Consistency, Heterogeneity, and Granularity of Individual Behavior under Uncertainty*

Syngjoo Choi[†]
UCL

Raymond Fisman[‡]
Columbia University

Douglas Gale[§]
NYU

Shachar Kariv[¶]
UC Berkeley

January 17, 2007

Abstract

By using graphical representations of budget sets over bundles of state-contingent commodities, we generate a very rich data set well-

*Some of the results reported here were previously distributed in a paper titled "Substantive and Procedural Rationality in Decisions under Uncertainty." This research was supported by the Experimental Social Science Laboratory (X-Lab) at UC Berkeley. We are grateful to David Ahn, Jim Andreoni, Dan Ariely, Colin Camerer, Andrew Caplin, Liran Einav, Paul Glimcher, Daniel Kahneman, Tom Palfrey, Drazen Prelec, Matthew Rabin, Ariel Rubinstein, Aldo Rustichini, Andrew Schotter, Dan Silverman, and Bill Zame for helpful discussions. We also thank to three anonymous referees for their comments. This paper has also benefited from suggestions by the participants of seminars at Berkeley, Duke, Georgetown, LSE, NYU, Princeton, Pittsburgh, Rutgers, Stanford, Yale, UCL, the Communication and Incentives Workshop at UCSB, and the Conference on Econometrics and Experimental Economics at Northwestern. We would also like to thank Brenda Naputi and Lawrence Sweet from the X-Lab for their valuable assistance, and Roi Zemmer for writing the computer programs. We acknowledge the Center on the Economics and Demography of Aging (CEDA) at UC Berkeley for financial support. Kariv is grateful to the hospitality of the Institute for Advances Studies School of Social Science.

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

[‡]Graduate School of Business, Columbia University, Uris 823, New York, NY 10027 (E-mail: rf250@columbia.edu, URL: http://www-1.gsb.columbia.edu/faculty/rfisman/).

[§]Department of Economics, New York University, 269 Mercer Street, #507, New York, NY, 10003 (E-mail: douglas.gale@nyu.edu, URL: http://www.econ.nyu.edu/user/galed).

[¶]Department of Economics, University of California, Berkeley, Evans Hall # 3880, Berkeley, CA 94720 (E-mail: kariv@berkeley.edu, URL: http://socrates.berkeley.edu/~kariv/).

suited to studying behavior under uncertainty at the level of the individual subject. We test the data for consistency with the maximization hypothesis, and we estimate preferences using a two-parameter utility function based on Faruk Gul (1991). This specification provides a good interpretation to the data at the level of the individual subject and can account for the highly heterogeneous behaviors observed in the laboratory. The parameter estimates jointly describe attitudes toward risk and allow us to characterize the distribution of risk preferences in the population.

JEL Classification Numbers: D81, C91.

Key Words: uncertainty, revealed preference, Expected Utility Theory, loss/disappointment aversion, experiment.

We report the results of a series of experiments using an innovative graphical interface that we apply to studying decision making under uncertainty. In our experimental design, subjects see a graphical representation of a standard budget constraint on a computer screen. This can be interpreted either as a portfolio choice problem (the allocation of wealth between two risky assets) or a consumer decision problem (the selection of a bundle of contingent commodities subject to a standard budget constraint). Subjects use the mouse to choose a portfolio by pointing-and-clicking on the budget line. This intuitive and user-friendly interface allows for the quick and efficient elicitation of many decisions per subject from a wide variety of budget constraints. The result is a rich individual-level data set that is at the core of this paper's contribution.

The benefits of our approach are immediately evident from inspecting the scatter diagrams of data from individual subjects, which reveal many distinct patterns in subjects' behaviors. Sometimes they choose safe portfolios that guarantee the same return in each state of nature. Sometimes they bet everything on one state, as if they were risk neutral. Sometimes they respond smoothly to changes in the risk-return tradeoff. The behavior of subjects is generally complex and impossible to classify in a simple taxonomy. Nonetheless, in almost every case, one or more of these distinct patterns can be clearly distinguished. We call this patterning "granularity," for want of a better term. (In computer science, granularity is a measure of the size or number of components in a system, with a course-grained system having relatively few, large components. Here, granularity refers to the number of distinct aspects or features displayed by individual behavior.)

Although individual behavior is granular (made up of distinct behavior patterns), the second striking fact is the high level of consistency in the individual level decisions. That is, most subjects behave as if they were maximizing a complete, transitive preference ordering over lotteries (portfolios). A well-known theorem of Sidney N. Afriat (1967) states that an individual's choices from a finite number of budget sets are consistent with maximization of a well behaved utility function if and only if they satisfy the Generalized Axiom of Revealed Preference (GARP). The broad range of budget sets that our experiment involves provides a rigorous test of GARP. In particular, the changes in endowments and relative prices are such that budget lines cross frequently. This means that our data lead to high power tests of revealed preference conditions. Our subjects attain very high scores on standard measures of consistency and most are close to the ideal of perfectly rational behavior.

The consistency of individual decisions naturally leads us to ask what kind of preferences are consistent with the observed choices. Our third discovery is that the data is well explained by a preference ordering in which the indifference curves have a kink at the 45 degree line. One interpretation of this preference ordering is that it displays loss or disappointment aversion (Eddie Dekel, 1986; Gul, 1991). Expected Utility Theory (EUT) is a special case of this theory. The family of utility functions we estimate is characterized by two parameters, one of which measures loss or disappointment aversion. Over half of our subjects have a significant degree of loss or disappointment aversion. The remainder appear to be well approximated by preferences consistent with EUT (John von Neumann and Oskar Morgenstern, 1947; Leonard J. Savage, 1954).

Because preferences are characterized by two parameters, we cannot easily summarize attitudes toward risk by a single number. However, we can compute a risk premium based on the difference between the expected value of a gamble and its certainty equivalent. Comparing the risk premium to a standard measure of risk aversion suggests that our estimates are within the range found by other researchers (cf. Kay-Yen Chen and Charles R. Plott, 1998; Charles A. Holt and Susan K. Laury, 2002; Jacob K. Goeree, Holt, and Thomas R. Palfrey, 2003, 2004; Goeree and Holt, 2004).

The rest of the paper is organized as follows. Section 1 provides a discussion of closely related literature. Section 2 describes the experimental design and procedures. Section 3 illustrates some important features of the data and establishes the consistency of the data with utility maximization. Section 4 provides the econometric analysis, and Section 5 concludes. Experimental instructions, technical details, and individual-level data are gathered in appendices.

1 Related literature

The experimental literature on choice under uncertainty is vast and cannot be summarized here. Colin F. Camerer (1995) provides a comprehensive discussion of the experimental and theoretical work, and Chris Starmer (2000) provides a more recent review that focuses on evaluating non-EUT theories. The typical experimental design presents subjects with a number of binary choices. The objective is to test the empirical validity of particular axioms or to compare the predictive abilities of competing theories. These theories tend to be systematically disconfirmed by the data. This has motivated researchers to develop more descriptive models, and the investigation of these models has led to the discovery of new empirical regularities in the laboratory.

Typically, the criterion used to evaluate a theory is the fraction of choices it predicts correctly. A theory is "rejected" when the pattern of violations appears to be systematic. More recently, following the seminal work of John D. Hey and Chris Orme (1994) and David W. Harless and Camerer (1994), a number of papers compare models while allowing for randomness. In these studies, randomness can be interpreted as the effect of a trembling hand, calculation error, and so forth. While Harless and Camerer (1994) fit models to aggregate data, Hey and Orme (1994) use data derived from decisions over a very large menu of binary choices and estimate functional forms for individual subjects. They test EUT as a restriction on non-EUT theories and find that EUT appears to fit as well as non-EUT alternatives for almost 40 percent of their subjects and that violations of EUT decay with repetition.

A few other studies, such as Imran S. Currim and Rakesh K. Sarin (1989), Richard L. Daniels and L. Robin Keller (1990), and Pamela K. Lattimore, Joanna R. Baker and A. Dryden Witte (1992) have also estimated parametric utility functions for individual subjects. These studies find that many subjects obey EUT, with considerable variation in risk aversion across subjects. Our paper – both in its experimental method and theoretical apparatus – substantially extends this research program by providing new techniques and larger samples that enable more precise estimation and better predictions. Camerer (1995) emphasizes the need for such improvements in advancing the research program in this area.

The distinctive features of the present paper are the new experimental design and the application of tools from consumer demand theory to individual decision making in the laboratory. This experimental design generates data that are better suited in a number of ways to estimating risk preferences. First, the choice of a portfolio from a convex budget set provides more information about preferences than a discrete choice.[1] Second, the large amount of individual-level data generated by this design allows us to apply statistical models to individual data rather than pooling data or assuming homogeneity across subjects. Hence, we may generate better individual-level estimates of risk aversion. Third, these decision problems are representative, both in the statistical sense and in the economic sense, rather than, as in existing methods, being designed to test a particular theory.

Syngjoo Choi, Raymond Fisman, Douglas M. Gale, and Shachar Kariv (forthcoming) illustrate the application of the experimental platform that has been developed in this paper facilitates the analysis of decisions under uncertainty at the individual level. Their analysis builds on revealed preference techniques to determine whether the choices of hypothetical subjects are consistent with utility maximization and to recover their underlying preferences. The experimental technique can also be applied to many types of individual choice problems. For example, Fisman, Kariv, and Daniel Markovits (2006) employ a similar experimental methodology to study social preferences. While the papers share a similar experimental methodology, they address very different questions and produce very different behaviors.

