

Who Is (More) Rational?

Author(s): Syngjoo Choi, Shachar Kariv, Wieland Müller and Dan Silverman

Source: *The American Economic Review*, Vol. 104, No. 6 (JUNE 2014), pp. 1518-1550

Published by: American Economic Association

Stable URL: <http://www.jstor.org/stable/42920858>

Accessed: 16-01-2018 06:30 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://about.jstor.org/terms>



JSTOR

*American Economic Association* is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

## Who Is (More) Rational?<sup>†</sup>

By SYNGJOO CHOI, SHACHAR KARIV, WIELAND MÜLLER, AND DAN SILVERMAN\*

*Revealed preference theory offers a criterion for decision-making quality: if decisions are high quality then there exists a utility function the choices maximize. We conduct a large-scale experiment to test for consistency with utility maximization. Consistency scores vary markedly within and across socioeconomic groups. In particular, consistency is strongly related to wealth: A standard deviation increase in consistency is associated with 15–19 percent more household wealth. This association is quantitatively robust to conditioning on correlates of unobserved constraints, preferences, and beliefs. Consistency with utility maximization under laboratory conditions thus captures decision-making ability that applies across domains and influences important real-world outcomes. (JEL D12, D14, D81, D83, D91, G11)*

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 test whether individual behavior in a choice under risk experiment is consistent with the utility maximization model. We conducted the experiment using the CentERpanel, a panel study of a large representative sample of households in the Netherlands that collects a wide range of individual sociodemographic and economic information about its members.

\*Choi: Department of Economics, University College London, Gower Street, London WC1E 6BT, UK (e-mail: syngjoo.choi@ucl.ac.uk); Kariv: Department of Economics, University of California–Berkeley, 530 Evans Hall 3880, Berkeley, CA 94720-3880 (e-mail: kariv@berkeley.edu); Müller: Department of Economics and VCEE, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria, and Department of Economics, CentER, and TILEC, Tilburg University (e-mail: wieland.mueller@univie.ac.at); Silverman: Department of Economics, W. P. Carey School of Business, Arizona State University, PO Box 879801, Tempe, AZ 85287-9801, and NBER (e-mail: dsilver3@asu.edu). We thank three anonymous referees for their important input on the paper, and Douglas Gale and Raymond Fisman for many 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, 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. Approval for this study was obtained from the Institutional Review Board at the UC Berkeley. We acknowledge financial support from the Economic and Social Research Council (ESRC) Grant No. RES-061-25-0348 (Choi); the National Science Foundation (SES-0962543), the Multidisciplinary University Research Initiative (MURI), and 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) (Müller). The authors have no financial or other material interests related to this research to disclose.

<sup>†</sup>Go to <http://dx.doi.org/10.1257/aer.104.6.1518> to visit the article page for additional materials and author disclosure statement(s).

The article provides three types of analysis. First, we offer a purely descriptive overview of the consistency of the experimental data with the utility maximization model. Second, we analyze the correlation between the levels of consistency and socioeconomic characteristics. In this way we address the question: “who is (more) rational?” Third, we test the validity of a causal interpretation of the relationship between our proposed measure of decision-making quality—the consistency of the experimental data with utility maximization—and wealth accumulation in the real world.

In accordance with Samuelson’s approach, traditional economic analysis assumes that individual behavior can be rationalized by a utility function. In this standard view, heterogeneity in choices is attributed to heterogeneity in preferences, constraints, information, or beliefs. More recently, several strands of research consider heterogeneity in choices driven also by differences in decision-making ability. Different from traditional analysis, this literature allows that the choices that some people actually make may be different from the choices they would make if they had the skills or knowledge to make better decisions. This research thus takes the view that those with lower decision-making *ability* may make choices of lower decision-making *quality*.

The idea that people vary in their decision-making ability, and therefore make choices of different decision-making quality, has intuitive appeal and important consequences for economics. It is difficult, however, to make definitive judgments about which choices exhibit low decision-making quality, and which people have inferior decision-making ability, due to twin problems of *identification* and *measurement*. The identification problem is to distinguish differences in decision-making ability from unobserved differences in preferences, constraints, information, or beliefs. The measurement problem is to define and implement a practical, portable, quantifiable, and economically interpretable measure of decision-making quality. The two problems are conceptually distinct, but tightly linked in practice.

The identification problem emerges because it is usually unclear whether those with lower decision-making ability—as evidenced by less education, lower cognitive abilities, or less financial literacy—are making choices of lower decision-making quality. They might have different preferences over the same outcomes, or face different but unobserved incentives and constraints, or have different information, or hold different beliefs. The measurement problem emerges because only rarely are the relevant incentives so clear and the data quality so high that classifying some choices as of low decision-making quality is straightforward and uncontroversial. Typically, a measure of decision-making quality is challenging to formalize, quantify, and make practical for application in a variety of choice environments.

We offer a new approach to the challenges posed by the identification and measurement problems. Our point of departure is a proposal to measure decision-making quality by the consistency of choices with economic rationality, in the sense of a complete and transitive preference ordering. Adopting this standard for decision-making quality, we present individuals with an economic choice experiment which provides a stringent test of utility maximization. The measure thus has a well-established economic interpretation, and revealed preference theory tells us whether we have enough data to make it statistically useful. In addition, the analytical techniques and experimental platform are easily portable to a variety of choice problems. Our approach thus addresses the measurement problem.

Furthermore, an experiment like ours, to the extent possible, holds information and beliefs constant within subject and controls the relevant constraints. If decision-making ability is defined simply as the capacity to make choices of higher decision-making quality, then our approach addresses the identification problem inside the laboratory. It can distinguish individual heterogeneity in decision-making ability from unobserved differences in preferences, constraints, information, or beliefs.

Our interest in decision-making quality under laboratory conditions largely derives, however, from the possibility that it reflects decision-making ability that affects important outcomes outside the laboratory. We evaluate this possibility first by examining the correlation between decision-making quality in the experiment and socioeconomic characteristics. The goal of this analysis is not to establish causation. If, however, we find a significant correlation between decision-making quality and certain characteristics, this lends a basic level of credence to the idea that people with these characteristics tend to make different choices not just because they face different constraints or have different preferences, but also because they tend to have different levels of decision-making ability.

To evaluate a causal interpretation of the link between decision-making quality in the experiment and important economic outcomes in the real world, we examine whether our measure of decision-making quality from the experiment can independently and robustly explain real-world economic outcomes, conditional on socioeconomic characteristics including income, age, education, and occupation. If heterogeneity in decision-making ability is an important source of heterogeneity in real-world outcomes, and if decision-making quality in the experiment is a good proxy for decision-making ability, then differences in the experiment-based measure across subjects should help explain differences in their real-world outcomes.

We chose wealth as the outcome of interest because the task of explaining wealth provides a strong test of the idea that decision-making quality in the experiment reflects a more general form of decision-making ability. The test is strong because wealth is determined by countless decisions, made over time in many different settings, and involving many different trade-offs, thus increasing our chance of rejecting a relationship. We are also motivated to study wealth by prior research that documents large wealth differentials among households with similar lifetime 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, Skinner, and Weinberg 2001; Ameriks, Caplin, and Leahy 2003; and Scholz, Seshadri, and Khitatrakun 2006 for different perspectives).

A causal interpretation of the relationship between decision-making quality in the experiment and wealth depends importantly on the estimated correlation between these two measures being quantitatively robust to conditioning on additional correlates of unobserved preferences, constraints, information, and beliefs. This assessment of robustness is not an effort to “control for everything.” Instead, in the spirit of Altonji, Elder, and Taber (2005), we examine whether the estimated correlation is much affected by the inclusion of additional controls that, *a priori*, should be correlated with economic outcomes through their correlation with unobserved or misspecified variables. If these unobservables are indeed important sources of the

observed correlation between consistency and economic outcomes, then adding the controls should have a substantial effect on the estimated correlation coefficients.

We measure the decision-making quality in the experiment by evaluating the consistency of individual choices with the Generalized Axiom of Revealed Preference (GARP). We assess how nearly individual behavior complies with GARP using standard measures of consistency that have been proposed for quantifying the extent of violations. There is considerable heterogeneity within and across sociodemographic groups. Taking advantage of the large and heterogeneous CentERpanel sample, we find that high-income and high-education subjects display greater levels of consistency than lower-income and lower-education subjects. In addition, men are more consistent than women, and young subjects tend more toward utility maximization than those who are old. The magnitudes imply that low-income subjects on average “leave on the table” as much as 3.3 percentage points more of their earnings, relative to high-income subjects, by making inconsistent choices. The corresponding numbers for low-education subjects, females, and old subjects are 2.6, 2.4, and 5.1, respectively.

We also find an economically large and statistically significant correlation between consistency in the experiment and household wealth. The point estimates indicate that, conditional on measures of current income, age, education, occupation, basic demographic characteristics, and household structure, a standard deviation increase in the consistency score of the person who is primarily responsible for household financial matters is associated with 15–19 percent more household wealth. As important, this estimated correlation is quantitatively robust to conditioning on many additional correlates of unobserved preferences, constraints, information, and beliefs. We interpret the economically large, statistically significant, and quantitatively robust relationship between decision-making quality in the experiment—the consistency of the experimental data with the utility maximization model—and household wealth as evidence of decision-making ability that applies across choice domains and affects important real-world outcomes.

We also show that alternative measures of decision-making quality from the experiment and decision-making ability from the CentERpanel survey are not substitutes for our measure of compliance with GARP. The alternatives include (i) a stronger notion of decision-making quality in our experiment that measures the extent of violations of both GARP and first-order stochastic dominance; (ii) parametric estimates of a tendency to “tremble” in an experiment by von Gaudecker, van Soest, and Wengström (2011) with an overlapping sample of CentERpanel members; and (iii) scores from Frederick’s (2005) Cognitive Reflection Test that, in other samples, is well correlated with measures of cognitive ability. We find that these three alternatives either have no independent power to predict wealth, or are not well correlated with compliance with GARP in the experiment. Finally, we investigate the correlation between decision-making quality in the experiment and details of household saving allocations that influence wealth.

The rest of the article is organized as follows. Section I describes the experimental design and procedures. Section II describes decision-making quality in the experimental data. Section III contains analysis of the correlation between decision-making quality and socioeconomic characteristics. Section IV discusses the relationship between wealth differentials and decision-making ability. Section V

describes the margins along which we extend the previous literature. Section VI contains some concluding remarks. The article also includes six online data and technical Appendices for the interested reader.

## I. The Experiment

### A. Sample

The experiment uses the CentERpanel, an online, weekly, and stratified survey of a sample of over 2,000 households and 5,000 individual members. The sample is designed to be representative of the Dutch-speaking population in the Netherlands. Via the Internet, the survey instrument allows researchers to implement experiments and collects a great deal of individual demographic and economic information from its respondents. The subjects in the experiment were recruited at random from the entire CentERpanel sample. The experiment was conducted online with 1,182 CentERpanel adult members. Table 1 provides summary statistics of individual characteristics. We present the data for *participants* (completed the experiment), *dropouts* (logged in but quit the experiment), and *nonparticipants* (recruited for the experiment but never logged in). In later analysis we will control for sample selection using the Heckman (1979) model.

