model that accounts for the censored distribution of the CCEI. The two specifications yield similar results.[16] The results reported in column (1) and (2) show significant correlations. We obtain statistically significant coefficients in all demographic categories, ranging in absolute values from about 0.025 to just over 0.050. These magnitudes are large, implying that demographic differences can account for significant differential changes in wealth 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 by making inefficient choices.[17] In columns (3) and (4), we repeat the OLS and Tobit estimations reported in columns (1) and (2), after screening out the 124 subjects (10.5 percent) with FOSD scores below 0.9. In columns (5) and (6), we repeat the estimations, after screening out the 341 subjects (28.8 percent) with FOSD scores below 0.95. In columns (3) and (4), we find that on this reduced samples, the corresponding estimates are of comparable magnitude and in all categories still statistically significant. In columns (5) and (6), we also find that the corresponding estimates are of comparable magnitude and still statistically significant in the age and education categories.

[Table 4 here]

[15] The tables based on the indices proposed by Varian (1990, 1991) and Houtman and Maks (1985) are presented in Online Appendix II (http://emlab.berkeley.edu/~kariv/CKM_I_A2.pdf). In practice, all indices yield similar conclusions.

[16] To test for a potential misspecification, we calculated the RESET test of Ramsey (1969) 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).

[17] Agarwal, et al. (2009) document a U-shaped relationship between age and mistakes in financial decision making, suggesting that although 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 life-cycle.

We next turn to regression analyses that examine the patterns in the data more systematically. 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 may create a missing data problem. We correct for the possible sample selection bias in our econometric analysis below, using Heckman's (1979) method.[18] This method has been widely applied in empirical work because of its relative ease of use. Our *exclusion restriction* variable is the number of completed CentERpanel questionnaires as a fraction of the total invitations to participate in the three months prior to our experiment. Our identifying assumption is that this "participation ratio" influences the participation in our experiment but does not influence the laboratory outcomes of interest. The estimation results are reported in Table 5 below. In column (1), we omit the non-participants, focusing on the subsample of participants and dropouts in the data. In column (2), we repeat the estimation reported in columns (1) after adding the non-participants. We obtain very similar results on the reduced sample and the entire sample. More interestingly, the estimated parameters from the OLS and the sample selection estimations are virtually *identical*. Finally, testing the null 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 so there is no bias, as desired. This obviates the familiar concern of self-selection into the experiment. It is also noteworthy that in both specifications the coefficient on the exclusion restriction variable is positive and significant and that many demographic categories are positively correlated with participation.

[Table 5 here]

4.2 Dominance

We next turn to regression analyses that examine the patterns of FOSD violations in the data. Table 6 below presents the estimation results. In column (1), we present estimates using OLS. In column (2), we repeat the estimation reported in column (1) using a Tobit model. The two specifications yield similar results.[19] We obtain statistically significant coefficients only in

[18] We also use Heckman's sample selection model to analyze the correlates of the Varian (1990, 1991) measure. For the third measure, proposed by Houtman and Maks (1985), we estimate the sample selection model of Terza (1998). These results are provided in Online Appendix II.

[19] We again calculated the RESET test of Ramsey (1969) by adding the squared and cubed fitted values of the regression equation as additional regressors, and found no evi-

the age and education categories. We again use Heckman's (1979) method to correct for the possible sample selection bias. Our exclusion restriction variable is again the number of completed CentERpanel questionnaires as a fraction of the total invitations to participate in the three months prior to our experiment. The estimation results are reported in Table 7 below. In column (1), we omit the non-participants, focusing on the subsample of participants and dropouts in the data. In column (2), we repeat the estimation reported in columns (1) after adding the non-participants. We again find that ρ is indistinguishable from zero so there is no bias, as desired. Only the age coefficients are negative and significant. Overall, subject choices generally satisfy the dominance principle.[20]

[Table 6 here]
[Table 7 here]

5 Conclusion

Are choices rational? Behavioral economics raises intriguing questions about the rationality of individual choice. Nevertheless, rationality implicates every field of economics. It is meant to serve as a normative guide for choice and also as a descriptive model of how individuals choose. Developing methods for appropriately confronting the theory of rational choice with empirical or experimental evidence therefore has implications in many areas of economic theory and policy. A new field experimental design – employing graphical representations of standard consumer decision problems and utilizing a rich pool of subjects – enables us to collect richer data than has heretofore been possible. These data allow us to thoroughly analyze the correlates of individual levels of rationality. This is the unique and distinctive feature of the paper.