2 Experimental design and procedures

2.1 Design

In the experimental task we study, individuals make decisions under conditions of uncertainty about the objective parameters of the environment. In our preferred interpretation, there are two *states of nature* denoted by $s = 1, 2$ and two associated *Arrow securities*, each of which promises a payoff of one unit of account in one state and nothing in the other. We consider the problem of allocating an individual's wealth between the two Arrow securities. Let x_s denote the demand for the security that pays off in state s and let p_s denote its price. We normalize the individual's wealth to 1. The budget constraint is then $p_1x_1 + p_2x_2 = 1$ and the individual can choose any portfolio $(x_1, x_2) \geq 0$ that satisfies this constraint.

[1] In Loomes (1991) subjects also allocate wealth in a portfolio of risky assets. The focus of this paper is on providing a test of the independence axiom, so the results are not directly comparable to those presented here. Loomes (1991) showed that most subjects made nearly rational choices but systematically violated the intendance axiom, and that the observed behavior cannot be accommodated by a number of non-EUT alternatives.

An example of a budget constraint defined in this way is the straight line AB drawn in Figure 1. The axes measure the future value of a possible portfolio in each of the two states. The point C, which lies on the 45 degree line, corresponds to a portfolio with a certain payoff. By contrast, point A (point B) represents a portfolio in which all wealth is invested in the security that pays off in state 1 (state 2). A portfolio such as C is called a *safe* portfolio and portfolios such as A and B are called *boundary* portfolios. A portfolio that is neither a safe nor a boundary portfolio is called an *intermediate* portfolio. Notice that, given the objective probabilities of each state, positions on AB do not represent *fair* bets (portfolios with the same expected value as the a safe portfolio). If π is the probability of state 1 and the slope of the budget line $-p_1/p_2$ is steeper than $-\pi/(1-\pi)$, positions along AC have a higher payoff in state 1, a lower payoff in state 2, and a lower expected portfolio return than point C.

[Figure 1 here]

2.2 Procedures

The experiment was conducted at the Experimental Social Science Laboratory (X-Lab) at UC Berkeley under the X-Lab Master Human Subjects Protocol. The 93 subjects in the experiment were recruited from undergraduate classes and staff at UC Berkeley. After subjects read the instructions (reproduced in Appendix I), the instructions were read aloud by an experimenter. Each experimental session lasted about one and a half hours. Payoffs were calculated in terms of tokens and then converted into dollars. Each token was worth \$0.5. A \$5 participation fee and subsequent earnings, which averaged about \$19, were paid in private at the end of the session.

Each session consisted of 50 independent decision rounds. In each round, a subject was asked to allocate tokens between two accounts, labeled x and y. The x account corresponds to the x-axis and the y account corresponds to the y-axis in a two-dimensional graph. Each choice involved choosing a point on a budget line of possible token allocations. Each round started by having the computer select a budget line randomly from the set of lines that intersect at least one axis at or above the 50 token level and intersect both axes at or below the 100 token level. The budget lines selected for each subject in his decision problems were independent of each other and of the budget lines selected for other subjects in their decision problems.

The x-axis and y-axis were scaled from 0 to 100 tokens. The resolution compatibility of the budget lines was 0.2 tokens. At the beginning of each decision round, the experimental program dialog window went blank and the entire setup reappeared. The appearance and behavior of the pointer were set to the Windows mouse default and the pointer was automatically repositioned randomly on the budget line at the beginning of each round. To choose an allocation, subjects used the mouse or the arrows on the keyboard to move the pointer on the computer screen to the desired allocation. Subjects could either left-click or press the Enter key to record their allocations. No subject reported difficulty understanding the procedures or using the computer interface. (The computer program dialog window is shown in the experimental instructions which are reproduced in Appendix I.)

At the end of the round, the computer randomly selected one of the accounts, x or y. Each subject received the number of tokens allocated to the account that was chosen. We studied a *symmetric* treatment (subjects ID 201-219 and 301-328), in which the two accounts were equally likely ($\pi = 1/2$) and two *asymmetric* treatments (subjects ID 401-417, 501-520 and 601-609) in which one of the accounts was selected with probability 1/3 and the other account was selected with probability 2/3 ($\pi = 1/3$ or $\pi = 2/3$). The treatment was held constant throughout a given experimental session. Subjects were not informed of the account that was actually selected at the end of each round. At the end of the experiment, the computer selected one decision round for each participant, where each round had an equal probability of being chosen, and the subject was paid the amount he had earned in that round.

3 From data to preferences

3.1 Data description

We begin with an overview of some important features of the experimental data. We will focus on the symmetric treatment, where the regularities in the data are very clear, and select a small number of subjects who illustrate salient features of the data. One must remember, however, that for most subjects the data are much less regular. Figure 2 depicts, for each subject, the relationship between the log-price ratio $\ln(p_1/p_2)$ and the token share $x_1/(x_1+x_2)$. The figures for the full set of subjects are available in Appendix II, which also shows the portfolio choices (x_1, x_2) as points in a scatterplot, and the relationship between the log-price ratio $\ln(p_1/p_2)$ and the budget share p_1x_1 (prices are normalized by income so that $p_1x_1 + p_2x_2 = 1$). Clearly, the distinction between token share and budget share is only relevant in the presence of price changes.

[Figure 2 here]

Figure 2A depicts the choices of a subject (ID 304) who always chose nearly safe portfolios $x_1 = x_2$. This behavior is consistent with infinite risk aversion. Figure 2B shows the choices of the only subject (ID 303) who, with a few exceptions, made nearly equal expenditures $p_1x_1 = p_2x_2$. This behavior is consistent with a logarithmic von Neumann-Morgenstern utility function. This is a very special case, where the regularity in the data is very clear. We also find many cases of subjects who implemented "smooth" responsiveness of portfolio allocations to prices, albeit less precisely. Among these subjects, we find considerable heterogeneity in price sensitivity. Perhaps most interestingly, no subject in the symmetric treatment allocated all the tokens to x_1 if $p_1 < p_2$ and to x_2 if $p_1 > p_2$. This is the behavior that would be implied by pure risk neutrality, for example. Nevertheless, boundary portfolios $(x_1, 0)$ and $(0, x_2)$ were used in combination with other portfolios by many subjects, as we will see below.[2]

Another interesting regularity is illustrated in Figure 2C, which depicts the decisions of a subject (ID 307) who allocated all of his tokens to x_1 (x_2) for values of $\ln(p_1/p_2)$ that give a flat (steep) budget line. This aspect of his behavior would be consistent with risk neutrality. However, for a variety of intermediate prices corresponding to $\ln(p_1/p_2)$ around zero, this subject chose nearly safe portfolios $x_1 = x_2$. This aspect of his choice behavior is consistent with infinite risk aversion. So this subject is apparently switching between behaviors that are individually consistent with EUT, but mutually inconsistent. In fact, as we will see in the econometric analysis below, this subject's preferences exhibit loss or disappointment aversion (where the safe portfolio $x_1 = x_2$ is taken to be the reference point).

There are yet more fine-grained cases where the behavior is less stark, such as the subject (ID 216) whose choices are depicted in Figure 2D. This subject combines intermediate portfolios for a variety of intermediate relative prices with boundary portfolios for prices that give sufficiently flat or steep budget lines. Further, the subject (ID 318) whose choices are depicted in Figure 2E combines safe, intermediate and boundary portfolios. There is something distinctly discontinuous in the behavior of these subjects and their choices are clearly not consistent with the standard interpretation of

[2] A single subject (ID 508) almost always chose $x_1 = 0$ if $p_1 > p_2$ and $x_2 = 0$ otherwise. However, he participated in the asymmetric treatment ($\pi = 2/3$) and thus his choices do not correspond to risk neutrality. Three subjects (ID 205, 218 and 320) chose a minimum level of consumption of ten tokens in each state, and allocated the residual to the less expensive security.

EUT.

These are of course special cases, where the regularities in the data are very clear. There are many subjects for whom the behavioral rule is much less clear and there is no taxonomy that allows us to classify all subjects unambiguously. But even in cases that are harder to classify, we can pick out the safe, intermediate, and boundary portfolios described above. Overall, a review of the full data set reveals striking regularities *within* and marked heterogeneity *across* subjects.

3.2 Testing rationality

Before proceeding to a parametric analysis of the data, we want to check whether the observed data are consistent with any preference ordering, EU or non-EU. To answer this question, we need to make use of some results from the theory of revealed preference. A well-known result, due to Afriat (1967), tells us that a *finite* data set generated by an individual's choices can be rationalized by a well-behaved (piecewise linear, continuous, increasing and concave) utility function, if and only if the data satisfies the Generalized Axiom of Revealed Preference (GARP).[3] GARP requires that if a portfolio x is revealed preferred to x' then x' is not strictly revealed preferred to x. So, in order to show that the data are consistent with utility-maximizing behavior, we can simply check whether it satisfies GARP (simple in theory, though difficult in practice for moderately large data sets).

Since GARP offers an exact test (either the data satisfy GARP or they do not) and choice data almost always contain at least some violations, we also wish to measure the *extent* of GARP violations. We report measures of GARP violations based on an index proposed by Afriat (1972). Afriat's *critical cost efficiency index* (CCEI) measures the amount by which each budget constraint must be adjusted in order to remove all violations of GARP. Figure 3 illustrates one such adjustment for a simple violation of GARP involving two portfolios, x^1 and x^2.[4] It is clear that x^1 is revealed preferred to x^2 because x^2 is cheaper than x^1 at the prices at which x^1 is purchased, and x^2 is revealed preferred to x^1, since x^1 is cheaper than x^2 at the prices at which x^2 is purchased. If we shifted the budget constraint through x^2 as shown, the violation would be removed. In this case, the CCEI would equal A/B ($A/B > C/D$).