### B. Design

In our experiment, we present subjects with a sequence of decision problems under risk. Each decision problem was presented as a choice from a two-dimensional budget line. A choice of the allocation from the budget line represents an allocation of points 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; the subject received the points allocated to one of the accounts  $x$  or  $y$ , determined at random and equally likely. An example of a budget line defined in this way is the line  $AB$  drawn in Figure 1. The point  $C$ , which lies on the 45-degree line, corresponds to the equal allocation with certain outcome, whereas points  $A$  and  $B$  represent allocations in which all points are allocated to one of the accounts. Notice that points along  $AC$  are risky—they have a lower payoff in state  $x$  and a higher payoff in state  $y$ —but because the slope of the budget line  $AB$  is steeper than  $-1$ , they have higher expected return than point  $C$ . By contrast, points along  $BC$  have lower expected return than point  $C$ .

### C. Procedures

The procedures described below are identical to those used by Choi et al. (2007), with the exception that the experiment described here consisted of 25, rather than 50, decision problems.<sup>1</sup> We also made some minor changes to accommodate the online

<sup>1</sup>The number of individual decisions is still higher than usual in the literature, and revealed preference analysis presented below shows the experiment provides a dataset consisting of enough individual decisions over a sufficiently wide range of budget lines to provide a powerful test of consistency.

TABLE 1—SOCIODEMOGRAPHIC INFORMATION

	Participants	Dropouts	Nonparticipants
<i>Female</i>	45.43	37.89	50.00
<i>Age</i>			
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
<i>Education</i>			
Low	33.59	42.63	30.99
Medium	29.70	22.63	31.61
High	36.72	34.74	37.40
<i>Household monthly income</i>			
€0–2,500	22.76	40.00	22.62
€2,500–3,499	25.55	22.11	18.18
€3,500–4,999	29.19	15.79	28.51
€5,000+	22.50	22.11	30.68
<i>Occupation</i>			
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
<i>Household composition</i>			
Partner	80.88	67.89	82.64
Number of children	0.84	0.32	1.09
Observations	1,182	190	968

*Notes:* 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 prevocational secondary education, preuniversity 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 levels of completed education and occupations are based on the categorization of Statistics Netherlands (Centraal Bureau voor de Statistiek).

experimental setting. Each decision problem started with the computer selecting a budget line randomly from the set of budget lines that intersect with at least one of the axes at 50 or more points, but with no intercept exceeding 100 points. The 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.<sup>2</sup> Choices were made using the computer mouse to move the pointer on the computer screen to the desired point and then clicking the mouse or hitting the enter key. More information and full experimental instructions, including the computer program dialog window, are available in online Appendix I.

<sup>2</sup>Like Choi et al. (2007), we restricted choices to allocations on the budget line so that subjects could not dispose of payoffs. In Fisman, Kariv, and Markovits (2007), each decision involved choosing a point on a graph representing a budget set that included interior 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 their choices that were on the budget line), we restricted choices to allocations on the budget constraint to make the computer program easier to use.

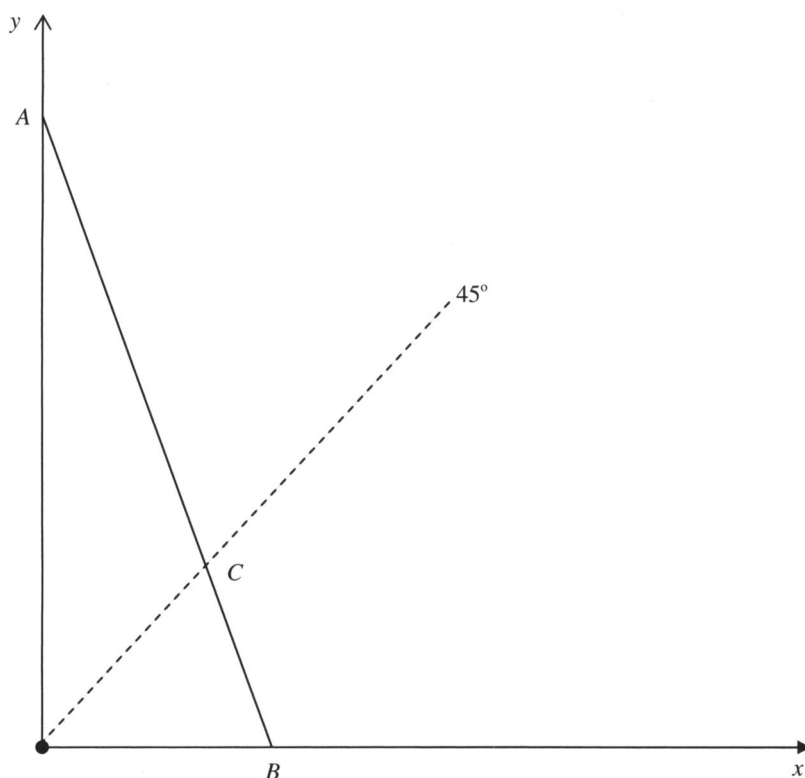


FIGURE 1. AN ILLUSTRATION OF THE BUDGET LINE

During 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. 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.

## II. Decision-Making Quality

We propose to measure decision-making quality by the consistency of choices with economic rationality, and we described a simple economic choice experiment in which we can measure decision-making quality with a high degree of precision, and separate it from other sources of heterogeneity in choice. Specifically, we employ the GARP to test whether the finite set of observed price and quantity data that our experiment generated may be rationalized by a utility function. GARP generalizes various revealed preference tests. It requires that if allocation  $\mathbf{x}^i$  is revealed preferred to  $\mathbf{x}^j$ , then  $\mathbf{x}^j$  is not *strictly* and directly revealed preferred to  $\mathbf{x}^i$ ;



that is, allocation  $\mathbf{x}^i$  must cost at least as much as  $\mathbf{x}^j$  at the prices prevailing when  $\mathbf{x}^j$  is chosen.<sup>3</sup>

If choices are generated by a nonsatiated utility function, then the data must satisfy GARP. Conversely, the result due to Afriat (1967) tells us that if a *finite* dataset generated by an individual's choices satisfies GARP, then the data can be rationalized by a utility function. Consistency with GARP has long been a touchstone for rationality, but it demands only a complete and transitive preference ordering. It places no restrictions on the utility function and makes no assumptions about what is reasonable to maximize. Since it is possible to “pump” an indefinite amount of money out of an individual making intransitive decisions, consistency with GARP provides a crucial test of decision-making quality.

An individual's decisions may violate GARP and thus be of low decision-making quality for a number of reasons. First, violations of GARP can result from “trembles.” Subjects may compute payoffs incorrectly, execute intended choices incorrectly, or err in other less obvious ways. Second, inconsistency can result from bounded rationality or cognitive biases such as “framing effects” and “mental accounting” (Kahneman and Tversky 1984). The resources needed for determining optimal choices are limited. Thus, especially in complex or unfamiliar environments, the cost of computing an optimal decision can be high. Some subjects may therefore adopt simple decision rules, and this “simplification” may cause their choices to be inconsistent.<sup>4</sup> Third, if the data  $(\mathbf{p}^i, \mathbf{x}^i)$  satisfy GARP, then they can be rationalized by an outcome-based utility function  $U(x_1, x_2)$ . However, unobserved factors could enter the utility function so the “true” underlying preference ordering is represented by a utility function  $U(x_1, x_2, \omega)$  parametrized by  $\omega$ . If the data  $(\mathbf{p}^i, \mathbf{x}^i)$  are generated by a utility function  $U(x_1, x_2, \omega)$  and  $\omega$  is fixed, then the data will still satisfy GARP. An example of this is the disappointment aversion model proposed by Gul (1991), where the safe allocation  $x_1 = x_2$  is the reference point. If, however,  $\omega$  is not fixed then subjects can exhibit “preference reversals,” and the data  $(\mathbf{p}^i, \mathbf{x}^i)$  might not satisfy GARP (cf. Bernheim and Rangel 2009 where  $\omega$  may be time). The variable  $\omega$  may also be interpreted, as in Kőszegi and Rabin (2006), as a *dynamic* reference point determined endogenously by the environment.

### A. Consistency with GARP

Although testing conformity with GARP is conceptually straightforward, there is a 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 the Afriat (1972) Critical Cost Efficiency Index (CCEI), which measures the fraction by which all budget constraints must be shifted in order

<sup>3</sup>Without loss of generality, assume the individual's payout is normalized to 1. The budget line is then  $p_1x_1 + p_2x_2 = 1$ , and the individual can choose any allocation  $\mathbf{x}$  that satisfies this constraint. Let  $(\mathbf{p}^i, \mathbf{x}^i)$  be the data generated by an individual's choices, where  $\mathbf{p}^i$  denotes the  $i$ th observation of the price vector and  $\mathbf{x}^i$  denotes the associated allocation. An allocation  $\mathbf{x}^i$  is directly revealed preferred to  $\mathbf{x}^j$  if  $\mathbf{p}^i\mathbf{x}^i \geq \mathbf{p}^i\mathbf{x}^j$ . An allocation  $\mathbf{x}^i$  is revealed preferred to  $\mathbf{x}^j$  if there exists a sequence of allocations  $\{\mathbf{x}^k\}_{k=1}^K$  with  $\mathbf{x}^1 = \mathbf{x}^i$  and  $\mathbf{x}^K = \mathbf{x}^j$ , such that  $\mathbf{x}^k$  is directly revealed preferred to  $\mathbf{x}^{k+1}$  for every  $k = 1, \dots, K - 1$ .

<sup>4</sup>Consistency is also endogenous: subjects can make decisions that are consistent with GARP in a complex decision problem because they adopt simple choice rules to cope with complexity. But in that case, the “revealed” preference ordering may not be the “true” underlying preference ordering.

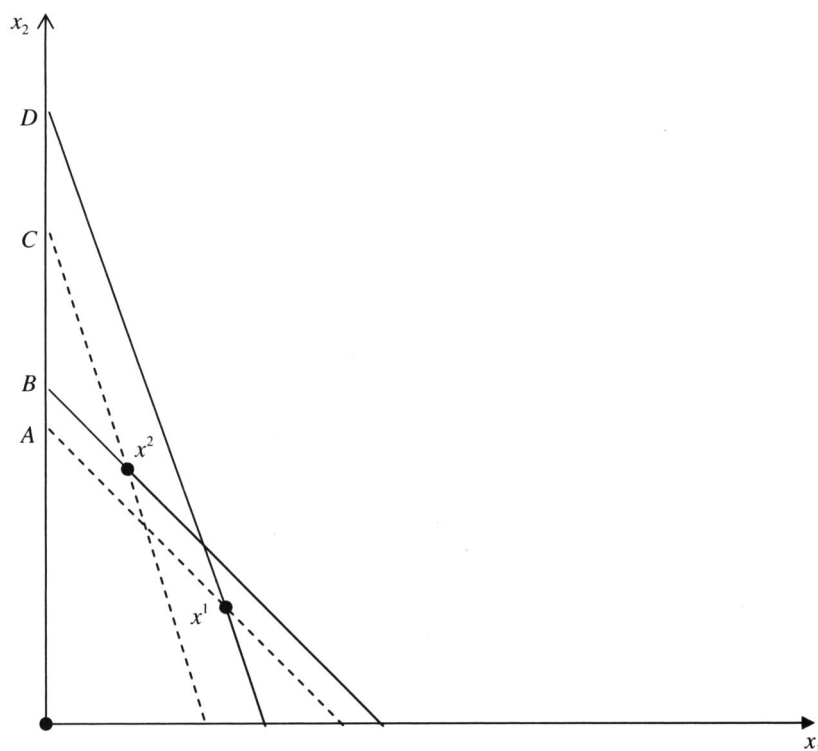


FIGURE 2. THE CONSTRUCTION OF THE CCEI FOR A SIMPLE VIOLATION OF GARP

*Notes:* Here we have a violation of the Weak Axiom of Revealed Preference (WARP) since  $x^1$  is directly revealed preferred to  $x^2$  and  $x^2$  is directly revealed preferred to  $x^1$ . A perturbation  $A/B < C/D$  of the budget line through allocation  $x^1$  removes the violation.

to remove *all* violations of GARP. Put precisely, for any number  $0 \leq e \leq 1$ , define the direct revealed preference relation

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

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  $e^*$  associated with the dataset.