Empirical revealed preference research can tap either real-world data from large-scale markets or small-scale experimental data. Revealed preference tests have been applied to aggregate consumption data. However, real-world data do not 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

dence of misspecification (p-value = 0.2593).

[20] Charness, et al. (2007) study violations of first-order stochastic dominance in individual and group decision making. They find a significant number of violations and that the violation rates of groups are substantially lower and decreasing with group size.

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. We note that even random behavior can appear consistent if the sample size is small, as it often is in experimental studies.

The setup used in this experiment has the highest Bronars value of one (all random subjects had violations). Our large and rich menu of budget lines provides more opportunities to violate GARP and thus improves the power of nonparametric tests of revealed preference theory. We also use a much richer pool of subjects than the usual collection of undergraduate students, and the relation of consistency scores to individual characteristics also enables us to shed some light on the external validity of our findings. Finally, in the experimental task we study, subjects make decisions under conditions of uncertainty, which enters every realm of individual decision-making. Past research has typically investigated the rationality of choice in simpler decision tasks under certainty. When facing more complex decision tasks, individuals might not have the cognitive ability necessary to discover their optimal choice.

The relation of our experimental results to individual characteristics enables us to shed some 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. In addition to providing external validation for the results, our personal data creates the opportunity to analyze the correlates of experimental outcomes. Our paper thus contributes to the emerging literature on the relation of laboratory behaviors to cognitive ability, typically measured using IQ tests or SAT scores (see, for example, Benjamin, et al., 2006, and Dohmen, et al., 2010). By contrast, we use the single crucial test for "economic cognition" – consistency with utility-maximizing behavior – and investigate the correlation between consistency scores under laboratory conditions and demographic and economic characteristics. Also related to our choice under risk design, but somewhat further afield, there is a large and growing 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. In a concurrent paper, we use the same data set to investigate the correlation between risk attitudes and individual characteristics and attempt to explain heterogeneity in risk preferences in terms of demographic variables.[21]

The first striking fact from the experiment is the high level of consistency with utility maximization in the individual level decisions. Standard tests suggest that nearly half of our subjects exhibit behavior that appears to be almost optimizing in the sense that their choices are close to satisfying GARP. Additionally, subjects systematically display behaviors that are also sufficiently consistent with first-order stochastic dominance. Nevertheless, there is also substantial heterogeneity. The second striking fact is correlation between consistency levels and demographic and economic characteristics. This knowledge may ultimately prove to be useful for the formulation of economic policy. To name just a couple of relatively straightforward examples, the relationships between socio-demographics and levels of rationality (specifically through their heterogeneity) can better inform designing social programs (Manski, 2001) and paternalistic policies (Thaler and Sunstein, 2003).

References

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

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

[6] 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.

[21] von Gaudecker, et al. (forthcoming) also conducted risk experiments with CentER-panel 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."

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

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

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

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

[11] 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.

[12] 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.

[13] 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.

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

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

[16] 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.

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

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

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

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

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

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

[23] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Table 1. Socio-demographic 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			
Partner	80.88	67.89	82.64
# of kids	0.84	0.32	1.09
# of obs.	1182	190	968

Table 2. CCEI scores

	Mean	Sd	Percentiles 10	25	50	75	90	# of obs.
All	0.881	0.141	0.676	0.808	0.930	0.998	1.000	1,182
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								
Partner	0.878	0.142	0.673	0.802	0.927	0.998	1.000	956
Kids	0.899	0.128	0.704	0.835	0.959	1.000	1.000	490