[3] This statement of the result follows Hal R. Varian (1982), who replaced the condition Afriat called *cyclical consistency* with GARP.

[4] In fact, here we have a violation of the Weak Axiom of Revealed Preference (WARP). Note that if choices violate WARP they need not violate GARP.

[Figure 3 here]

By definition, the CCEI is a number between 0 and 1, where a value of 1 indicates that the data satisfy GARP perfectly. There is no natural threshold for determining whether subjects are close enough to satisfying GARP that they can be considered utility maximizers. Varian (1991) suggests a threshold of 0.95 for the CCEI, but this is purely subjective. A more scientific approach, proposed by Stephen G. Bronars (1987), calibrates the various indices using a hypothetical subject whose choices are uniformly distributed on the budget line. We generated a random sample of 25,000 subjects and found that their scores on the Afriat CCEI indices averaged 0.60.[5] Furthermore, all 25,000 random subjects violated GARP at least once, and none had a CCEI score above Varian's 0.95 threshold. If we choose the 0.9 efficiency level as our critical value, we find that only 12 of the random subjects had CCEI scores above this threshold.

Figure 4 compares the distributions of the CCEI scores generated by the sample of 25,000 hypothetical subjects (gray) and the distributions of the scores for the actual subjects (black).[6] The horizontal axis shows the value of the index and the vertical axis measures the percentage of subjects corresponding to each interval. The histograms clearly show that a significant majority of the subjects did much better than the randomly generated subjects and only a bit worse than an ideal (rational) subject. Our experiment is thus sufficiently powerful to exclude the possibility that consistency is the accidental result of random behavior. As a practical note, the consistency results presented above suggest that subjects did not have any difficulties in understanding the procedures or using the computer program.

[Figure 4 here]

The power of the experiment is very sensitive to the number of observations for each subject. To illustrate this point, we simulated the choices of random subjects in two experiments which used the design of this paper except that in one, subjects made 10 choices and in the other, they made 25 choices. In each case, the simulation was based on 25,000 random subjects.

[5] Each of the 25,000 random subjects makes 50 choices from randomly generated budget sets, in the same way that the human subjects do.

[6] To allow for small trembles resulting from the slight imprecision of subjects' handling of the mouse, all the results presented below allow for a narrow confidence interval of one token (for any i and $j \neq i$, if $\left|x^i, x^j\right| \leq 1$ then x^i and x^j are treated as the same portfolio). We generate virtually identical results allowing for a narrower confidence interval.

In the simulated experiment with 25 choices, 4.3 percent of random subjects were perfectly consistent, 14.3 percent had CCEI scores above Varian's 0.95 threshold, and 28.9 percent had values above 0.90. In the simulated experiment with only 10 choices, the corresponding percentages were 20.2, 37.3, and 50.6. In other words, there is a very high probability that random behavior will pass the GARP test if the number of individual decisions is as low as it usually has been in earlier experiments. We refer the interested reader to Choi, Fisman, Gale and Kariv (forthcoming) for further details on the power of tests for consistency with GARP.

Appendix III lists, by subject, the number of violations of WARP and GARP, and also reports the values of the three indices according to descending CCEI scores. Although it provides a summary statistic of the overall consistency of the data with GARP, the CCEI does not give any information about which of the observations are causing the most severe violations. We refer the interested reader to Appendix III for precise details on testing for consistency with GARP and other indices that have been proposed for this purpose by Varian (1991) and Martijn Houtman and J. A. H. Maks (1985). The various indices are all computationally intensive for even moderately large data sets. (The computer program and details of the algorithms are available from the authors upon request.)

4 Econometric analysis

4.1 Specification

The near consistency of subjects' choices tells us that there exists a well-behaved utility function that rationalizes *most* of the data. Additionally, because of the nature of the data, particularly the clustering at the safe and boundary portfolios, EUT cannot provide a plausible fit for the data at the individual level. The particular patterns observed in the data lead us to consider the theory of loss/disappointment aversion proposed by Gul (1991), which implies that the utility function over portfolios (x_1, x_2) takes the form

$$\min\left\{\alpha u\left(x_1\right)+u\left(x_2\right), u\left(x_1\right)+\alpha u\left(x_2\right)\right\},$$

where $\alpha \geq 1$ is a parameter measuring loss/disappointment aversion (where the safe portfolio $x_1 = x_2$ is taken to be the reference point) and $u(\cdot)$ is the utility of consumption in each state. If $\alpha > 1$ there is a kink and if $\alpha = 1$ we have the standard EUT representation. This formulation thus embeds EUT as a parsimonious and tractable special case and allows for the estimation of the parameter values in our empirical analysis below.

To implement this approach, we assume that $u(\cdot)$ takes the power form commonly employed in the analysis of choice under uncertainty,

$$u(x) = \frac{x^{1-\rho}}{(1-\rho)},$$

where ρ is the Arrow-Pratt measure of relative risk aversion. The parameters in this two-parameter specification, α and ρ, jointly describe the attitudes toward risk and allow us to characterize the distribution of risk preferences in the population.

The use of the power function has one limitation, however, in that the function is not well defined for the boundary portfolios. We incorporate the boundary observations $(1/p_1, 0)$ or $(0, 1/p_2)$ into our estimation using strictly positive portfolios where the zero component is replaced by a small consumption level such that the demand ratio x_1/x_2 is either $1/\omega$ or ω, respectively. The minimum ratio is chosen to be $\omega = 10^{-3}$. The selected level did not substantially affect the estimated coefficients for any subject.

With this adjustment, maximizing the utility function subject to the budget constraint yields a non-linear relationship between $\ln(p_1/p_2)$ and $\ln(x_1/x_2)$, which is illustrated in Figure 5 below. If the security prices are very different, then the optimum is the boundary portfolio with the larger expected payoff. If the security prices are very similar (log-price ratios are close to zero), then the optimum is the safe portfolio. In these cases, the optimal choice is insensitive to small price changes. For log-price ratios that are neither extreme nor close to zero, the optimum is an intermediate portfolio and the choice is sensitive to small changes in the risk-return tradeoff.

[Figure 5 here]

The subject's demand will belong to one of five possible cases: (*i*) a corner solution in which $x_1 = \omega\bar{x}_2$ if $x_1/x_2 < \omega$; (*ii*) an interior solution where $\omega \leq x_1/x_2 < 1$; (*iii*) a corner solution where $x_2 = \omega\bar{x}_1$ if $1/\omega < x_1/x_2$; (*iv*) an interior solution where $1 < x_1/x_2 \leq 1/\omega$; and (*v*) a solution at the kink where $x_1/x_2 = 1$.[7] The two interior solutions are characterized by first-order conditions in the form of equations; the two corner solutions and the kink are characterized by inequalities. Combining these cases, we can define

[7]Intuitively, these conditions set the ratio of demands x_1/x_2 equal to ω or $1/\omega$ when observations are near to the boundary.

an individual-level econometric specification for each subject n separately, and generate estimates of $\hat{\alpha}_n$ and $\hat{\rho}_n$ using nonlinear least squares (NLLS).

The data generated by an individual's choices are $\left\{\left(\bar{x}_1^i, \bar{x}_2^i, x_1^i, x_2^i\right)\right\}_{i=1}^{50}$, where $\left(x_1^i, x_2^i\right)$ are the coordinates of the choice made by the subject and $\left(\bar{x}_1^i, \bar{x}_2^i\right)$ are the endpoints of the budget line, (so we can calculate the relative prices $p_1^i/p_2^i = \bar{x}_2^i/\bar{x}_1^i$ for each observation i). Next, we identify the five different cases discussed above (corner solutions, interior solutions, kink). The first-order conditions at the optimal choice $\left(x_1^{i*}, x_2^{i*}\right)$, given $\left(\bar{x}_1^i, \bar{x}_2^i\right)$, can thus be written as follows (here we have taken logs of the first-order conditions and then replaced prices with the observed values):

$$\ln\left(\frac{x_1^{i*}}{x_2^{i*}}\right) = f\left[\ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right); \alpha, \rho, \omega\right] = \begin{cases} \ln\omega & \text{if } \ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right) \geq \ln\alpha - \rho\ln\omega, \\ -\frac{1}{\rho}\left[\ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right) - \ln\alpha\right] & \text{if } \ln\alpha < \ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right) < \ln\alpha - \rho\ln\omega, \\ 0 & \text{if } -\ln\alpha \leq \ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right) \leq \ln\alpha, \\ -\frac{1}{\rho}\left[\ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right) + \ln\alpha\right] & \text{if } -\ln\alpha + \rho\ln\omega < \ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right) < -\ln\alpha, \\ -\ln\omega & \text{if } \ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right) \leq -\ln\alpha + \rho\ln\omega. \end{cases}$$

Then, for each subject n, we choose the parameters, α and ρ, to minimize

$$\sum_{i=1}^{50}\left[\ln\left(\frac{x_1^i}{x_2^i}\right) - f\left(\ln\left(\frac{\bar{x}_2^i}{\bar{x}_1^i}\right); \alpha, \rho, \omega\right)\right]^2.$$

Before proceeding to estimate the parameters, we omit the nine subjects with CCEI scores below 0.80 (ID 201, 211, 310, 321, 325, 328, 406, 504 and 603) as their choices are not sufficiently consistent to be considered utility-generated. We also exclude the three subjects (ID 205, 218 and 320) who almost always chose a minimum level of consumption of ten tokens in each state, and the single subject (ID 508) who almost always chose a boundary portfolio. This leaves a total of 80 subjects (86.0 percent) for whom we recover preferences by estimating the model. Finally, we note that out of the 80 subjects, 33 subjects (41.3 percent) have no boundary observations and this increases to a total of 60 subjects (75.0 percent) if we consider subjects with less than five boundary observations.