Figure 2 illustrates the construction of the CCEI for a simple violation of the GARP involving two allocations,  $\mathbf{x}^1$  and  $\mathbf{x}^2$ . If we shift the budget line through  $\mathbf{x}^1$  as shown ( $A/B < C/D$ ) the violation would be removed so the CCEI score associated with this violation is  $A/B$ . By definition, the CCEI is between zero and one—the closer the CCEI is to one, the smaller the perturbation of the budget constraints required to remove all violations and the closer the data are to satisfying GARP. The CCEI thus provides a summary statistic of the overall consistency with GARP, reflecting the minimum adjustment required to eliminate all violations of GARP associated with the dataset.

In our experiment, the CCEI scores averaged 0.881, which implies that on average budget lines needed to be reduced by about 12 percent to eliminate a subject's

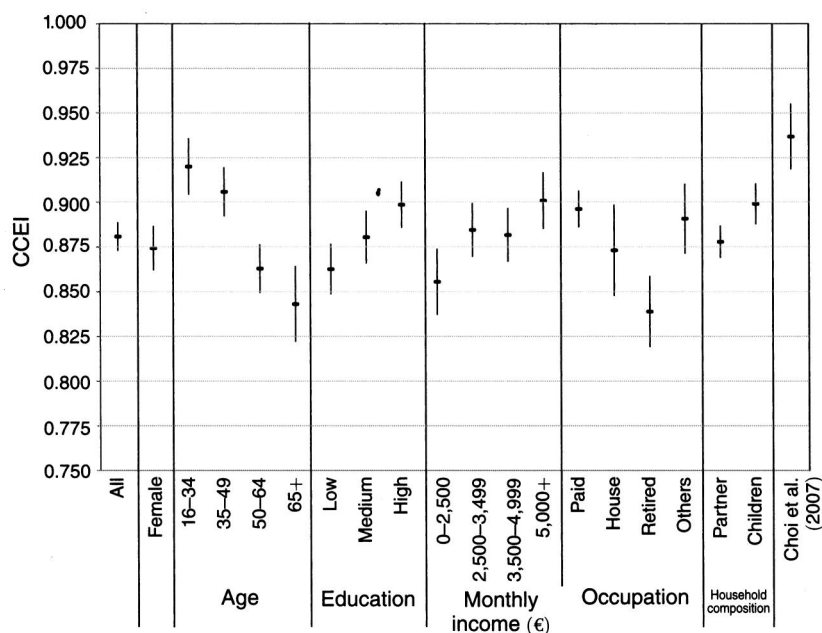


FIGURE 3. MEAN CCEI SCORES

GARP violations.<sup>5</sup> There is also marked heterogeneity in the CCEI scores within and across socioeconomic groups. Figure 3 summarizes the mean CCEI scores and 95 percent confidence intervals across selected socioeconomic categories. On average, high-income and high-education subjects display greater levels of consistency than lower-income and lower-education subjects. Men are more consistent than women, and young subjects tend more toward utility maximization than those who are old. The magnitudes imply that, in order to eliminate their GARP violations, low-income subjects require an average contraction of their budgets that is 3.3 percentage points larger than that of high-income subjects. The corresponding numbers for low-education subjects, females, and old subjects are 2.6, 2.4, and 5.1, respectively.<sup>6</sup> We will further analyze the relationship between consistency scores and socioeconomic characteristics, and address sample selection, in our regression analysis below.

A key advantage of the CCEI is its tight connection to economic theory. This connection makes the CCEI economically quantifiable and interpretable. Moreover, the same economic theory that inspires the measure also tells us when we have enough data to make it statistically useful. Thus, this theoretically grounded measure of decision-making quality helps us design and interpret the experiments in several

<sup>5</sup>In contrast, Choi et al. (2007) 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 0.954 (see Figure 1). The subjects of Choi et al. (2007) were undergraduate students and staff at UC Berkeley, and the experiment was conducted in the laboratory. We note that the subjects of Choi et al. (2007) were given a menu of 50 budget sets which provides a more stringent test of GARP.

<sup>6</sup>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).

ways. In online Appendix II, we provide more details on testing for consistency with GARP, discuss the power of the revealed preference tests, explain other indices that have been proposed for this purpose, and describe the related empirical literature on revealed preferences. In reporting our results, we focus on the CCEI. The results based on alternative indices are presented in online Appendix III. In practice, all indices yield similar conclusions.

### B. Beyond Consistency

*Stochastic Dominance.*—Consistency with GARP requires consistent preferences over all possible alternatives, but any consistent preference ordering is admissible. In this way, we see consistency as a necessary, but not sufficient, condition for choices to be considered of high decision-making quality. Indeed, choices can be consistent with GARP and yet fail to be reconciled with any utility function that is normatively appealing. For example, consider an individual who always allocates all points to the same account as measured by  $x_1$ . This behavior is consistent with maximizing the utility function  $U(x_1, x_2) = x_1$  and would generate a CCEI score of one.

Such preferences are, however, hard to justify in this setting because for many of the budget lines that a subject faces, always allocating all points to the same account means allocating all points to the more expensive account, a violation of monotonicity with respect to first-order stochastic dominance. Violations of 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. Stochastic dominance is, thus, a compelling criterion for decision-making quality and is generally accepted in decision theory (Quiggin 1990, and Wakker 1993).

To provide a unified measure of the extent of GARP violations *and* violations of stochastic dominance (for a given subject), we combine the actual data from the experiment and the *mirror image* of these data obtained by reversing the prices and the associated allocation for each observation. That is, we assume that if  $(x_1, x_2)$  is chosen subject to the budget constraint  $p_1x_1 + p_2x_2 = 1$ , then  $(x_2, x_1)$  would have been chosen subject to the mirror-image budget constraint  $p_2x_1 + p_1x_2 = 1$ . We then compute the CCEI for this combined dataset and compare that number to the CCEI for the actual data. The CCEI score for the combined data consisting of 50 observations can be no bigger than the CCEI score for the actual data.<sup>7</sup>

For example, recall the preferences represented by  $U(x_1, x_2) = x_1$ , where the individual always allocates all points to  $x_1$ . Such choices are perfectly consistent with GARP, but they generate severe violations when combined with their mirror images. Indeed, as Figure 4 illustrates, any decision to allocate *fewer* points to the cheaper account will generate a violation of GARP involving the mirror image of this decision. Hence, the CCEI score for the combined data provides a summary statistic of the overall consistency with GARP *and* first-order stochastic dominance.

<sup>7</sup> Put precisely, the data generated by an individual's choices are  $\{(p_1^i, p_2^i, x_1^i, x_2^i)\}_{i=1}^{25}$ , and the mirror-image data are  $\{(p_2^i, p_1^i, x_2^i, x_1^i)\}_{i=1}^{25}$ .

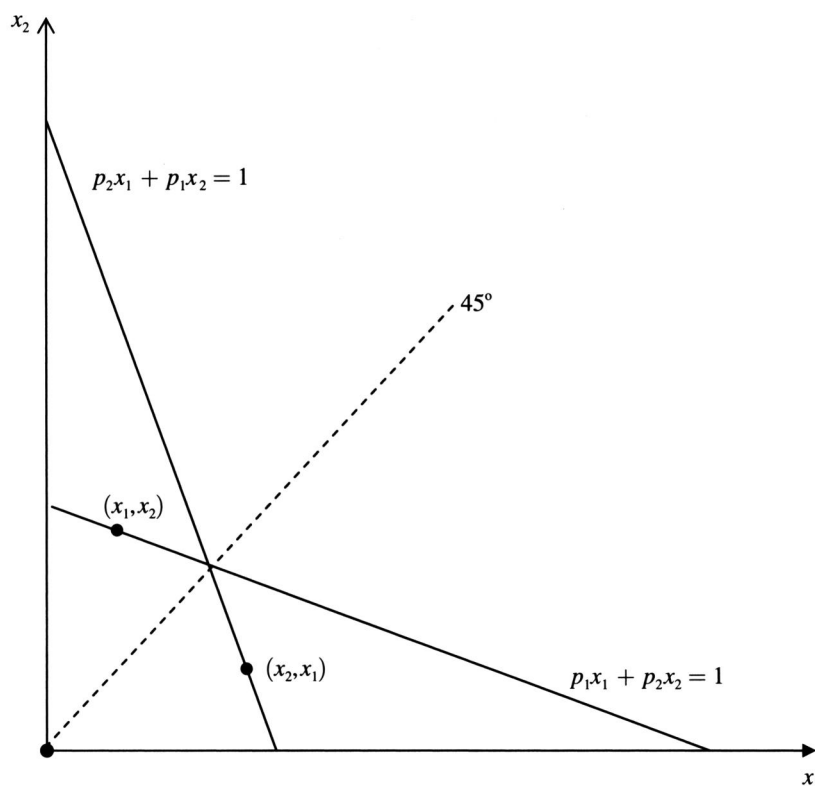


FIGURE 4. A VIOLATION OF GARP INVOLVING THE MIRROR-IMAGE BUDGET LINE

*Notes:* An individual choosing allocation  $(x_1, x_2)$  subject to budget constraint  $p_1x_1 + p_2x_2 = 1$  violates first-order stochastic dominance. This decision generates a violation of the Weak Axiom of Revealed Preference involving the mirror-image decision of choosing  $(x_1, x_2)$  subject to budget constraint  $p_2x_1 + p_1x_2 = 1$ .

On average, the CCEI score is 0.733 for the combined data compared to 0.881 for the actual data.<sup>8</sup> In our econometric analysis below, we use both the CCEI scores for the actual dataset and for the combined dataset. In online Appendix IV, we assess how closely individual choice behavior complies *only* with first-order stochastic dominance using an alternative measure based on expected payoff calculations. Replacing the CCEI score for the combined data with this measure of stochastic dominance in the econometric analysis below yields similar conclusions.

**Risk Attitudes.**—Our experimental task delivers measures of both decision-making quality and (risk) preferences from a single realm of decision-making. We summarize an individual's attitudes toward risk with a single statistic: the fraction of total points he allocated to the cheaper account. We choose this measure because in each problem that a subject faces, each account is equally likely to be chosen, and the budget line is drawn from a symmetric distribution. Thus, the only behavior consistent with infinite risk aversion is always allocating the points equally between

<sup>8</sup>Of the 1,182 subjects in the experiment, only 29 subjects (2.5 percent) almost always allocated all points to one of the accounts by choosing the same endpoint of the budget line.