Table 3. FOSD scores

	Mean	Sd	Percentiles					# of obs.
			10	25	50	75	90	
All	0.959	0.951	0.998	0.992	0.977	0.944	0.897	1,182
Female	0.961	0.957	0.998	0.991	0.977	0.945	0.905	537
Age								
16-34	0.966	0.951	1.000	0.997	0.986	0.953	0.904	219
35-49	0.969	0.958	0.999	0.995	0.985	0.963	0.910	309
50-64	0.953	0.949	0.996	0.988	0.967	0.937	0.896	421
65+	0.949	0.948	0.995	0.988	0.965	0.926	0.874	233
Education								
Low	0.953	0.951	0.996	0.989	0.969	0.936	0.886	397
Medium	0.961	0.956	0.998	0.991	0.977	0.948	0.906	351
High	0.963	0.947	1.000	0.995	0.984	0.948	0.901	430
Household monthly inco								
€0-2500	0.955	0.953	0.996	0.988	0.972	0.937	0.888	269
€2500-3499	0.960	0.953	0.997	0.991	0.977	0.948	0.909	302
€3500-4999	0.958	0.948	0.999	0.993	0.978	0.941	0.892	345
€5000+	0.962	0.948	0.999	0.994	0.982	0.953	0.897	266
Occupation								
Paid work	0.964	0.954	0.999	0.993	0.982	0.949	0.907	628
House work	0.957	0.952	0.999	0.991	0.976	0.941	0.888	137
Retired	0.948	0.948	0.995	0.986	0.963	0.928	0.876	247
Others	0.957	0.944	0.999	0.992	0.978	0.946	0.887	170
Household composition								
Partner	0.958	0.951	0.998	0.992	0.977	0.942	0.896	956
Kids	0.962	0.951	0.999	0.993	0.982	0.952	0.901	490

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

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	Tobit	OLS	Tobit	OLS	Tobit
Constant	.887***	.916***	.904***	.929***	.937***	.960***
	(.022)	(.029)	(.021)	(.026)	(.018)	(.024)
Female	-.024***	-.031***	-.017**	-.022**	-.011*	-.014
	(.009)	(.011)	(.008)	(.010)	(.007)	(.009)
Age						
35-49	-.016	-.026	-.016	-.026*	-.022**	-.033***
	(.011)	(.016)	(.011)	(.014)	(.009)	(.012)
50-64	-.052***	-.074***	-.045***	-.065***	-.036***	-.054***
	(.011)	(.016)	(.010)	(.014)	(.009)	(.013)
65+	-.051**	-.067***	-.045**	-.059***	-.026*	-.037*
	(.020)	(.024)	(.018)	(.022)	(.015)	(.020)
Education						
Medium	.009	.013	.008	.011	.002	.005
	(.011)	(.013)	(.010)	(.012)	(.008)	(.011)
High	.026**	.036***	.024**	.032***	.018**	.025**
	(.011)	(.013)	(.010)	(.012)	(.008)	(.011)
Income						
€2500-3499	.026**	.035**	.018	.026*	.008	.016
	(.012)	(.014)	(.011)	(.013)	(.009)	(.012)
€3500-4999	.020	.028*	.019	.026*	.011	.018
	(.013)	(.015)	(.012)	(.014)	(.010)	(.012)
€5000+	.033**	.047***	.031**	.043***	.010	.018
	(.014)	(.017)	(.013)	(.015)	(.011)	(.014)
Occupation						
Paid work	.028	.039**	.020	.028	.022	.033*
	(.018)	(.020)	(.015)	(.018)	(.013)	(.017)
House work	.047**	.064***	.036**	.050**	.027*	.040**
	(.021)	(.023)	(.018)	(.020)	(.016)	(.019)
Others	.037*	.046**	.013	.018	.005	.010
	(.019)	(.022)	(.017)	(.021)	(.015)	(.019)
Household composition						
Partner	-.026**	-.039***	-.023**	-.032**	-.017**	-.025**
	(.011)	(.014)	(.010)	(.013)	(.008)	(.011)
# of kids	.001	.003	.001	.003	.003	.004
	(.004)	.006	(.004)	(.005)	(.003)	(.004)
σ		.165		.142		.112
		(.004)		(.004)		(.003)
R^2/ log-likelihood	.068	117.52	.066	218.95	.061	262.15
# of obs.	1182	1182	1058	1058	841	841

Omitted categories: male, age under 35, low education (primary and lower secondary education), household gross monthly income under €2500, retired, and not having a partner. Standard errors in parentheses. *, **, *** indicate 10, 5, 1 percent significance levels, respectively.

Table 5. The correlation between the 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				
Partner	-.026**	.262**	-.027**	.123
	(.011)	(.119)	(.011)	(.092)
# of kids	.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	

Omitted categories: male, age under 35, low education (primary and lower secondary education), household gross monthly income under €2500, retired, and not having a partner. Standard errors in parentheses. *, **, *** indicate 10, 5, 1 percent significance levels, respectively.