Appendix IV presents the results of the estimations $\hat{\alpha}_n$ and $\hat{\rho}_n$ for the full set of subjects. Table 1 below displays summary statistics for the estimation results. Of the 80 subjects listed in Appendix IV, 56 subjects (70.0 percent) exhibit kinky preferences ($\hat{\alpha}_n > 1$). Also, a significant fraction of our subjects in both treatments have moderate levels of $\hat{\rho}_n$. However, our specification allows the kink (α) to "absorb" some of the curvature in the indifference curves (ρ). More importantly, because the model has two parameters, α and ρ, it is not obvious how to define (a single measure of) risk aversion. In the next section, we define one particularly useful measure and discuss its properties.

[Table 1 here]

Figure 6 presents, in graphical form, the data from Appendix IV by showing a scatterplot of $\hat{\alpha}_n$ and $\hat{\rho}_n$ split by symmetric (black) and asymmetric (white) treatments. Nine subjects with high values for $\hat{\rho}_n$ (ID 206, 210, 304, 306, 404, 407, 515, 516, and 606) and one subject (ID 203) with a high value of $\hat{\alpha}_n$ are omitted to facilitate presentation of the data. The most notable features of the distributions in Figure 6 are that both the symmetric and asymmetric subsamples exhibit considerable heterogeneity in both $\hat{\alpha}_n$ and $\hat{\rho}_n$ and that their values are not correlated ($r^2 = 0.000$).

[Figure 6 here]

Finally, Figure 7 shows the relationship between $\ln(p_1/p_2)$ and $\ln(\hat{x}_1/\hat{x}_2)$ for the same group of subjects (ID 304, 303, 307, 216, and 318) that we followed in the non-parametric analysis. Figure 7 also depicts the actual choices (x_1, x_2). The figures for the full set of subjects are available in Appendix V. An inspection of the estimation results against the observed data reveals that the fit is quite good for most subjects. However, it also shows that the specification has difficulty dealing with the subject (ID 307) who combines safe portfolios for values of $\ln(p_1/p_2)$ close to zero with boundary portfolios for values of $\ln(p_1/p_2)$ that give steep or flat budget lines. His estimated parameters $\hat{\alpha} = 1.043$ and $\hat{\rho} = 0.076$ may be reasonable given the fact that boundary portfolios are chosen also for intermediate values of $\ln(p_1/p_2)$, but leaves the safe portfolio choices largely unexplained. For similar reasons, the estimated curve does not pick up the apparent kink in the scatterplot of the subject (ID 318) with $\hat{\alpha} = 1.056$ and $\hat{\rho} = 0.173$ that often chose safe portfolios. Clearly, no continuous relationship could replicate these patterns.

The estimation also seems sensitive to "outliers," as can be seen in the case of the subject (ID 303) with $\hat{\alpha} = 1.641$ and $\hat{\rho} = 0.284$, who is the only subject that very precisely implemented logarithmic preferences, apart from a small number of deviations. Although his behavior is very regular and fits with standard preferences, the attempt to fit the outlying observations exaggerates the non-linearity and leads to the insertion of a spurious kink. Apart from this subject, the individual-level relationship between $\ln(p_1/p_2)$

and $\ln(\hat{x}_1/\hat{x}_2)$ does not have a kink unless one is clearly identifiable in the data. In fact, a review of our full set of subjects shows that the estimation is more likely to ignore a kink that is evident in the data than to invent one that is not there. Perhaps most notably, the estimation fits the "switch" points, when they exist, quite well.

[Figure 7 here]

Finally, we note that while we have followed prior literature in using a constant relative risk aversion (CRRA) specification, we are concerned that our estimates may be sensitive to this assumption. In particular, one difficulty with assuming CRRA is that behavior depends on the initial level of wealth ω_0, and since ω_0 is unobserved, the model is not completely identified. In the analysis above, we have followed the standard procedure of setting $\omega_0 = 0$. We have also estimated the model with the assumption of constant absolute risk aversion (CARA). The CARA utility function has two advantages. First, it allows us to get rid of the nuisance parameter ω_0 (which bedevils most attempts estimate power utility functions). Secondly, it easily accommodates boundary portfolios.

The problem with CARA is that it implies a (non-linear) relationship between $\log(p_1/p_2)$ and $x_1 - x_2$. Since the variation in $\log(p_1/p_2)$ is quite small relative to the variation in $x_1 - x_2$, the estimated individual-level regression coefficients are bound to be small. This implies that the estimated coefficients of absolute risk aversion and loss/disappointment aversion will be small too, but this seems unlikely given the behavior of the subjects, which suggests a non-negligible degree of risk aversion. We refer the interested reader to Appendix VI for precise details on the CARA specification and results of the estimations.

4.2 Measuring risk aversion

Since we have estimated a two-parameter utility function, risk aversion cannot be represented by a single univariate measure. To summarize the risk aversion of our subjects, we use the concept of the *risk premium*. Specifically, we propose a gamble over wealth levels which offers $50-50$ odds of winning or losing some fraction $0 < h < 1$ of the individual's initial wealth ω_0. The risk premium for h is the fraction of wealth r that satisfies the certainty equivalence relationship

$$(1+\alpha)u(\omega_0(1-r)) = \alpha u(\omega_0(1-h)) + u(\omega_0(1+h)).$$

Substituting the power function yields

$$(1+\alpha)(1-r)^{1-\rho} = \alpha(1-h)^{1-\rho} + (1+h)^{1-\rho},$$

which is independent of the initial wealth level ω_0. This equation can be rearranged to yield

$$r(h) = 1 - \left[\frac{\alpha(1-h)^{1-\rho} + (1+h)^{1-\rho}}{1+\alpha}\right]^{\frac{1}{1-\rho}}.$$

To help us understand the meaning of the parameters α and ρ, Figure 8 below plots the risk premium $r(h)$ for different values of α and ρ. Note that an increase in α makes the risk premium curve $r(h)$ steeper and an increase in ρ makes it more convex.

[Figure 8 here]

To see the role of α and ρ more clearly, we consider the second-order approximation of $r(h)$. Direct calculation yields

$$\begin{aligned} r(h) &\approx r(0) + r'(0)h + r''(0)\frac{h^2}{2} \\ &= 0 + \frac{\alpha-1}{\alpha+1}h + \rho\frac{2\alpha}{(\alpha+1)^2}h^2, \end{aligned}$$

which reduces to the usual case $r(h) \approx \rho\frac{h^2}{2}$ when $\alpha = 1$. The approximation clearly tells us that α has a first-order effect on the risk premium r while ρ has a second-order effect, so the standard practice of considering small gambles is inadequate. Motivated by the second-order approximation of $r(h)$, we calculate the following weighted average of ρ and α:

$$\begin{aligned} r(1) &\approx \frac{\alpha-1}{\alpha+1} + \rho\frac{2\alpha}{(\alpha+1)^2} \\ &\approx \frac{\alpha-1}{\alpha+1} + \rho, \end{aligned}$$

which reduces to the Arrow-Pratt measure of relative risk aversion ρ when $\alpha = 1$. We will use $r(1)$ as a summary measure of risk aversion.

Although there is no strong theoretical rationale for adopting this formula as our summary measure of risk aversion, it agrees with other measures of risk aversion. As a benchmark, we use the "low-tech" approach of estimating an individual-level power utility function directly from the data.

By straightforward calculation, the solution to the maximization problem (x_1^*, x_2^*) satisfies the first-order condition

$$\frac{\pi}{1-\pi}\left(\frac{x_2^*}{x_1^*}\right)^{\rho} = \frac{p_1}{p_2}$$

and the budget constraint $p \cdot x^* = 1$. This generates the following individual-level econometric specification for each subject n:

$$\log\left(\frac{x_{2n}^i}{x_{1n}^i}\right) = \alpha_n + \beta_n \log\left(\frac{p_{1n}^i}{p_{2n}^i}\right) + \epsilon_n^i$$

where ϵ_n^i is assumed to be distributed normally with mean zero and variance σ_n^2. We generate estimates of $\hat{\alpha}_n$ and $\hat{\beta}_n$ using ordinary least squares (OLS), and use this to infer the values of the underlying parameter $\hat{\rho}_n = 1/\hat{\beta}_n$.

Before proceeding to the estimations, we again omit the nine subjects with CCEI scores below 0.80 as well the four subjects (ID 307, 311, 324 and 508) for whom the simple power formulation is not well defined. This leaves the group of 77 subjects (82.8 percent) for whom we estimated parameters. For these subjects, we discard the boundary observations, for which the power function is not well defined, using a narrow confidence interval of one token (if $x_1^i \leq 1$ or $x_2^i \leq 1$ then x^i is treated as a boundary portfolio). This results in many fewer observations for a small number of subjects.

Appendix VII lists the estimated risk measures $\hat{r}_n$ and values of $\hat{\rho}_n$ derived from the simple OLS estimation for the full set of subjects except for the two subjects (ID 508 and 603) who always choose boundary portfolios. The last column of Appendix VII reports the number of observations per subject in the OLS estimation. Table 2 below displays summary statistics. Most notably, the distribution shifts to the left when calculated using the $\hat{r}_n$ estimates as compared to the distribution calculated using the OLS $\hat{\rho}_n$ estimates. The reason may be the upward bias in the OLS estimates due to the omission of boundary observations.