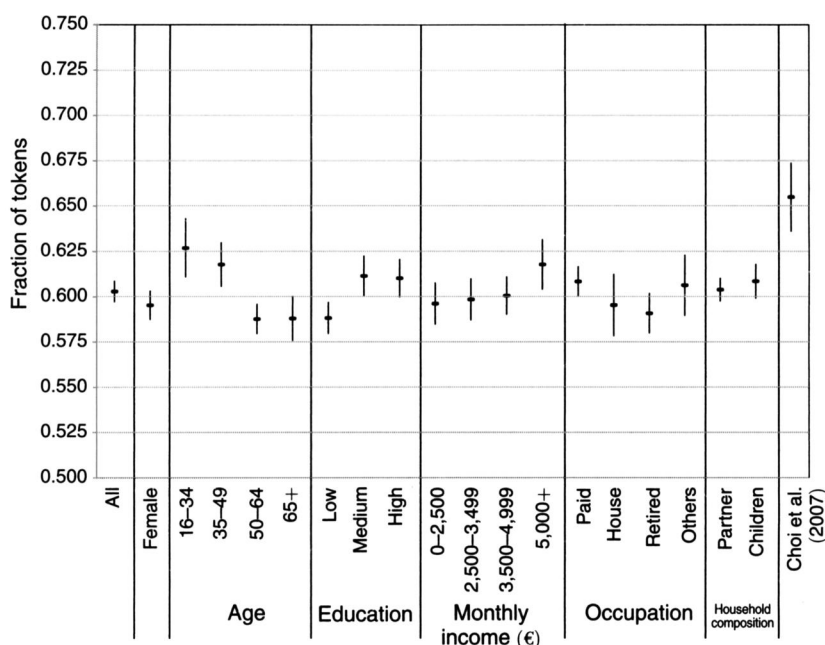


FIGURE 5. THE AVERAGE FRACTION OF TOKENS ALLOCATED TO THE CHEAPER ASSET

the two accounts. Conversely, always allocating all points to the cheaper account is the behavior that would be implied by risk neutrality. More generally, subjects who are less averse to risk will allocate a larger fraction of points to the cheaper account. Like the revealed preference tests, an advantage of this measure is that it is nonparametric. It measures attitudes toward risk without making assumptions about the parametric form of the underlying utility function.<sup>9</sup> Figure 5 displays the mean fraction of points allocated to the cheaper account and 95 percent confidence intervals across the socioeconomic categories. We 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 effectively uncorrelated ( $\rho = 0.113$ ).<sup>10,11</sup>

### III. Decision-Making Quality and Socioeconomics

We next perform an analysis of the correlation between decision-making quality—the consistency of the experimental data with GARP—and socioeconomic

<sup>9</sup>In parametric estimation, beyond the scope of this article, we find that the choice data of many subjects are well explained by a preference ordering in which the indifference curves have a kink at the 45-degree line, corresponding to an allocation with a certain payoff. One interpretation of this preference ordering is the disappointment aversion model proposed by Gul (1991). This finding corroborates the results in Choi et al. (2007) with undergraduate students.

<sup>10</sup>As Figure 5 shows, our individual-level measures of risk aversion are higher than the measures reported in Choi et al. (2007), but they are within the range of estimates from recent studies (see Choi et al. 2007 for a discussion of these studies).

<sup>11</sup>The seminal paper by Holt and Laury (2002) also reports substantial low-stakes risk aversion in the lab, as does Andersen et al. (2006) in the field. Risk aversion over moderate stakes contradicts the validity of Expected Utility over wealth (Rabin 2000, and Rabin and Thaler 2001). We emphasize that GARP does not imply the Savage (1954) axioms on which Expected Utility is based, and Expected Utility need not be assumed to investigate the decision-making quality of choice under uncertainty.

characteristics. Table 2 below presents the results of the econometric analysis. In column 1, we present estimates with the CCEI scores for the actual dataset using ordinary least squares.<sup>12</sup> The results show significant correlations. We obtain statistically significant coefficients in nearly all socioeconomic categories, ranging in absolute values from about 0.025 to just over 0.050. Most notably, female, low-education, low-income, and older 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 inconsistent choices. In column 2 we repeat the estimation reported in column 1 using the CCEI scores for the combined dataset. The two scores are highly correlated ( $\rho = 0.645$ ) and, unsurprisingly, the results are qualitatively similar.

The preceding analysis is based on the nonrandomly 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 evaluate sample selection bias in our econometric analysis using the Heckman (1979) method. Our exclusion restriction involves the number of completed CentERpanel questionnaires out of the total invitations to participate in the three months prior to our experiment. This variable enters the participation equation but, we assume, is conditionally uncorrelated with the CCEI (see Bellemare, Kröger, and van Soest 2008). To economize on space, the estimation results are reported in online Appendix V. The estimated parameters from the OLS and the sample selection estimations are virtually identical. We interpret these results to indicate that self-selection is not importantly driving the results.

#### IV. Wealth Differentials and Decision-Making Ability

The preceding analysis investigates the correlation between the consistency scores and the sociodemographic characteristics of subjects. This analysis makes no effort to evaluate a causal interpretation of these relationships. Nevertheless, the higher CCEI scores among high-income, high-education, and younger subjects suggest that these groups may have better economic outcomes, not only because they face fewer constraints or have more normative preferences, but also because they tend to have superior decision-making ability.

We next evaluate a causal interpretation of the correlation between the CCEI scores and important economic outcomes in the real world. If heterogeneity in decision-making ability is an important source of heterogeneity in economic outcomes, and if decision-making quality in the experiment as measured by the CCEI is a good proxy for decision-making ability, then differences in the CCEI scores across subjects should independently and robustly explain differences in their real-world outcomes.

We focus on household wealth as the real-world economic outcome of interest. As we argued, the investigation of wealth offers a strong test of the idea that decision-making quality in the experiment captures decision-making ability. Wealth is also of special interest because of a debate about the sources of its dispersion. Bernheim,

<sup>12</sup>To test for a potential misspecification, we used the Ramsey (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).

TABLE 2—THE CORRELATION BETWEEN CCEI SCORES AND SUBJECTS' INDIVIDUAL CHARACTERISTICS (OLS)

	(1)	(2)
Constant	0.887*** (0.022)	0.735*** (0.037)
<i>Female</i>	−0.024*** (0.009)	−0.011 (0.015)
<i>Age</i>		
35–49	−0.016 (0.011)	−0.007 (0.020)
50–64	−0.052*** (0.011)	−0.077*** (0.020)
65+	−0.051** (0.020)	−0.081** (0.032)
<i>Education</i>		
Medium	0.009 (0.011)	0.021 (0.017)
High	0.026** (0.011)	0.060*** (0.018)
<i>Income</i>		
€2,500–3,499	0.026** (0.012)	0.026 (0.019)
€3,500–4,999	0.020 (0.013)	0.006 (0.020)
€5,000+	0.033** (0.014)	0.017 (0.022)
<i>Occupation</i>		
Paid work	0.028 (0.018)	0.03 (0.026)
House work	0.047** (0.021)	0.039 (0.030)
Others	0.037* (0.019)	0.035 (0.030)
<i>Household composition</i>		
Partner	−0.026** (0.011)	−0.023 (0.018)
Number of children	0.001 (0.004)	0.001 (0.007)
$R^2$	0.068	0.058
Observations	1,182	1,182

*Notes:* Omitted categories: male, age under 35, low education (primary and lower secondary education), household gross monthly income under €2,500, retired, and not having a partner. Since the CCEI is a number between zero and one, we repeated the estimations reported in columns 1 and 2 using a fractional regression model (Papke and Wooldridge 1996). The two specifications yield similar results. Standard errors in parentheses.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

Skinner, and Weinberg (2001) and Ameriks, Caplin, and Leahy (2003) show substantial differences in wealth even among households with very similar lifetime incomes and provide evidence that differences in decision-making ability drive wealth differentials. Scholz, Seshadri, and Khitatrakun (2006) counter this by showing that, with detailed data on household-specific earnings, a standard life-cycle



model with homogeneous preferences can account for more than 80 percent of the cross-sectional variation in wealth.

Our analysis of wealth proceeds in four steps. We first establish the correlation between the CCEI and household wealth by estimating regressions of the log of household wealth on socioeconomic variables (including a flexible function of age), the log of household contemporaneous income, and the consistency score of the person who is primarily responsible for household financial matters. Importantly, we follow a well-established tradition in life-cycle analysis and treat income, education, and family structure as predetermined. We also evaluate the robustness of the conditional correlation between the CCEI and wealth to changes in the functional form of the estimating wealth equation.

Second, we demonstrate that this correlation is quantitatively robust to the inclusion of additional controls for unobserved constraints, preferences, and beliefs. If these unobservables were important sources of the observed correlation, then adding the controls should have a substantial effect on the estimated correlation between the CCEI and wealth. This analysis does not seek the impossible goal of “controlling for everything” that might influence wealth. Instead, in the spirit of Altonji, Elder, and Taber (2005), we examine whether the conditional correlation we see between CCEI and wealth in the basic specifications is much affected by the inclusion of additional controls that, *a priori*, should be correlated with wealth through their correlation with unobserved or misspecified variables.<sup>13</sup>

Third, we show that alternative measures of decision-making quality (from the experiment) and decision-making ability (from the CentERpanel survey) are not substitutes for the CCEI. The available alternatives either have no independent power to predict wealth, or are not well-correlated with consistency with GARP. Last, to account for the sources of the correlation and to better understand which real-world decisions cause those with higher CCEI scores to accumulate more wealth, we investigate the relationship between the CCEI and the details of household saving allocations.

#### *A. Establishing the Correlation between CCEI and Wealth*

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 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 household member. The inventory covers checking and saving accounts, stocks, bonds and other financial assets, real estate, business assets, mortgages, loans, and lines of credit. Our analysis focuses on nonpension household wealth, calculated by summing net worth over household members

<sup>13</sup>The reference to Altonji, Elder, and Taber (2005) underscores an analogy to the literature on education. Economists have long struggled to identify and interpret the various effects of education. The principal difficulty is that exogenous variation in education is rare and limited in scope. But the topic is so important that it justifies many approaches to inference and interpretation. We view our article in a similar vein, though, in the case of decision-making quality, both the empirical and the theoretical literatures are much less mature. Even the unconditional correlations presented above were previously unknown.

and taking the household's average over 2008 and 2009. The 703 households with wealth data and a CCEI score from the household's financial respondent had an average household wealth of €164,130. Percentile values (in thousands of euros) are provided below.<sup>14</sup>

Min	1	5	10	25	50	75	90	95	Max
-180.7	-68.2	-4.8	0.0	10.8	93.0	242.1	523.8	955.6	3,984.2

In our baseline specification, the sample size drops from 703 to 517 households (73.5 percent). This decline derives largely from three sources. First, 54 households (7.7 percent) have negative or missing household income in 2008, and 74 households (10.5 percent) have negative wealth and, thus, a missing dependent variable. Second, younger households face incentives to hold less wealth as they borrow in order to invest or to smooth lifetime consumption. With that in mind, we drop the 49 households (7.0 percent) whose financial respondent is less than 35 years old. Finally, to reduce the importance of extreme outliers, we drop the seven households that represent the top and bottom half of 1 percent of the wealth distribution and the bottom half of 1 percent of the CCEI distribution. Two additional households are dropped due to missing data on education. Our basic estimation results are reported in Table 3.