Table 6. The correlation between the FOSD scores and subjects' individual characteristics (OLS and Tobit)

	(1) OLS	(2) Tobit
Constant	.957***	.959***
	(.008)	(.008)
Female	.003	.002
	(.003)	(.003)
Age		
35-49	.002	.002
	(.004)	(.005)
50-64	-.013***	-.015***
	(.004)	(.005)
65+	-.014**	-.016**
	(.007)	(.007)
Education		
Medium	.005	.006
	(.004)	(.004)
High	.007*	.008**
	(.004)	(.004)
Income		
€2500-3499	.004	.004
	(.004)	(.004)
€3500-4999	-.000	-.000
	(.004)	(.004)
€5000+	.002	.002
	(.005)	(.005)
Occupation		
Paid work	.006	.007
	(.006)	(.006)
House work	.005	.005
	(.006)	(.007)
Others	.004	.004
	(.007)	(.007)
Household composition		
Partner	-.002	-.002
	(.004)	(.004)
# of kids	-.001	-.002
	(.002)	(.002)
σ		.050
		(.001)
R^2/ log-likelihood	0.037	1772.31
# of obs.	1182	1182

Omitted categories: male, age under 35, low education (primary and lower secondary education), household gross monthly income under €2500, retired, and not having a partner. Standard errors in parentheses. *, **, *** indicate 10, 5, 1 percent significance levels, respectively.

Table 7. The correlation between the FOSD scores and subjects' individual characteristics (sample-selection)

	(1)		(2)	
	Outcome	Selection	Outcome	Selection
Constant	.964*** (.009)	.544* (.314)	.963*** (.009)	-2.062*** (.208)
Female	.002 (.003)	.084 (.104)	.003 (.003)	-.033 (.068)
Age				
35-49	.004 (.005)	-.554** (.223)	.001 (.004)	-.138 (.101)
50-64	-.009 (.005)	-1.023*** (.220)	-.014*** (.004)	-.398*** (.102)
65+	-.005 (.009)	-1.557*** (.258)	-.013* (.007)	-.822*** (.153)
Education				
Medium	.004 (.004)	.191 (.120)	.005 (.004)	-.037 (.081)
High	.006 (.004)	.169 (.117)	.007 (.004)	.006 (.084)
Income				
€2500-3499	.002 (.005)	.304** (.127)	.003 (.004)	.274*** (.094)
€3500-4999	-.003 (.005)	.428*** (.138)	-.001 (.004)	.171* (.094)
€5000+	.002 (.005)	.065 (.145)	.003 (.005)	.072 (.106)
Occupation				
Paid work	.008 (.006)	-.203 (.173)	.007 (.006)	-.033 (.130)
House work	.004 (.007)	.109 (.205)	.004 (.006)	.075 (.147)
Others	.003 (.007)	.081 (.193)	.003 (.007)	.111 (.145)
Household composition				
Partner	-.004 (.004)	.261** (.115)	-.003 (.004)	.127 (.091)
# of kids	-.002 (.002)	.145** (.062)	-.001 (.001)	.027 (.035)
Participation ratio		1.230*** (.234)		3.373*** (.126)
ρ	-.503		-.193 (.107)	
R^2/ log-likelihood	0.039	-471.961	850.503	
# of obs.	1372		2340	

Omitted categories: male, age under 35, low education (primary and lower secondary education), household gross monthly income under €2500, retired, and not having a partner. Standard errors in parentheses. *, **, *** indicate 10, 5, 1 percent significance levels, respectively.