[Table 2 here]

Figure 9 shows a scatterplot of $\hat{r}_n$ and $\hat{\rho}_n$, split by symmetric (black) and asymmetric (white) treatments. Subjects with high values for $\hat{r}_n$ or $\hat{\rho}_n$ (ID 203, 204, 206, 210, 304, 306, 314, 404, 407, 408, 413, 515, 516, 606 and 607) are omitted to facilitate presentation of the data. Note that we obtain once more very similar distributions for the symmetric and asymmetric subsamples, and that there is a strong correlation between the estimated $\hat{r}_n$ parameters and individual-level estimates of $\hat{\rho}_n$ that come from a simple expected-utility model ($r^2 = 0.488$).

[Figure 9 here]

Much of the existing evidence about risk preferences is based on laboratory experiments. Our individual-level measures of risk aversion are very similar to some recent estimates that come out of the simple expected-utility model. For comparison, Chan and Plott (1998) and Goeree, Holt and Palfrey (2002) report, respectively, $\rho = 0.48$ and 0.52 for private-value auctions. Goeree, Holt and Palfrey (2003) estimate $\rho = 0.44$ for asymmetric matching pennies games, and Goeree and Holt (2004) report $\rho = 0.45$ for a variety of one-shot games. Holt and Laury (2002) estimate individual degrees of risk aversion from ten paired lottery-choices under both low- and high-money payoffs. Most of their subjects in both treatments exhibit risk preferences in the $0.3 - 0.5$ range.

5 Conclusion

We present a set of experimental results which build on a graphical computer interface that contains a couple of important innovations over previous work. The primary contribution is an experimental technique for collecting richer data on choice under uncertainty than was previously possible. Perhaps the most interesting aspect of the data set generated by this approach is the granularity of behavior. Subjects' behavior appears to be made up of a small number of stylized patterns of behavior, sometime choosing safe portfolios, sometimes choosing boundary portfolios, and sometimes choosing intermediate portfolios. In the present paper, we have shown that this behavior can be rationalized by "kinky" preferences that are consistent with loss or disappointment aversion. The potential of this data set to teach us about individual behavior has not been exhausted, however. One aspect of the data that invites further scrutiny is the rather curious switching between stylized behavior patterns exhibited by some subjects. We plan to explore this and other themes in future work based on further extensions of our experimental design.

References

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

[2] **Afriat, Sidney N. 1972**. "Efficiency Estimation of Production Functions." *International Economic Review* 13(3): 568-598.

[3] **Bronars, Stephen G.** 1987. "The power of nonparametric tests of preference maximization." *Econometrica* 55(3): 693-698.

[4] **Camerer, Colin F.** 1995. "Individual Decision Making." In *Handbook of Experimental Economics*, ed. John H. Kagel and Alvin E. Roth. Princeton: Princeton University Press.

[5] **Chen, Kay-Yen, and Charles R. Plott.** 1998. "Nonlinear Behavior in Sealed Bid First Price Auctions." *Games and Economic Behavior* 25(1): 34-78.

[6] **Chew, Soo Hong. 1983.** "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox." *Econometrica* 51(4): 1065-1092.

[7] **Choi, Syngjoo, Fisman, Ray, Douglas M. Gale, and Shachar Kariv.** Forthcoming. "Revealing Preferences Graphically: An Old Method Gets a New Tool Kit." *American Economic Review* Papers and Proceedings.

[8] **Currim, Imran S., and Rakesh K. Sarin.** 1989. "Prospect versus Utility." *Management Science* 35(1): 22-41.

[9] **Daniels, Richard L., and L. Robin Keller.** 1990. "An Experimental Evaluation of the Descriptive Validity of Lottery-dependant Utility Theory." *Journal of Risk and Uncertainty* 3(2): 15-134.

[10] **Dekel, Eddie. 1986.** "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom." *Journal of Economic Theory* 40(2): 304-318.

[11] **Fisman, Ray, Shachar Kariv, and Daniel Markovits.** 2006. "Individual Preferences for Giving." mimeo.

[12] **Goeree, Jacob K., and Charles A. Holt.** 2004. "A Model of Noisy Introspection." *Games and Economic Behavior* 46(2): 365-382.

[13] **Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey.** 2002. "Quantal Response Equilibrium and Overbidding in Private Value Auctions," *Journal of Economic Theory* 104(1): 247-272.

[14] **Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey.** 2003. "Risk Averse Behavior in Generalized Matching Pennies Games." *Games and Economic Behavior* 45(1): 97-113.

[15] **Gul, Faruk. 1991.** "A Theory of Disappointment in Decision Making under Uncertainty." *Econometrica* 59(3) 667–686.

[16] **Harless, David W., and Colin F. Camerer.** 1994. "The Predictive Utility of Generalized Expected Utility Theories." *Econometrica*, 62(6): 1251-1289.

[17] **Hey, John D., and Chris Orme.** 1994. "Investigating Generalizations of Expected Utility Theory Using Experimental Data." *Econometrica*, 62(6): 1291-1326.

[18] **Holt, Charles A., and Susan K. Laury.** 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92(5): 1644-1655.

[19] **Houtman, Martijn, and J. A. H. Maks.** 1985. "Determining all Maximal Data Subsets Consistent with Revealed Preference." *Kwantitatieve Methoden* 19: 89-104.

[20] **Lattimore, Pamela K., Joanna R. Baker, and A. Dryden Witte.** 1992. "The Influence of Probability on Risky Choice: A parametric Investigation." *Journal of Economic Behavior and Organization* 26: 293-304.

[21] **Loomes, Graham.** 1991. "Evidence of a new violation of the independence axiom." *Journal of Risk and Uncertainty* 4(1): 91-108.

[22] **Neumann, John, and Oskar Morgenstern.** 1947. *The Theory of Games and Economic Behavior*, 2nd ed. Princeton: Princeton University Press.

[23] **Savage, Leonard.** 1954. *The Foundations of Statistics*. New York: Wiley.

[24] **Starmer, Chris.** 2000. "Developments in Non-Expected Utility Theory: The Hunt for a descriptive Theory of Choice under Risk." *Journal of Economic Literature* 38(2): 332-382.

[25] **Varian, Hal R.** 1982 "The Nonparametric Approach to Demand Analysis." *Econometrica* 50(4): 945-972.

[26] **Varian, Hal R.** 1991. "Goodness-of-Fit for Revealed Preference Tests." mimeo.

Table 1: Summary statistics of individual-level estimation

α	All	$\pi=1/2$	$\pi\neq1/2$
Mean	1.315	1.390	1.248
Std	0.493	0.584	0.388
p5	1.000	1.000	1.000
p25	1.000	1.000	1.000
p50	1.115	1.179	1.083
p75	1.445	1.477	1.297
p95	2.427	2.876	2.333

ρ	All	$\pi=1/2$	$\pi\neq1/2$
Mean	1.662	2.448	0.950
Std	7.437	10.736	1.206
p5	0.053	0.048	0.080
p25	0.233	0.165	0.290
p50	0.481	0.438	0.573
p75	0.880	0.794	0.990
p95	3.803	3.871	3.693

Table 2. Summary statistics of risk measures

r(1)	All	π=1/2	$\pi\neq$1/2
Mean	1.834	1.041	2.608
Std	7.580	1.211	10.594
p5	0.201	0.114	0.201
p25	0.432	0.411	0.465
p50	0.612	0.613	0.609
p75	0.990	1.076	0.983
p95	3.865	3.865	3.953

OLS	All	π=1/2	$\pi\neq$1/2
Mean	3.168	1.401	4.888
Std	15.025	1.362	21.060
p5	0.439	0.439	0.375
p25	0.648	0.597	0.700
p50	0.904	0.826	1.011
p75	1.434	1.426	1.533
p95	5.348	5.158	5.448

Figure 1: An example of a budget constraint with two states and two assets.