*Baseline.*—In column 1, we present estimates from our baseline specification using the sample of 517 households described above. The point estimate of 1.35 for the coefficient on the CCEI indicates that a standard deviation increase in the CCEI score of the household's financial respondent is associated with 18 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 = 0.017) with standard errors robust to heteroskedasticity.<sup>15</sup>

*Life Cycle.*—In column 2, we repeat the estimation reported in column 1 with the sample not restricted to households with financial respondents who are at least 35 years old. Using the entire analysis sample, we find that the point estimate on the CCEI is somewhat smaller, so a standard deviation increase in the CCEI score of the household's financial respondent is associated with about 15 percent more household wealth. The standard error on this point estimate implies that we can reject a null hypothesis of no relationship with considerable confidence ( $p$ -value = 0.038), but

<sup>14</sup>The CentERpanel data do not include information on pension wealth. Nearly all of the Dutch population is covered by the public pension system, whose benefits are a relatively simple function of family structure. A large majority of workers is also 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 earnings. Conditioning on family structure and earnings should, therefore, do much to control for the incentives these pensions create for nonpension wealth accumulation. See Alessie and Kapteyn (2001) and OECD (2009) for details about the pension systems in the Netherlands. While it is a necessity, studying nonpension wealth has the advantage of better isolating discretionary wealth accumulation.

<sup>15</sup>The individual coefficients on age, education, and occupation are economically large, but not statistically distinguishable from zero at standard levels of significance. A test of their joint significance, however, rejects a null hypothesis of no relationship with a  $p$ -value of 0.0002.

TABLE 3—THE RELATIONSHIP BETWEEN CCEI SCORES AND WEALTH

	(1)	(2)	(3)
CCEI	1.351** (0.566)	1.109** (0.534)	101,888.0* (52,691.9)
log 2008 household income	0.584*** (0.132)	0.606*** (0.126)	
2008 household income			1.776*** (0.4)
Female	−0.313* (0.177)	−0.356** (0.164)	−32,484.3* (17,523.9)
Partnered	0.652*** (0.181)	0.595*** (0.171)	46,201.9*** (17,173.7)
Number of children	0.090 (0.093)	0.109 (0.086)	14,078.6* (8,351.5)
Age	−0.303 (0.347)	−0.008 (0.208)	−19,148.5 (30,164.4)
Age <sup>2</sup>	0.007 (0.006)	0.002 (0.004)	468.7 (523.6)
Age <sup>3</sup>	0.000 (0.000)	0.000 (0.000)	−2.9 (2.9)
<i>Education</i>			
Prevocational	0.269 (0.464)	0.245 (0.462)	14,137.4 (43,449.1)
Preuniversity	0.634 (0.478)	0.562 (0.476)	59,035.0 (44,746.1)
Senior vocational training	0.416 (0.474)	0.421 (0.468)	28,318.7 (42,419.2)
Vocational college	0.490 (0.451)	0.527 (0.449)	31,341.2 (42,046.8)
University	0.725 (0.473)	0.685 (0.465)	77,578.8 (47,709.4)
<i>Occupation</i>			
Paid work	0.206 (0.322)	0.226 (0.321)	−12,657.2 (26,597.8)
House work	0.552 (0.406)	0.603 (0.413)	16,876.8 (31,114.3)
Retired	0.131 (0.318)	0.190 (0.318)	16,753.1 (35,165.2)
Constant	6.292 (6.419)	0.469 (3.598)	76,214.4 (559,677.5)
R <sup>2</sup>	0.179	0.217	0.188
Observations	517	566	568

Notes: The groupings of different levels of education are based on the categorization of Statistics Netherlands (Centraal Bureau voor de Statistiek). For a complete description see <http://www.centerdata.nl/en/centerpanel>. Standard errors in parentheses.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

we cannot reject a null hypothesis that the point estimates of the coefficient on the CCEI reported in columns 1 and 2 are the same.

*Levels.*—The log specification in columns 1 and 2 excludes households with negative wealth and may also cause small differences at positive but very low levels of wealth to have large effects on point 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) for the sample age 35 and older. We again see an economically large correlation between the CCEI and levels of wealth, though this relationship is estimated somewhat less precisely; the coefficient on the CCEI is significant only at the 10 percent level ( $p$ -value = 0.054).

### B. Evaluating a Causal Interpretation

We find an economically large and statistically significant correlation between the financial respondent's CCEI score in the experiment and household wealth. This lends a basic level of support to the idea that our measure of decision-making quality from the experiment can proxy for decision-making ability that applies across multiple real world choice domains. The correlation between the CCEI and wealth may, however, not reflect decision-making ability, but instead be due to a correlation between the CCEI and other standard, but so far unobserved or misspecified, sources of heterogeneity in choice that affect wealth. To evaluate a causal relationship, we study the robustness of the correlation with respect to the inclusion of additional controls for unobserved constraints, preferences, and beliefs. Tables 4 and 5 present the estimates of the coefficients of interest. The full-length tables are contained in online Appendix VI.

*Constraints.*—We begin by investigating the correlation of the CCEI with unobserved or misspecified constraints that affect the accumulation of wealth. The estimation results are reported in Table 4. In standard life-cycle models, wealth at a given age is a function of the constraints imposed by the path of income over a lifetime. There is a variety of reasonable specifications for a wealth regression, and our baseline specification reported in Table 3 above is an especially simple benchmark. In column 1 of Table 4, we assess the importance of the baseline linear-in-contemporaneous-income specification by allowing income to enter in the form of a cubic. We see virtually no change in the point estimate of the coefficient on the CCEI. We thus find no evidence that the simple specification of contemporaneous income drives the estimated relationship between the CCEI and wealth.

Another concern is that contemporaneous income is measured with error, and the estimated coefficient on this variable is therefore biased toward zero. The bias of this estimate then biases estimates of the coefficients on other variables, including the CCEI. Standard life-cycle models predict constant saving rates across lifetime income groups, and thus a unit elasticity of wealth with respect to income (Dynan, Skinner, and Zeldes 2004). If contemporaneous income is a good proxy for lifetime income, then these theories predict the coefficient on the log of contemporaneous income should equal one in the baseline specification. In column 2 of Table 4, we impose this restriction and see virtually no change in the point estimate of the coefficient on the CCEI.<sup>16</sup> There is, thus, no evidence that measurement error in contemporaneous income drives the main result.

<sup>16</sup>Brown (1976) shows that if theory makes a prediction about the coefficient on a variable that is measured with error then restricting that coefficient to take on the value predicted by theory will reduce the bias on the other coefficients being estimated.

TABLE 4—THE ROBUSTNESS OF THE CORRELATION BETWEEN CCEI SCORES AND WEALTH TO THE INCLUSION OF CONTROLS FOR UNOBSERVED CONSTRAINTS

	(1)	(2)	(3)	(4)	(5)
CCEI	1.322** (0.570)	1.318** (0.574)	1.925*** (0.672)	1.888*** (0.652)	1.441** (0.578)
log household income					
2008	19.770 (14.629)	1.000	0.544*** (0.137)	0.285* (0.165)	0.616*** (0.128)
2008 <sup>2</sup>	-2.194 (1.533)				
2008 <sup>3</sup>	0.082 (0.053)				
2006				0.232 (0.231)	
2004				0.215 (0.174)	
Demographics	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	No
Occupation	Yes	Yes	Yes	Yes	Yes
Constant	-47.059 (46.275)	0.864 (6.545)	5.354 (6.93)	3.016 (7.109)	6.398 (6.484)
R <sup>2</sup>	0.187		0.205	0.217	0.177
Observations	517	517	449	449	517

Notes: Standard errors in parentheses.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

A related concern is that contemporaneous income is a poor proxy for the *path* of lifetime income, and unobserved aspects of that path are correlated with the CCEI. We can evaluate this concern with some of the limited panel data available on household income. 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.<sup>17</sup> In column 3 of Table 4, we repeat the baseline specification reported in column 1 of Table 3, this time restricting attention to the 449 households (86.8 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 different from zero ( $p$ -value = 0.004). In column 4 of Table 4, we add controls for the log of household income in 2004 and 2006. As a result, the magnitude of the coefficient on the CCEI declines only slightly (by 0.037). We interpret this to indicate that, while some of the correlation between the CCEI and wealth may be attributable to a correlation between the CCEI and unobserved past income, the available CentERpanel data on income provide little evidence that this is the case.

<sup>17</sup> 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. In cases where we have two out of the three income measures, we use linear extrapolation to fill in the third.

An alternative approach is to take completed education as a proxy for lifetime income. All of our specifications so far include indicators for each level of education completed. It may be, however, that unobserved aspects of education, such as the quality of schooling, are correlated with unobserved elements of lifetime income which are, in turn, correlated with the CCEI score in the experiment. If so, and if these unobserved constraints are important sources of the observed correlation between CCEI and wealth, then conditioning on completed education should have a substantial effect on the estimated coefficient on the CCEI. In column 5 of Table 4, we repeat the baseline specification reported in column 1 of Table 3 after omitting the controls for the education of the financial respondent. Comparing the estimates from these two specifications, we see that removing the education controls increases the estimated coefficient on the CCEI only modestly (by 0.090). In this way, we find little evidence that unobserved aspects of education are driving the correlation between the CCEI and wealth.

*Preferences.*—We have evaluated whether the correlation between the CCEI and wealth is due to a relationship between the CCEI and unobserved or misspecified income constraints. The results on education, which we took to proxy for unobserved income but could also proxy for attitudes toward risk and time, suggest that the relationship between the CCEI and unobserved preferences is unlikely to play an important role. Nevertheless, we next turn to analyze the possibility that the correlation between the CCEI and wealth is driven by a relationship between the CCEI and unobserved preferences that determine wealth. The estimation results are reported in Table 5 below.

Our experimental task delivers individual-level measures of both decision-making quality and risk preferences from a single realm of decision-making. In column 1 of Table 5 we add to the baseline specification reported in column 1 of Table 3 a control for risk attitudes by including a nonparametric measure from the experiment discussed above—the average fraction of points the financial respondent allocated to the cheaper account.<sup>18</sup> The point estimate on this *quantitative* measure indicates that risk tolerance in the experiment is negatively associated with wealth. The coefficient is economically large—a standard deviation increase in the fraction placed in the cheaper account is associated with about 7 percent less household wealth—but imprecisely estimated. We cannot reject a null hypothesis of no relationship ( $p$ -value = 0.282) or a null of an economically large and positive relationship. Given that risk attitudes and CCEI scores in the experiment are effectively uncorrelated ( $\rho = 0.113$ ), it is unsurprising that the inclusion of this control leaves the point estimate of the coefficient on the CCEI little changed.

It may be, however, that risk attitudes that influence wealth are not well correlated with risk preferences revealed over the small stakes of the experiment. If so, *qualitative* measures of risk tolerance taken from the CentERpanel survey instead of the experiment may do better. In column 2, we evaluate this possibility by also

<sup>18</sup>To avoid the influence of extreme outliers, and to place this variable on more equal footing with the CCEI, we also estimated a specification that omitted the top and bottom half of 1 percent of the distribution of this risk attitude measure, a total of six households. The results are qualitatively similar.