Figure 1. Mean CCEI scores

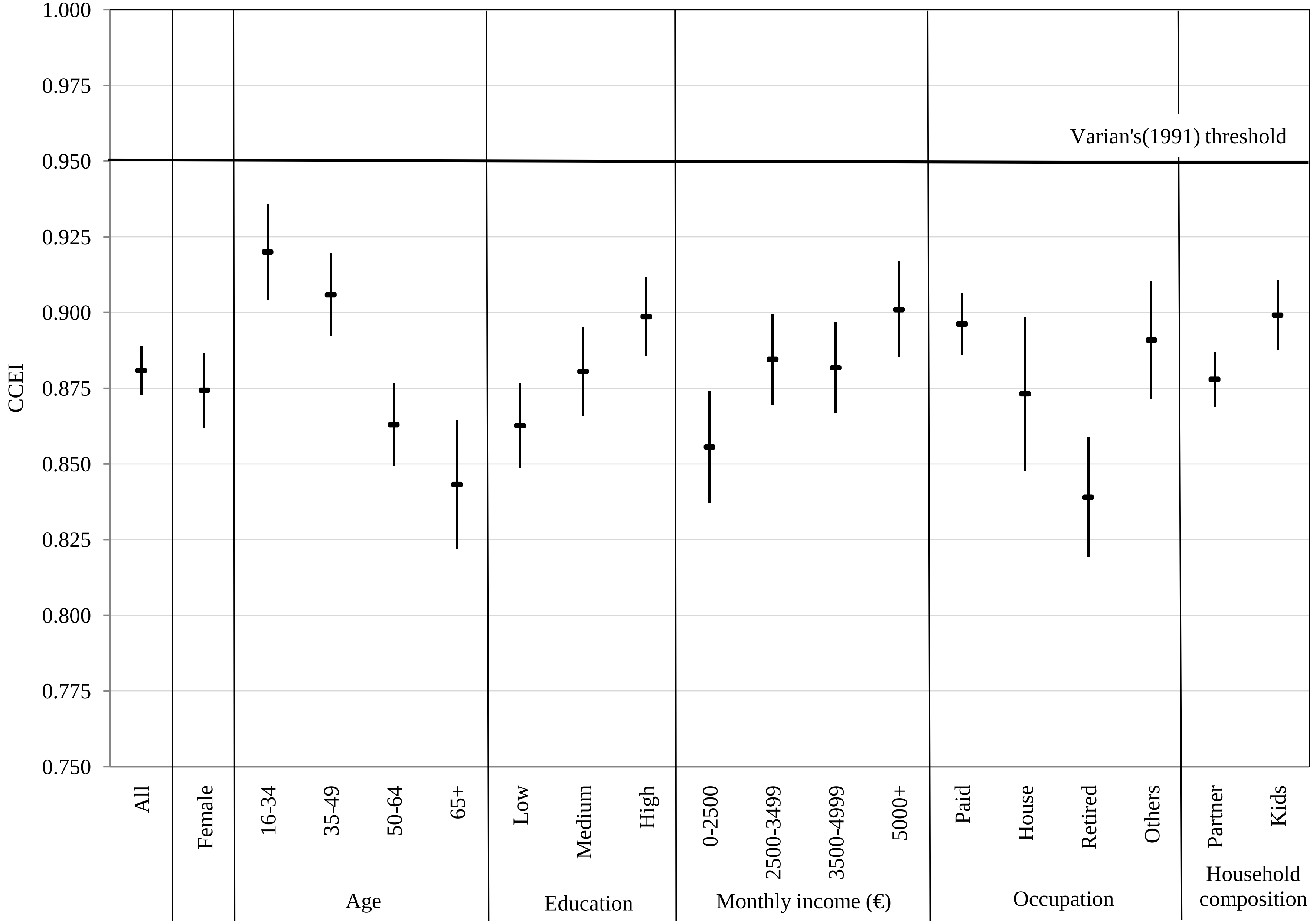


Figure 2. The distributions of CCEI scores of simulated subjects

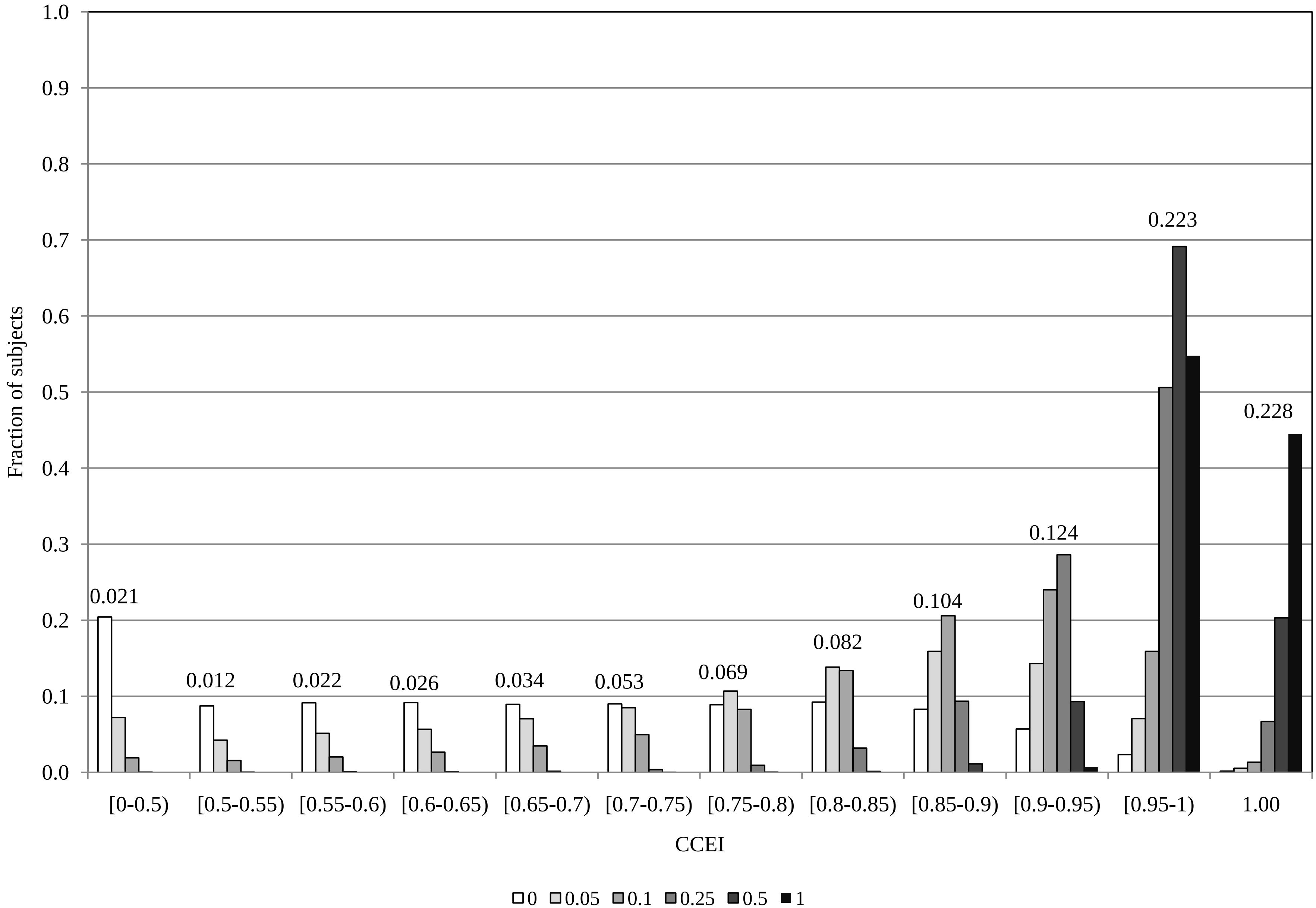