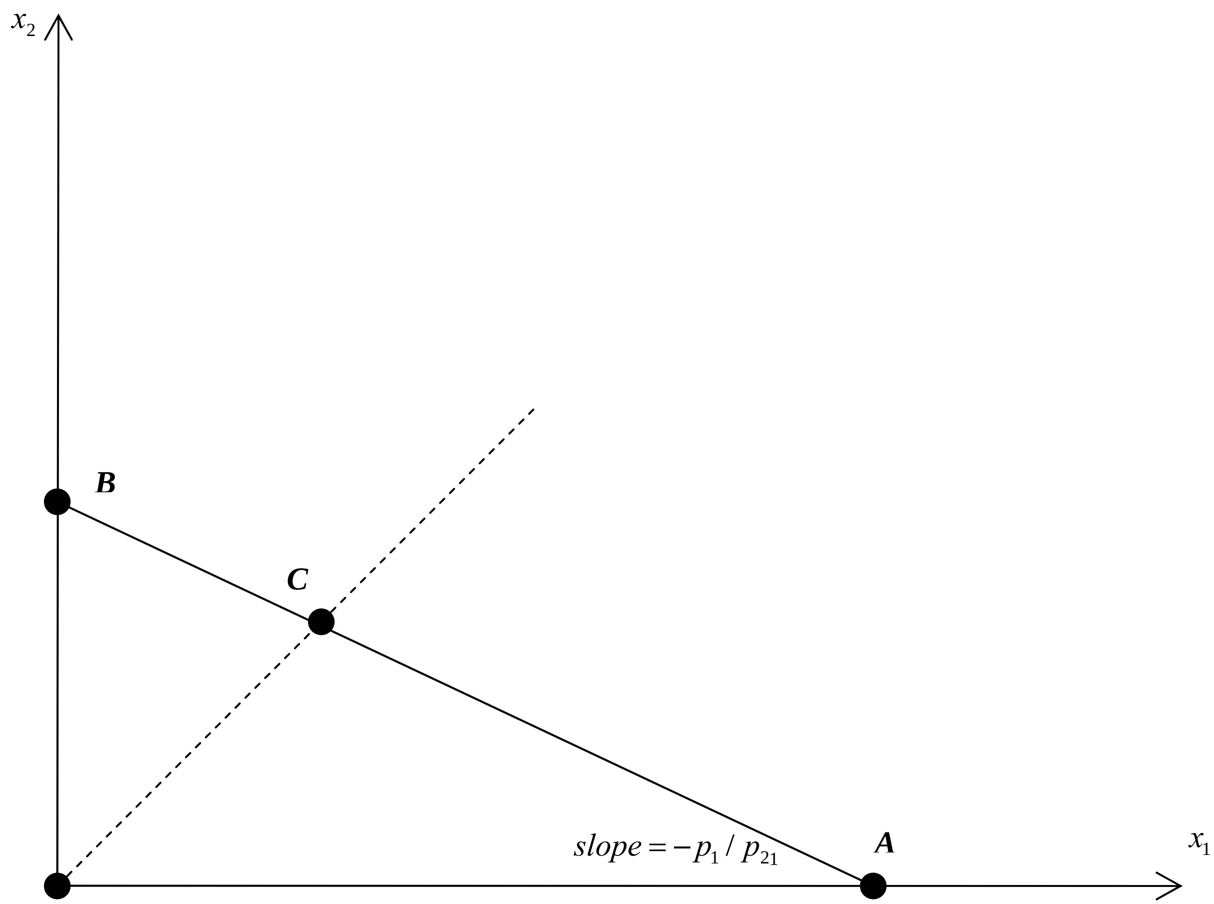

Figure 2: The relationship between the log-price ratio $\ln(p_1 / p_2)$ and the token share $x_1 / (x_1 + x_2)$ for selected subjects.

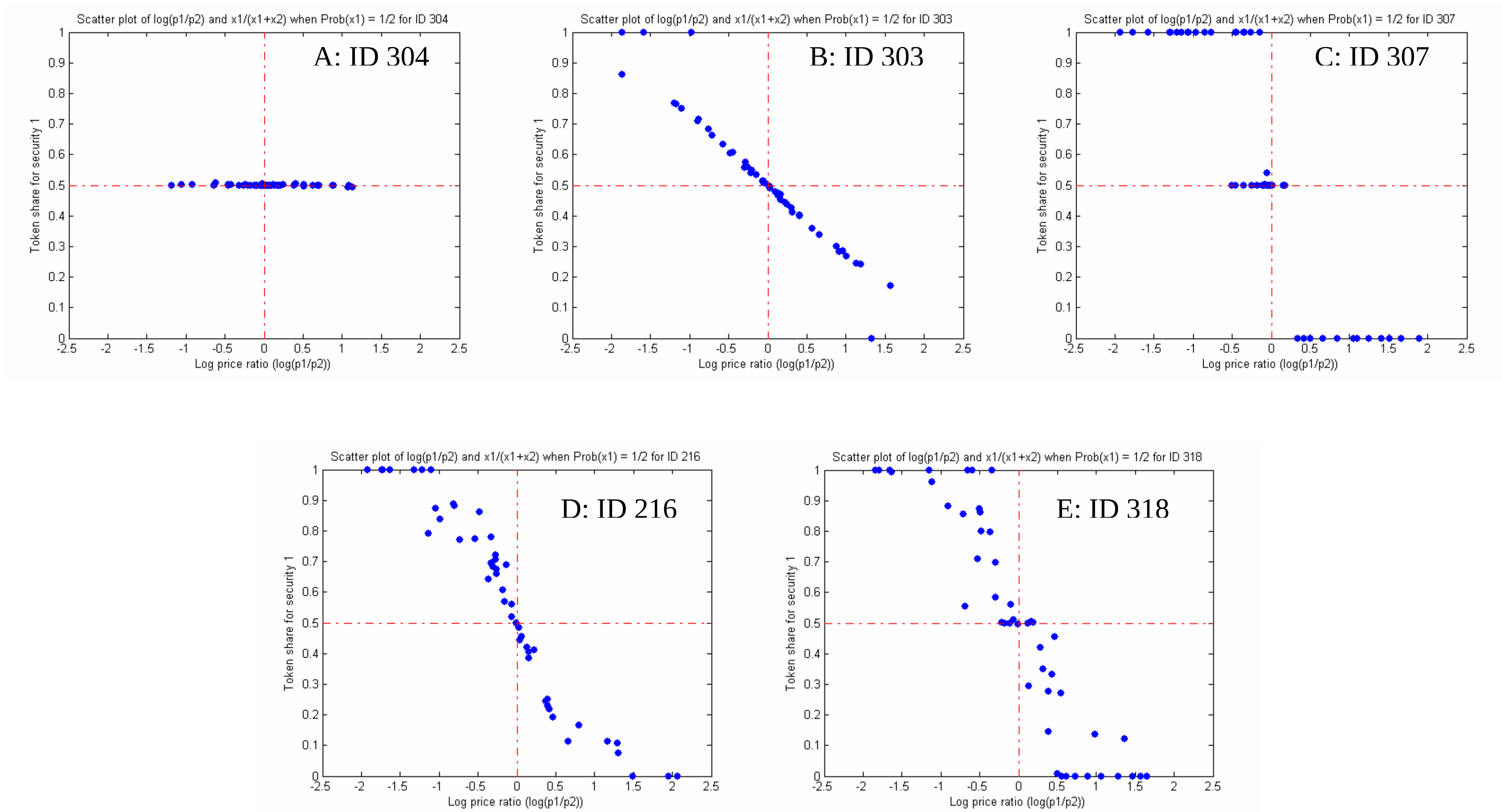

Figure 3: The construction of the CCEI for a simple violation of GARP.

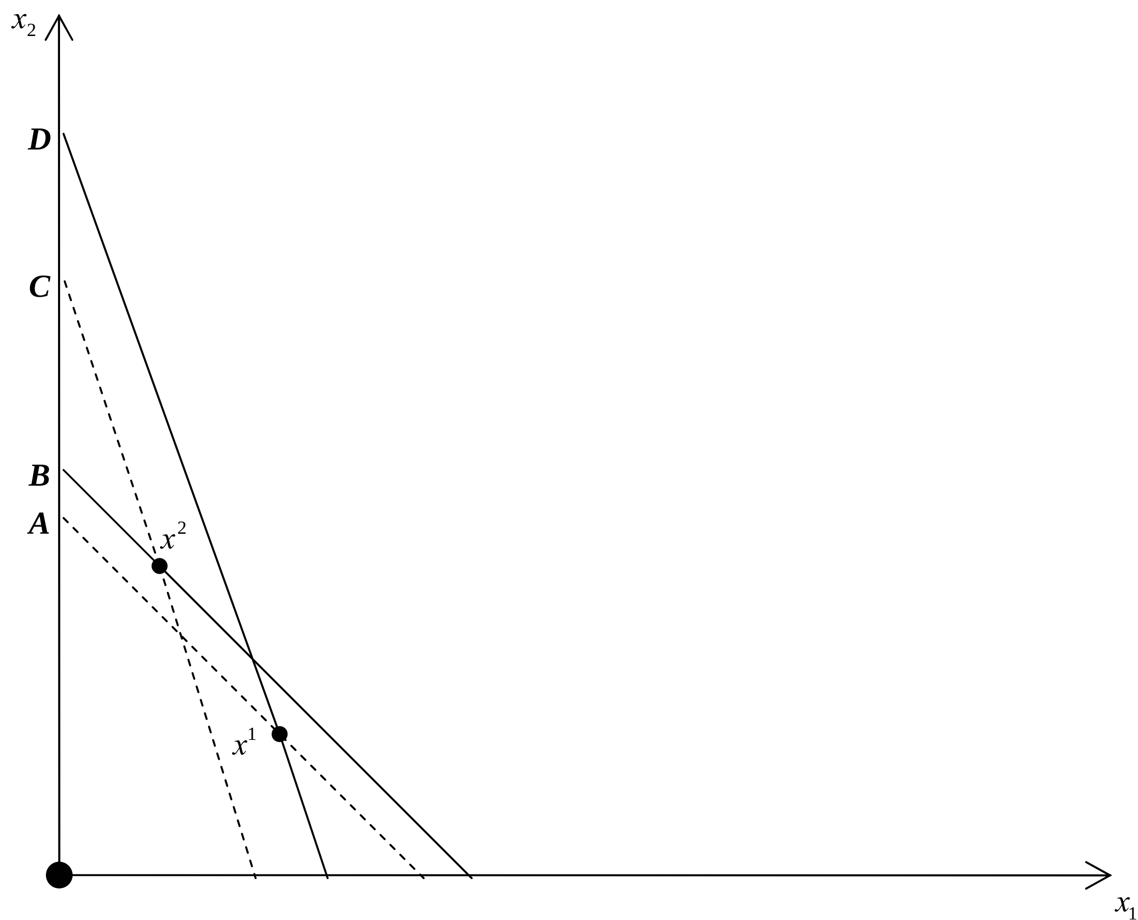

Figure 4: The distributions of GARP violations Afriat's (1972) efficiency index (CCEI).