TABLE 5—THE ROBUSTNESS OF THE CORRELATION BETWEEN CCEI SCORES AND WEALTH TO THE INCLUSION OF CONTROLS FOR UNOBSERVED PREFERENCES AND BELIEFS

	(1)	(2)	(3)	(4)	(5)
CCEI	1.379** (0.568)	1.396** (0.568)	1.404** (0.569)	1.214* (0.625)	1.237** (0.623)
Risk tolerance					
Quantitative (experiment)	−0.768 (0.714)	−0.808 (0.711)	−0.766 (0.718)		
Qualitative (survey)		0.017 (0.074)	0.023 (0.076)		
Qualitative (survey) missing		−0.190 (0.335)	−0.162 (0.482)		
Conscientiousness			0.089 (0.072)		
Conscientiousness missing			−0.040 (0.668)		
Longevity expectations					−0.034 (0.040)
log 2008 household income	0.589*** (0.132)	0.578*** (0.131)	0.572*** (0.133)	0.443*** (0.123)	0.434*** (0.123)
Demographics	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes	Yes
Constant	6.840 (6.361)	6.883 (6.357)	6.496 (6.395)	3.777 (15.258)	4.411 (15.256)
R <sup>2</sup>	0.179	0.176	0.176	0.163	0.163
Observations	517	517	517	414	414

Notes: Risk aversion in the experiment is measured by the average fraction of tokens allocated to the cheaper asset. Standard errors in parentheses.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

including a normalized measure of risk-taking in investments.<sup>19</sup> To preserve sample size, we also include a variable to indicate whether the respondent provided a complete answer to these questions. The results reinforce those from the previous specification. The point estimate of the coefficient on the qualitative risk tolerance measure is economically small but imprecisely estimated. As important, the inclusion of a qualitative measure of risk attitudes leaves the estimated coefficient on the CCEI virtually unchanged. We thus find no evidence that these qualitative measures of risk attitudes from the survey are better able to capture unobserved preferences, correlated with the CCEI, that influence wealth.

In a final effort to evaluate the importance of a correlation between the CCEI and unobserved preferences, in column 3 we add to the list of controls a conscientiousness

<sup>19</sup>The CentERpanel survey contains six statements related to investment risk and return such as “I think it is more important to have safe investments and guaranteed returns, than to take a risk to have a chance to get the highest possible returns,” and “I want to be certain my investments are safe.” Respondents are asked to evaluate the accuracy of these statements as descriptions of themselves on a seven point scale. We sum the responses for each respondent, when necessary reordering them so that higher scores reflect greater risk tolerance. We then normalize the scores to have sample mean 0 and standard deviation 1.

measure from the “Big Five” test used for personality research in psychology.<sup>20</sup> Again, to preserve sample size, we also include a variable to indicate whether the respondent completed the conscientiousness questions. If not, we set his score to the sample mean (zero). The magnitude of the coefficient on conscientiousness is large. A standard-deviation increase in conscientiousness is associated with about 9 percent more wealth. The standard error on the conscientiousness coefficient is also relatively large, however, and we cannot reject a null hypothesis of no correlation. Most important, adding the control for conscientiousness has almost no effect on the coefficient on the CCEI. Thus, we again find no evidence that the relationship between the CCEI and wealth is driven by a correlation between the CCEI and unobserved preferences that influence wealth.

*Beliefs.*—Standard life-cycle models predict that beliefs, such as expectations for longevity, income, or asset returns, should affect household wealth levels. The CentERpanel collects relatively little information about respondents’ beliefs, but the survey does ask questions about expected longevity. We can therefore use these data to evaluate the extent to which the correlation between the CCEI and wealth accumulation is attributable to a correlation between the CCEI and some unobserved beliefs that influence wealth. To strike a balance between capturing more information and maintaining the sample sizes, we consider a measure of longevity expectations based on the question answered by the largest number of respondents.<sup>21</sup>

In column 4 of Table 5, we repeat the baseline specification reported in column 1 of Table 3, this time restricting attention to the 414 households (80.0 percent) for whom the financial respondent answered this question about longevity. In this smaller sample, the point estimate of the coefficient on the CCEI remains economically large, though a larger standard error reduces the statistical significance ( $p$ -value = 0.053). In column 5 of Table 5, we add the control for longevity expectations. This measure, itself, has little power to predict wealth levels, and including it increases the estimate of the coefficient on the CCEI very slightly (by 0.023). Thus, while we have limited ability to explore this issue with available data, we find no evidence that a relationship between the CCEI and unobserved beliefs drives the correlation between the CCEI and wealth.

<sup>20</sup>The other Big Five personality traits are openness, extraversion, agreeableness, and neuroticism. Among the Big Five, conscientiousness has the strongest correlation with economic success (see, for examples, Barrick and Mount 1991, and Tett, Jackson, and Rothstein 1991). Conscientious people are described as “thorough, careful, reliable, organized, industrious, and self-controlled” (Duckworth et al. 2007). These terms suggest more patience, less risk tolerance, and less taste for leisure. Conscientiousness may thus proxy for unobserved preferences that influence wealth. The CentERpanel survey contains ten statements related to conscientiousness. The statements include: “I do chores right away,” “I am accurate in my work,” among others. Respondents are asked to evaluate the accuracy of these statements as descriptions of themselves on a five-point scale. For each respondent, we sum his or her responses to the ten statements, and then normalize the scores to have sample mean 0 and standard deviation 1. When necessary, we reordered the responses so that higher scores reflect greater conscientiousness. Simultaneously adding other measures from the Big Five yields the same conclusion. We omit those results for the sake of brevity.

<sup>21</sup>The question answered by the largest number of respondents asks “How likely is it that you will attain (at least) the age of 80?” Responses are recorded on a scale from 0 to 10, and respondents are instructed to interpret 0 to mean “no chance at all” and 10 to mean “absolutely certain.”



TABLE 6—EVALUATING ALTERNATIVE MEASURES OF DECISION-MAKING QUALITY

	(1)	(2)	(3)	(4)
CCEI	1.253* (0.712)	1.401* (0.729)	1.269* (0.729)	1.177** (0.583)
CCEI (combined dataset)	0.099 (0.380)			
von Gaudecker et al. (2011)			0.927* (0.485)	
Cognitive Reflection Test				0.120* (0.071)
Cognitive Reflection Test missing				−0.203 (0.237)
log 2008 household income	0.586*** (0.132)	0.388* (0.155)	0.383* (0.154)	0.577*** (0.132)
Demographics	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes
Constant	6.237 (6.424)	10.056 (6.976)	8.355 (6.990)	6.855 (6.464)
R <sup>2</sup>	0.177	0.225	0.232	0.181
Observations	517	326	326	517

Notes: The CCEI scores for the combined dataset are computed after combining the actual data from the experiment and the mirror-image data. Standard errors in parentheses.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

### C. Evaluating Alternatives to the CCEI

We have found an economically large, statistically significant, and quantitatively robust relationship between the CCEI scores in the experiment and wealth. We next evaluate whether alternative laboratory- and survey-based measures can substitute for the CCEI for the purposes of explaining wealth. Measuring decision-making quality by consistency with GARP has strong theoretical and methodological justifications, but augmenting GARP with additional, normative criteria might better capture decision-making quality. It is also possible that, while lacking in theoretical foundations, other proxies for decision-making ability are so well correlated with the CCEI that they may serve as substitutes for the CCEI. This would be especially useful if the other proxies are readily available on surveys or in administrative datasets. The estimation results are reported in Table 6. The full-length table is contained in online Appendix VI.

*Stochastic Dominance.*—We begin with consideration of a stronger notion of decision-making quality derived from our experiment. In column 1 of Table 6 we repeat the estimation of the baseline specification reported in column 1 of Table 3 after adding the CCEI scores for the combined dataset (combining the actual data from the experiment and the mirror-image data). As explained above, this test of decision-making quality is stronger because it demands both consistency with GARP and first-order stochastic dominance. We find no evidence that, conditional on the CCEI score from the actual data, the CCEI score for the combined dataset has an independent relationship with wealth. Adding the CCEI for the combined dataset

as a regressor has only a modest effect on the point estimate of the coefficient on the CCEI, though the standard error on this estimate increases. The point estimate of the coefficient on the CCEI for the combined dataset is small but imprecisely estimated. These results are consistent with the idea that the CCEI for the combined dataset, while requiring a compelling and generally accepted notion of decision-making quality, merely represents a noisier measure of the aspects of decision-making ability captured by the CCEI scores for the actual dataset.

*Trembling.*—von Gaudecker, van Soest, and Wengström (2011) conducted risk experiments with CentERpanel members using a multiple price list design (Andersen et al. 2006). They estimated a flexible parametric model that includes an individual “trembling” parameter  $\omega_i$  measuring “the propensity to choose randomly rather than on the basis of preferences.” von Gaudecker, van Soest, and Wengström (2011) conclude that “while many people exhibit consistent choice patterns, some have very high error propensities.” That parameter can be interpreted as a measure of decision-making quality as it captures the degree to which an individual’s choices are consistent both with rationality and with some assumptions about the functional form of utility. This contrasts with the CCEI, which makes no assumptions about the structure of preferences. The CCEI and the “trembling” parameter of von Gaudecker, van Soest, and Wengström (2011) are only moderately correlated ( $\rho = 0.178$ ) in the overlapping sample of 624 subjects (43.9 percent) who participated in both experiments. As a result, we can gain some insight into the relationship between wealth and consistency with the utility-maximizing model *versus* consistency with a class of utility functions commonly employed in the empirical studies.

In column 2 of Table 6, we repeat the estimation of the baseline specification reported in column 1 of Table 3, this time restricting attention to the 326 households (63.1 percent) with a financial respondent who participated in both experiments.<sup>22</sup> In column 3 of Table 6 we add to the list of regressors a variable equal to  $(1 - \omega_i)$  in von Gaudecker, van Soest, and Wengström (2011). The point estimate of the coefficient on this parameter is large—a standard-deviation increase is associated with 17 percent more household wealth—and we can reject a null hypothesis of no relationship at the 10 percent confidence level ( $p$ -value = 0.057). The estimated coefficient on the CCEI is reduced somewhat when we include this estimate but remains economically large and statistically significant at the 10 percent confidence level ( $p$ -value = 0.083). There are many possible reasons for the differences in these two measures and their relationship to wealth—they are based on different methods, derived from different designs, and elicited in different experiments. Intriguingly, however, the results suggest substantial differences between decision-making quality measured by the restrictions imposed by the utility-maximization model and additional restrictions imposed by various hypotheses concerning functional structure. This is an interesting avenue for future work with more data on both measures.

*Cognitive Ability.*—Tests of cognitive ability (IQ) might also capture aspects of decision-making ability (cf. Dohmen et al. 2010). Different from IQ tests,

<sup>22</sup>To place the two measures on more equal footing, we trim the lowest one-half of 1 percent of the distribution of each of the measures (two observations each) in the overlapping sample.

consistency with GARP offers a theoretically disciplined measure for decision-making quality that has a well-established economic interpretation. There is no comparable, theoretically disciplined means of using and interpreting an IQ test. Nevertheless, if the CCEI and IQ were very well correlated, then analysts interested in measuring decision-making ability might be able to replace the revealed preference tests with one of the many IQ tests and, in some circumstances, the conceptual distinctions between the measures would have little practical import. It seems likely that the capacity to make choices of high decision-making quality draws on skills of analysis and perception that also improve IQ test scores. But if the goal is to isolate the influence of these skills on decision-making ability, rather than on constraints, information, or beliefs, are the CCEI and IQ scores substitutes?