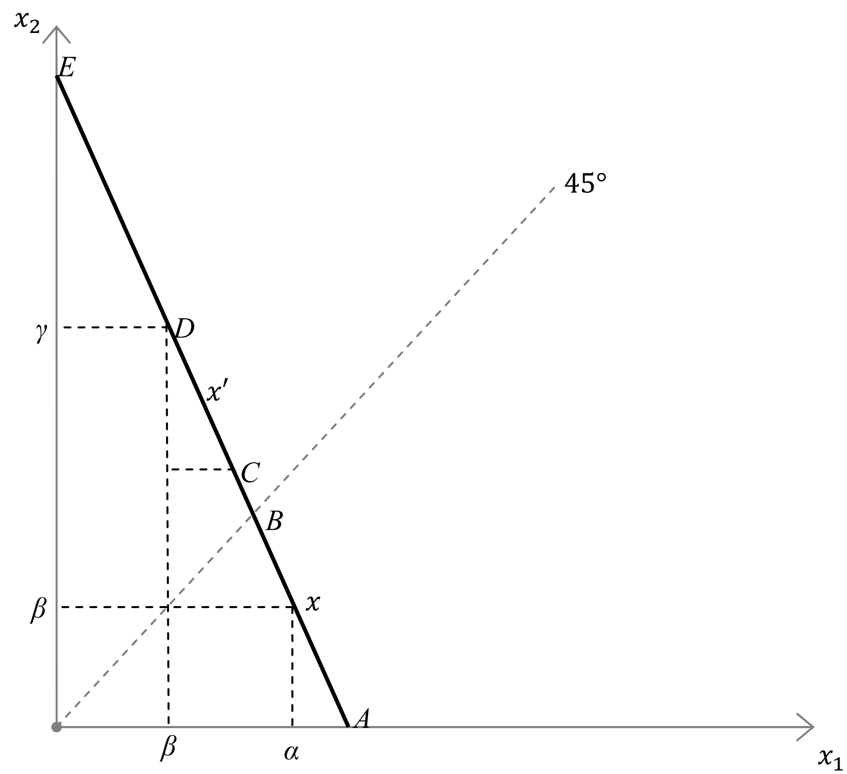


Figure 3. A violation of stochastic dominance



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 4. A scatterplot of CCEI and FOSD scores