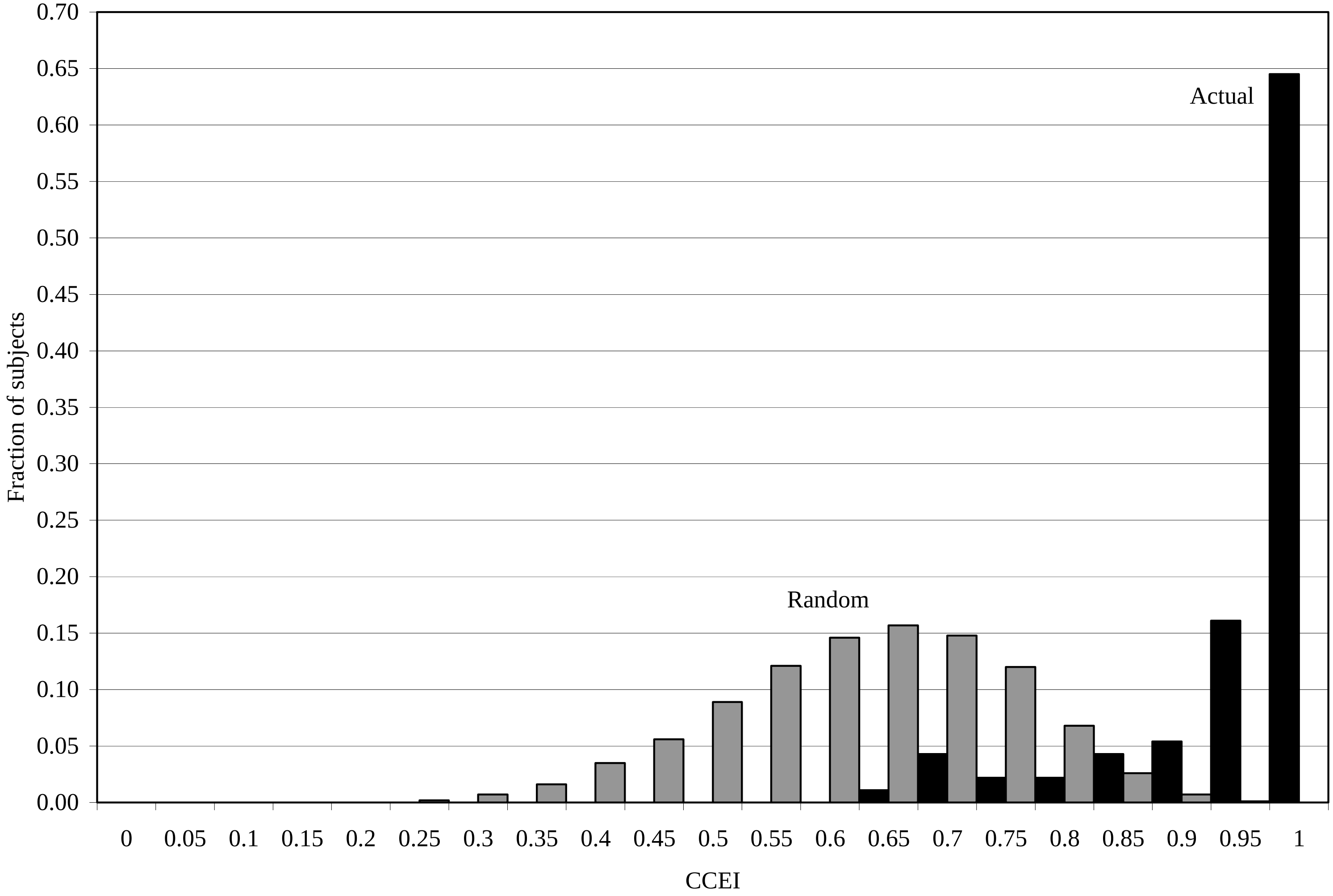

Figure 5: An illustration of the relationship between $\ln(p_1 / p_2)$ and $\ln(x_1 / x_2)$

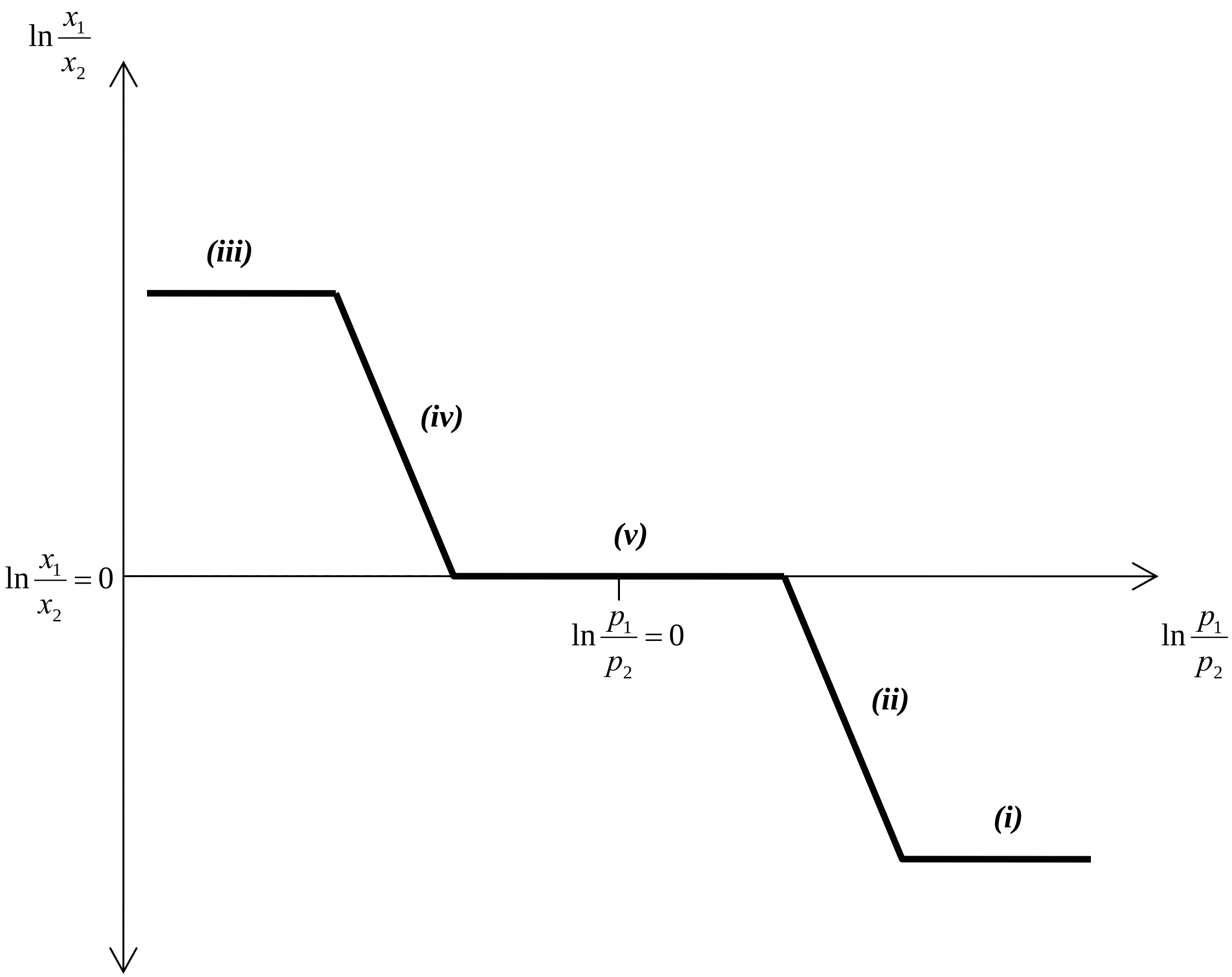

Figure 6: Scatterplot of the estimated parameters $\hat{\alpha}_n$ and $\hat{\rho}_n$.