Many IQ tests are precluded from wide use by intellectual property rights or are impossible to implement in Internet panels. The CentERpanel has not implemented any of the well-known and wide-ranging IQ instruments. However, in connection with our and other researchers' projects, the CentERpanel asked a sample of respondents to complete the Frederick (2005) Cognitive Reflection Test and a brief Raven's Progressive Matrices Test.<sup>23</sup> We omit the Raven's test because it generated effectively no variation in responses. Among the 467 subjects who completed the Cognitive Reflection Test and participated in our experiment, the correlation between the CCEI and the Cognitive Reflection Test is positive, but the variables are far from collinear ( $\rho = 0.193$ ). This result echoes Burks et al. (2009) who find a correlation of approximately 0.22 between IQ and compliance with monotonicity (measured by more than one switch point in multiple price list experiments regarding risk and time trade-offs). To assess the predictive content of the CCEI and the Cognitive Reflection Test, in column 4 of Table 6 we add the number of questions answered correctly in the Cognitive Reflection Test to the baseline specification. To preserve sample size, we also include a variable to indicate whether a Cognitive Reflection Test score was available for the household's financial respondent. For those who had no score, we substitute the mean of the distribution of the rest of the sample.

The point estimate of the coefficient on the Cognitive Reflection Test is economically large—answering one more of the three questions on the Cognitive Reflection Test correctly is associated with about 12 percent more wealth—and statistically significant at the 10 percent level ( $p$ -value = 0.094). Adding this measure, and the missing indicator, to the basic specification reduces the estimated coefficient on the CCEI somewhat, but about a fifth of the reduction is due to the inclusion of the indicator for a missing Cognitive Reflection Test score. These results suggest that the Cognitive Reflection Test captures some decision-making ability related both to decision-making quality in the experiment and wealth accumulation. The findings also indicate, however, that this measure of cognitive ability cannot be used as a simple substitute for the CCEI for the purposes of explaining wealth.<sup>24</sup>

<sup>23</sup>The Cognitive Reflection Test consists of three questions. Each question is designed to have an intuitive, but incorrect, answer. The intuitive answer tends to spring to mind and then require reflection in order to dismiss. One question asks "A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?" Here the intuitive answer is \$0.10, but the correct answer is \$0.05. Frederick (2005) shows that this test is well correlated with a 50-question test of general cognitive ability, as well as tests of achievement such as the Scholastic Assessment Test (SAT).

<sup>24</sup>Perhaps related to Cognitive Reflection Test and the tendency to reflect upon choices, adding a control for subjects' response times in the experiment has virtually no effect on the coefficient on the CCEI. This is expected

#### D. The Sources of Correlation between CCEI and Wealth

We found an economically large and statistically significant correlation between household wealth and the financial respondent's CCEI score in the experiment. We saw that this correlation is robust to the inclusion of controls for unobserved preferences, constraints, and beliefs, and that alternative measures of decision-making quality or decision-making ability elicited in an experiment or a survey are not substitutes for the CCEI. With these results in mind, we now turn to account for the sources of the correlation by investigating the relationship between the CCEI and the details of household saving allocations. Our goal is to better understand which important real-world decisions cause those who appear to have better decision-making ability to accumulate more wealth. These estimation results are reported in Table 7. The full-length table is contained in online Appendix VI.

*Portfolio.*—In columns 1–6 of Table 7, we present estimates that relate the CCEI score of the household's financial respondent to 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. To reduce the importance of extreme outliers, in all specification 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 results provide some evidence that, conditional on household characteristics, contemporaneous income, occupation, and education level, 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 2 and 4 of Table 7 are modest in magnitude but statistically significant at the 10 percent level ( $p$ -values of 0.083 and 0.095, 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 distinguishable from zero.

*Housing.*—Finally, the coefficients on the CCEI in columns 7 and 8 of Table 7 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 0.047 in the probability of owning a home. Similarly, a standard deviation increase in the CCEI is associated with an increase of 0.043 in the fraction of wealth held in housing. 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 gives home ownership an important advantage over renting and, other things equal, means wealth placed in

---

given the modest unconditional correlation between the CCEI scores and response times ( $\rho = -0.075$ ). We also cannot reject a null hypothesis of no correlation between the response time and wealth. These results are omitted in the interest of brevity.

TABLE 7—THE SOURCES OF THE RELATIONSHIP BETWEEN HOUSEHOLDS' NET WORTH AND CCEI SCORES

	(1)	(2)	(3)	(4)
	Have checking	Fraction in checking	Have saving	Fraction in saving
CCEI	0.03 (0.032)	−0.098* (0.057)	−0.047 (0.053)	−0.162* (0.097)
log 2008 household income	0.001 (0.002)	−0.029** (0.013)	0.003 (0.010)	−0.068*** (0.021)
Demographics	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes
Constant	0.998*** (0.172)	0.106 (0.822)	1.126 (0.848)	1.448 (1.288)
$R^2$	−0.007	0.021	−0.011	0.083
Observations	512	512	502	502
	(5)	(6)	(7)	(8)
	Have stocks	Fraction in stocks	Have a house	Fraction in house
CCEI	0.167 (0.163)	0.001 (0.050)	0.352** (0.152)	0.324** (0.129)
log 2008 household income	0.148*** (0.031)	0.013 (0.009)	0.134*** (0.029)	0.096*** (0.024)
Demographics	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes
Constant	−3.152* (1.856)	−0.317 (0.398)	−1.047 (1.760)	−1.151 (1.419)
$R^2$	0.079	0.002	0.148	0.123
Observations	514	514	479	479

Notes: Standard errors in parentheses.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

mortgaged housing pays a substantial premium.<sup>25</sup> So long as housing supply is somewhat elastic, and thus the incidence of these tax benefits is shared between buyers and sellers, this suggests that owning a home and placing more wealth in mortgaged housing are often high decision-making quality financial choices. 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 (financial) decision-making ability.

<sup>25</sup> About 69 percent of households in the sample own a home, and the average fraction of wealth held in housing is approximately 54 percent. In the Netherlands, assets held in owner-occupied housing are not subject to the usual capital income tax. If they were, 4 percent of housing value would be treated as implicit income and taxed at 30 percent. Instead, imputed rent is presumed to be very low (0.55 percent of housing value), is subject to the progressive tax on labor income, and that tax is not due unless the household claims a deduction for mortgage interest. Nominal mortgage interest is, in turn, fully deductible 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, Jacobs, and de Mooij (2007), the Netherlands offers by far the most favorable tax treatment of owner occupied housing in Western Europe.

## V. Related Literature

This paper is primarily concerned with a relatively new economics literature that emphasizes a gap between what people actually choose and what they would have chosen if they were fully attentive to their choices and had the skills and knowledge necessary to weigh costs and benefits in often complex settings. This research provides evidence that some choices are better than others and some individuals are better decision-makers than others.

This new literature contrasts standard economic analysis which takes a libertarian approach; in the absence of the complete data required, standard analysis assumes that all choices are of good decision-making quality and that everyone has sufficient decision-making ability. The libertarian approach has obvious appeal, and economists in the new literature have struggled to measure decision-making quality and to separately identify decision-making ability from standard sources of heterogeneity in important economic outcomes. That is, the new studies have been subject to identification and measurement problems and confronted them to different degrees.

Observational studies of market data have shown that some individuals make choices of such poor decision-making quality that they clearly leave “money on the table.” Agarwal et al. (2009) is a prominent example. That paper provides evidence of strictly dominated choices regarding the use of credit. Importantly, research like this requires an uncommonly high-quality (administrative) dataset that makes the relevant trade-offs very clear. The comprehensive nature of the data makes classifying some decisions as “mistakes” immediate and indisputable, and thus allows the analyst to avoid both the measurement and the identification problem. However, such administrative datasets provide information about the activity at just one retailer, or on just a few credit cards, or in just one form of saving. They thus capture only a slice of the economic activities of the individuals involved.<sup>26</sup> As a result, they cannot reveal whether a lower decision-making quality choice represents a minor bobble in otherwise sound decision-making, or a more fundamental problem in evaluating economic choices due to lack of decision-making ability.

Experiment- or survey-based studies take a different approach. They confront the identification problem using measures of education, cognitive and noncognitive skills, or financial literacy as proxies for decision-making ability and either instrumental variables or experimental manipulation to infer causal effects. A prominent example of this research is Ameriks, Caplin, and Leahy (2003). That paper provides evidence that differences in individuals’ planning abilities, rather than more standard sources of heterogeneity, explain important variation in wealth. Other examples of significant research in this vein include Bernheim and Garrett (2003) and Duflo and Saez (2003), who study the effect of employer-based financial education on saving. Research of this kind relaxes the imposing data requirements of observational studies that investigate decision-making quality directly, but it is

<sup>26</sup>Other examples of strictly observational studies of decision-making quality that focus on “slices” of economic activity include Miravete (2003) on cell phone calling plans, Choi, Laibson, and Madrian (2011) on employer-sponsored saving, Ketcham et al. (2012) on health insurance, and Lacetera, Pope, and Sydnor (2012) on car purchases.

silent on the measurement problem. It does not seek a portable, quantitative, and economically interpretable measure of decision-making quality.

Other related research relies importantly on surveys and is more strictly descriptive. Restricting attention just to financial decision-making, this research includes Lusardi and Mitchell (2007) who document very low levels of financial planning, financial literacy, and a positive correlation between literacy, financial planning and wealth; and Fang, Keane, and Silverman (2008), who find that, rather than measures of risk preferences, cognitive functioning stands out as a significant predictor of Medigap purchase. Agarwal and Mazumder (2013) correlate proxies of decision-making ability, namely mathematical and nonmathematical cognitive tests, with making low quality financial decisions involving the use of credit.<sup>27</sup>

Our approach follows the strictly observational studies by defining decision-making quality and measuring it directly. Different from that literature, we define decision-making quality by the consistency of choices with GARP, and we measure it by the extent of GARP violation using the CCEI and other standard indices that have been proposed for this purpose. Like the experimental literature, we seek to isolate decision-making ability from standard sources of heterogeneity in choice, in our case by presenting individuals with a theoretically grounded economic choice experiment designed to provide a stringent test of consistency with GARP.

Echenique, Lee, and Shum (2011) measure consistency with revealed preference conditions in individual-level data on grocery store expenditures and correlate the degree of consistency with some indicators of family size, income, age, and education. Their measure is based on the idea that an individual who violates GARP can be exploited as a “money pump.” We refer the interested reader to online Appendix II for the construction of the money pump index. To our knowledge, no other papers analyze the correlation between consistency with GARP and socioeconomic characteristics in a broad population. Echenique, Lee, and Shum (2011) conclude that “the hypothesis of consumer rationality cannot be rejected.” Such consumption data can, however, lack power to reject violations of GARP (Blundell, Browning, and Crawford 2003, 2008). The power of the test depends on the range of prices consumers face (the frequency with which budget lines cross) and the number of choices each consumer makes.