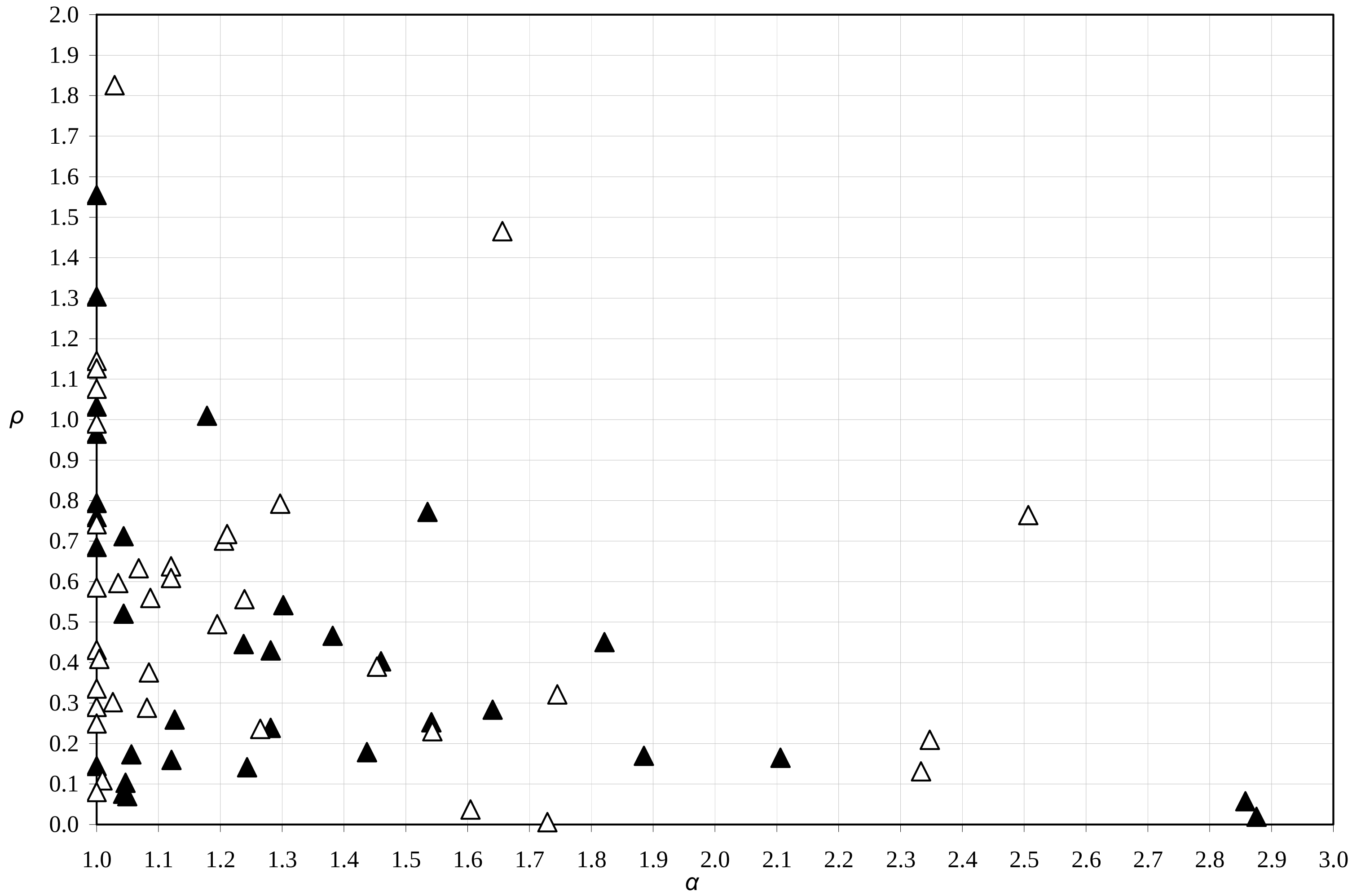

Figure 7: The relationship between $\ln(p_1 / p_2)$ and $\hat{x}_1 / \hat{x}_2$ for selected subjects.

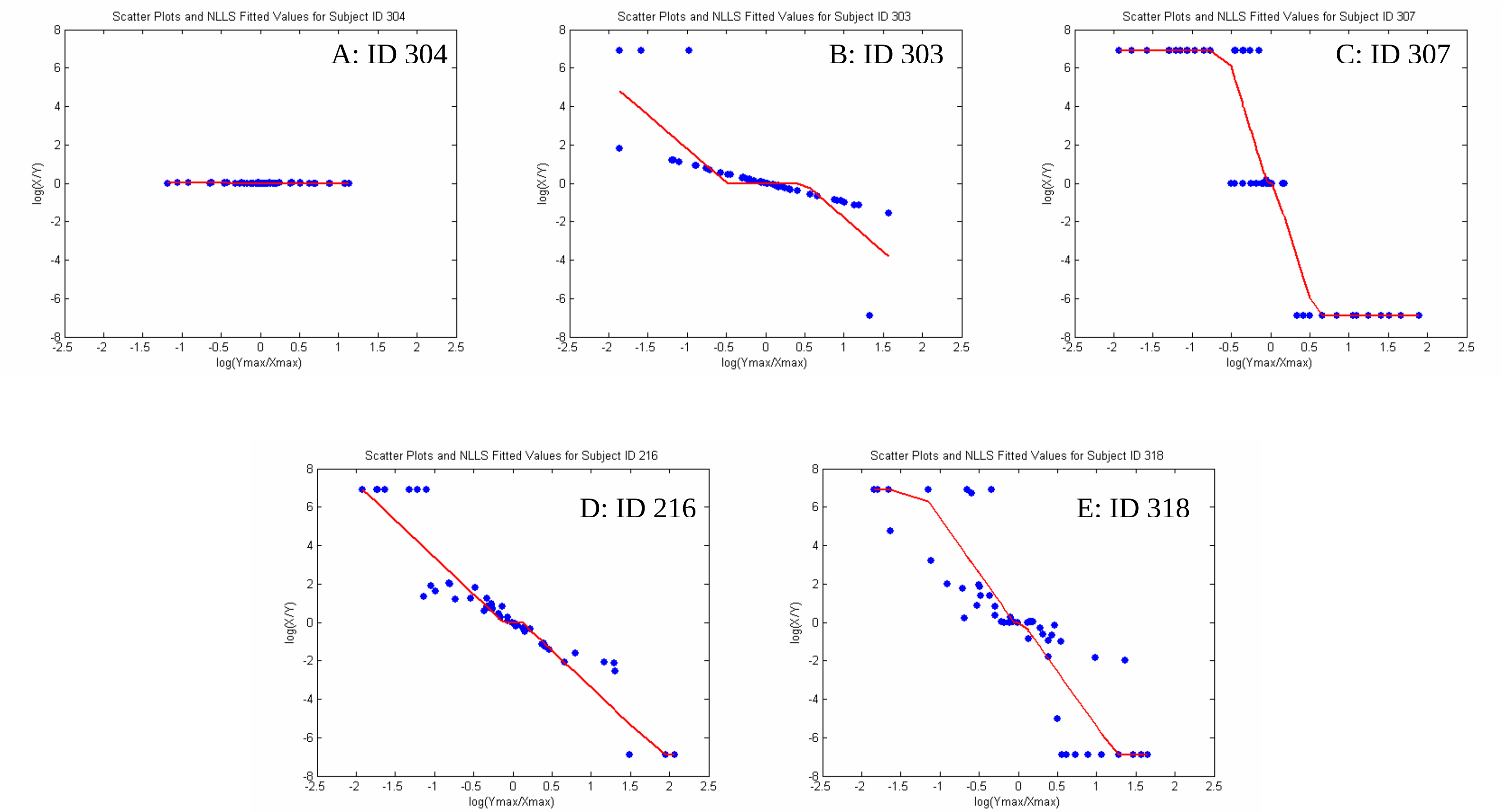

Figure 8: The risk premium $r(h)$ for different values of α and ρ.

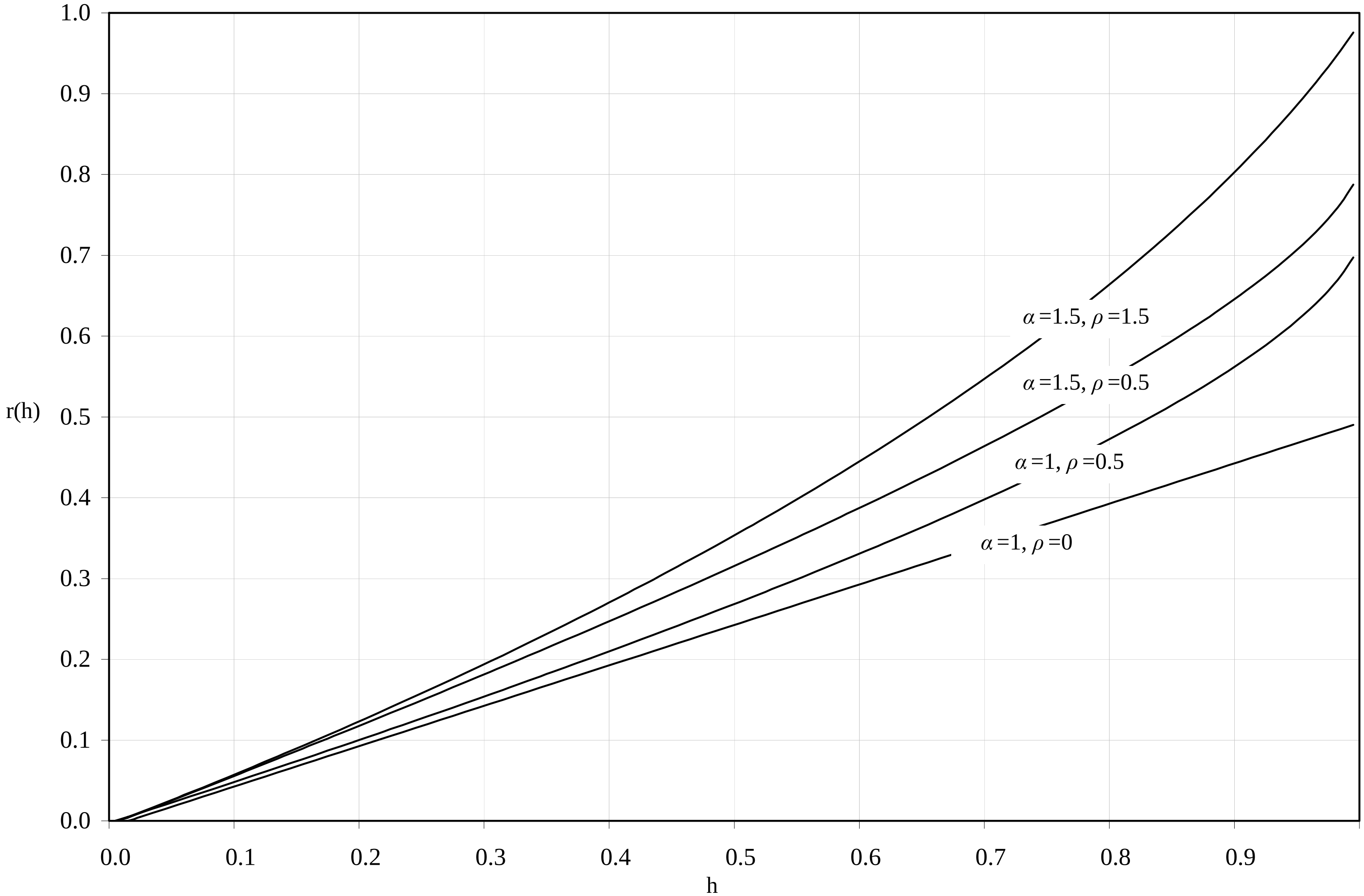

Figure 9: A Scatterplot of the risk measures $\hat{r}_n$ and values $\hat{\rho}_n$ derived from the simple OLS estimation.