## VI. Concluding Remarks

We offered a new large-scale experiment—employing graphical representations of standard consumer decision problems and using a diverse pool of subjects—that enables us to collect richer data than has been possible in the past. These data allow us to say that some sets of choices are of better decision-making quality than others, in that some choices are more rational than others. Because the data are provided by

<sup>27</sup> Special attention has been paid in this literature to older populations, reflecting a concern about decreasing decision-making ability later in life. The papers in the November 2010 issue of the *Economic Journal* (Volume 120, Issue 548) offer nice reviews: Banks (2010) summarizes research on the relationships between cognitive function, financial literacy, and financial outcomes at older ages; Smith, McArdle, and Willis (2010) and Banks, O’Dea, and Oldfield (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 causes of the differences in cognitive function and financial literacy in later life.

a large and heterogeneous sample, we can analyze the correlates of decision-making quality in the laboratory and relate it to important economic outcomes like wealth. We find that differences in the experimental measures of decision-making quality across subjects independently and robustly explain differential patterns of wealth across households. Since wealth accumulation is determined by countless decisions, made over time in many environments, and involving a host of different trade-offs, our findings indicate that our measure of decision-making quality captures aspects of decision-making ability that apply across many sorts of economic choice problems. This study therefore suggests a new path toward solving the twin problems of identification and measurement.

We conclude by underscoring three key features of the approach we have taken here. First, our measure of decision-making quality—the consistency of choices with GARP—dictates the experimental task: a canonical problem of selecting an allocation from a budget line. Given the task, the measure also provides the benchmark level of consistency necessary to provide a rigorous test. There are no comparable, theoretically disciplined means of quantifying, interpreting, and evaluating laboratory or survey measures of decision-making ability, namely tests of cognitive and noncognitive skills. Second, informed by economic theory, the single experimental task delivers measures of both decision-making quality and preferences from a unified realm of decision-making. Relevant preferences cannot be recovered from performance on standard psychological tests. Third, unlike many tests of cognitive abilities or personality traits, revealed preference tests are applicable to, and comparable across, all sorts of economic choice problems. Our approach can thus be transported, with relative ease, to different decision domains. We can make domain-specific predictions and provide a unified measure of decision-making quality across domains. In all of these ways, the theoretical foundation of our approach drives the design of the experiment and allows diverse and disciplined use of the resulting data.

Taken together, our findings provide new evidence on three important issues: (i) the validity of measuring decision-making quality by the consistency of choices with economic rationality; (ii) the feasibility of testing for consistency with rationality through a web-based survey on a large scale; (iii) the relative importance of heterogeneity in decision-making ability for understanding important economic outcomes. This last issue commands special attention because decision-making ability, unlike preferences, may be justifiably manipulated. If differences in decision-making ability are important sources of the heterogeneity in economic outcomes, then even quite costly policy changes aimed at “soft” or “libertarian” paternalism may hold substantial promise.<sup>28</sup>

## REFERENCES

- Afriat, Sidney N. 1967. “The Construction of Utility Functions from Expenditure Data.” *International Economic Review* 8 (1): 67–77.  
Afriat, Sidney N. 1972. “Efficiency Estimation of Production Function.” *International Economic Review* 13 (3): 568–98.

<sup>28</sup> The prominent forms of light or soft paternalism are libertarian paternalism (Thaler and Sunstein 2003) and asymmetric paternalism (Camerer et al. 2003). Loewenstein and Haisley (2008) and Kariv and Silverman (2013) provide relevant discussions.



- Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson. 2009. "The Age of Reason: Financial Decisions over the Life Cycle and Implications for Regulation." *Brookings Papers on Economic Activity* (2): 51–101.
- Agarwal, Sumit, and Bhashkar Mazumder. 2013. "Cognitive Abilities and Household Financial Decision Making." *American Economic Journal: Applied Economics* 5 (1): 193–207.
- Alessie, Rob, and Arie Kapteyn. 2001. "Savings and Pensions in the Netherlands." *Research in Economics* 55 (1): 61–82.
- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–84.
- Ameriks, John, Andrew Caplin, and John Leahy. 2003. "Wealth Accumulation and the Propensity to Plan." *Quarterly Journal of Economics* 118 (3): 1007–47.
- Andersen, Steffen, Glenn W. Harrison, Morten Igel Lau, and E. Elisabet Rutström. 2006. "Elicitation Using Multiple Price List Formats." *Experimental Economics* 9 (4): 383–405.
- Banks, James. 2010. "Cognitive Function, Financial Literacy and Financial Outcomes at Older Ages: Introduction." *Economic Journal* 120 (548): F357–62.
- Banks, James, Cormac O'Dea, and Zoe Oldfield. 2010. "Cognitive Function, Numeracy and Retirement Saving Trajectories." *Economic Journal* 120 (548): F381–410.
- Barrick, Murray R., and Michael K. Mount. 1991. "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis." *Personnel Psychology* 44 (1): 1–26.
- Bellemare, Charles, Sabine Kröger, and Arthur van Soest. 2008. "Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities." *Econometrica* 76 (4): 815–39.
- Bernheim, B. Douglas, and Daniel M. Garrett. 2003. "The Effects of Financial Education in the Workplace: Evidence from a Survey of Households." *Journal of Public Economics* 87 (7–8): 1487–519.
- Bernheim, B. Douglas, and Antonio Rangel. 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *Quarterly Journal of Economics* 124 (1): 51–104.
- Bernheim, B. Douglas, Jonathan Skinner, and Steven Weinberg. 2001. "What Accounts for the Variation in Retirement Wealth among US Households?" *American Economic Review* 91 (4): 832–57.
- Blundell, Richard, Martin Browning, and Ian A. Crawford. 2003. "Nonparametric Engel Curves and Revealed Preference." *Econometrica* 71 (1): 205–40.
- Blundell, Richard, Martin Browning, and Ian Crawford. 2008. "Best Nonparametric Bounds on Demand Responses." *Econometrica* 76 (6): 1227–62.
- Brown, Charles. 1976. "A Model of Optimal Human-Capital Accumulation and the Wages of Young High School Graduates." *Journal of Political Economy* 84 (2): 299–316.
- Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini. 2009. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences* 106 (19): 7745–50.
- Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin. 2003. "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism.'" *University of Pennsylvania Law Review* 151 (3): 1211–54.
- Choi, James J., David Laibson, and Brigitte C. Madrian. 2011. "\$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans." *Review of Economics and Statistics* 93 (3): 748–63.
- Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv. 2007. "Consistency and Heterogeneity of Individual Behavior under Uncertainty." *American Economic Review* 97 (5): 1921–38.
- Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman. 2014. "Who Is (More) Rational?: Dataset." *American Economic Review*. <http://dx.doi.org/10.1257/aer.104.6.1518>.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review* 100 (3): 1238–60.
- Duckworth, Angela L., Christopher Peterson, Michael D. Matthews, and Dennis R. Kelly. 2007. "Grit: Perseverance and Passion for Long-Term Goals." *Journal of Personality and Social Psychology* 92 (6): 1087–1101.
- Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 118 (3): 815–42.
- Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes. 2004. "Do the Rich Save More?" *Journal of Political Economy* 112 (2): 397–444.
- Echenique, Federico, Sangmok Lee, and Matthew Shum. 2011. "The Money Pump as a Measure of Revealed Preference Violations." *Journal of Political Economy* 119 (6): 1201–23.
- Fang, Hanming, Michael P. Keane, and Dan Silverman. 2008. "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy* 116 (2): 303–50.

- Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual Preferences for Giving." *American Economic Review* 97 (5): 1858–76.
- Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19 (4): 25–42.
- Gul, Faruk. 1991. "A Theory of Disappointment Aversion." *Econometrica* 59 (3): 667–86.
- Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1): 153–61.
- Holt, Charles A., and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92 (5): 1644–55.
- Jappelli, Tullio. 2010. "Economic Literacy: An International Comparison." *Economic Journal* 120 (548): F429–51.
- Kahneman, Daniel, and Amos Tversky. 1984. "Choices, Values, and Frames." *American Psychologist* 39 (4): 341–50.
- Kariv, Shachar, and Dan Silverman. 2013. "An Old Measure of Decision-Making Quality Sheds New Light on Paternalism." *Journal of Institutional and Theoretical Economics* 169 (1): 29–44.
- Ketcham, Jonathan D., Claudio Lucarelli, Eugenio J. Miravete, and M. Christopher Roebuck. 2012. "Sinking, Swimming, or Learning to Swim in Medicare Part D." *American Economic Review* 102 (6): 2639–73.
- Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121 (4): 1133–65.
- Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor. 2012. "Heuristic Thinking and Limited Attention in the Car Market." *American Economic Review* 102 (5): 2206–36.
- Loewenstein, George, and Emily Haisley. 2008. "The Economist as Therapist: Methodological Ramifications of "Light" Paternalism." In *The Foundations of Positive and Normative Economics*, edited by Andrew Caplin and Andrew Schotter, 210–48. New York: Oxford University Press.
- Lusardi, Annamaria, and Olivia S. Mitchell. 2007. "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics* 54 (1): 205–24.
- Miravete, Eugenio J. 2003. "Choosing the Wrong Calling Plan? Ignorance and Learning." *American Economic Review* 93 (1): 297–310.
- Organisation for Economic Co-operation and Development (OECD). 2009. *Pensions at a Glance 2009: Retirement-Income Systems in OECD Countries*. Paris: OECD.
- Papke, Leslie E., and Jeffrey M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics* 11 (6): 619–32.
- Quiggin, John. 1990. "Stochastic Dominance in Regret Theory." *Review of Economic Studies* 57 (3): 503–11.
- Rabin, Matthew. 2000. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica* 68 (5): 1281–92.
- Rabin, Matthew, and Richard H. Thaler. 2001. "Anomalies: Risk Aversion." *Journal of Economic Perspectives* 15 (1): 219–32.
- Ramsey, J. B. 1969. "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis." *Journal of the Royal Statistical Society* 31 (2): 350–71.
- Samuelson, Paul A. 1947. *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.
- Savage, Leonard J. 1954. *The Foundations of Statistics*. New York: Wiley.
- Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun. 2006. "Are Americans Saving 'Optimally' for Retirement?" *Journal of Political Economy* 114 (4): 607–43.
- Smith, James P., John J. McArdle, and Robert Willis. 2010. "Financial Decision Making and Cognition in a Family Context." *Economic Journal* 120 (548): F363–80.
- Tett, Robert P., Douglas N. Jackson, and Mitchell Rothstein. 1991. "Personality Measures as Predictors of Job Performance: A Meta-Analytic Review." *Personnel Psychology* 44 (4): 703–42.
- Thaler, Richard H., and Cass R. Sunstein. 2003. "Libertarian Paternalism." *American Economic Review* 93 (2): 175–79.
- van den Berg, Gerard J., Dorly J. H. Deeg, Maarten Lindeboom, and France Portrait. 2010. "The Role of Early-Life Conditions in the Cognitive Decline Due to Adverse Events Later in Life." *Economic Journal* 120 (548): F411–28.
- van Ewijk, Casper, Bas Jacobs, and Ruud De Mooij. 2007. "Welfare Effects of Fiscal Subsidies on Home Ownership in the Netherlands." *De Economist* 155 (3): 323–36.
- von Gaudecker, Hans-Martin, Arthur van Soest, and Erik Wengström. 2011. "Heterogeneity in Risky Choice Behavior in a Broad Population." *American Economic Review* 101 (2): 664–94.
- Wakker, Peter. 1993. "Savage's Axioms Usually Imply Violation of Strict Stochastic Dominance." *Review of Economic Studies* 60 (2): 487–93.