Appendix II
Testing for consistency

I. Afriat's (1967) Theorem

Let $\{(p^i, x^i)\}_{i=1}^{25}$ be the data generated by some individual's choices, where p^i denotes the i-th observation of the price vector and x^i denotes the associated allocation. More precisely, the data generated by an individual's choices are $\{(\bar{x}_1^i, \bar{x}_2^i, x_1^i, x_2^i)\}_{i=1}^{25}$, where (x_1^i, x_2^i) are the coordinates of the choice made by the subject and $(\bar{x}_1^i, \bar{x}_2^i)$ are the endpoints of the budget line, so we can calculate the budget line $x_1^i/\bar{x}_1^i + x_2^i/\bar{x}_2^i = 1$ for each observation i.

An allocation x^i is *directly revealed preferred* to an allocation x^j, denoted $x^i R^D x^j$, if $p^i \cdot x^i \geq p^i \cdot x^j$. An allocation x^i is *revealed preferred* to x^j, denoted $x^i R x^j$, if there exists a sequence of allocations $\{x^k\}_{k=1}^{K}$ with $x^1 = x^i$ and $x^K = x^j$, such that $x^k R^D x^{k+1}$ for every $k = 1, ..., K-1$.

The Generalized Axiom of Revealed Preference (GARP) requires that if $x^i R x^j$ then $p^j \cdot x^j \leq p^j \cdot x^i$; that is, if x^i is revealed preferred to x^j, then x^i must cost at least as much as x^j at the prices prevailing when x^j is chosen. It is clear that if the data are generated by a non-satiated utility function, then they must satisfy GARP. Conversely, the following result due to Afriat (1967) tells us that if a *finite* data set generated by an individual's choices satisfies GARP, then the data can be rationalized by a well-behaved utility function.

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

This statement of the theorem follows Varian (1982, 1983), who replaced the condition Afriat called *cyclical consistency* with GARP. Note that satisfying GARP entails only that choices are consistent with the utility maximization model. The further implication, that the choices may be rationalized by a well-behaved utility function, is a consequence of linear budget lines. Given that the budget constraints are linear, if a rationalizing utility function exists then we cannot reject the hypothesis that it is well-behaved.

II. Goodness-of-fit

In order to show that the data are consistent with utility-maximizing behavior we must check whether it satisfies GARP. Since GARP offers an exact test, it is desirable to measure the *extent* of GARP violations. We report measures of GARP violations based on three indices: Afriat's (1972) critical cost efficiency index (CCEI), Varian (1990, 1991), and Houtman and Maks (1985).

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

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

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

[Figure 1 here]

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

Varian (1990, 1991) Varian refined Afriat's CCEI to provide a measure that reflects the minimum adjustment required to eliminate the violations of GARP associated with each observation (p^i, x^i). In particular, fix an observation (p^i, x^i) and let e^i be the largest value of e such that $R(e)$ has no violations of GARP within the set of allocations x^j such that $x^i R(e) x^j$. The value e^i measures the efficiency of the choices when compared to the allocation x^i. Knowing the efficiencies $\left\{e^i\right\}$ for the entire set of observations $\left\{\left(p^i, x^i\right)\right\}$ allows us to say where the inefficiency is greatest or least. When a single number is desired, as here, Varian (1990, 1991) uses $e^* = \min\left\{e^i\right\}$. Thus, the Varian (1990, 1991) score associated with the violation of GARP depicted in Figure 1 above is also A/B. More generally, the Varian (1990, 1991) index is a lower bound on the CCEI.

Echenique et al. (2011) Echenique et al. also provide a disaggregated measure that indicates the amount of money one can extract from an individual for each violation of GARP. Their measure is based on the idea that an individual who violates GARP can be exploited as a "money pump." The construction of their money pump index for a simple violation of GARP is also illustrated in Figure 1 above. An "arbitrager" who chooses allocation x^1 at prices p^2 and allocation x^2 at prices p^1 could profitably trade x^1 with the individual at prices p^1 and x^2 at prices p^2, yielding a profit

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

Echenique et al. (2011) use money pump index to measure the extent of each GARP violation. To summarize consistency, they use the mean and median money pump scores across all violations of GARP (cyclic sequences of allocations). The reasons for the discrepancies between the CCEI and the Varian (1990, 1991) index and the money pump index are discussed in Echenique et al. (2011).

Houtman and Maks (1985) Houtman and Maks find the largest subset of choices that is consistent with GARP. This method has a couple of drawbacks. First, some observations may be discarded even if the associated GARP violations could be removed by small perturbations of the budget line. Second, since the algorithm is computationally very intensive, for a small number of subjects we report upper bounds on the consistent set. We compute the Houtman and Maks (1985) scores using the algorithm developed by Dean and Martin (2010).

In reporting our results, we focus on the CCEI, which offers a straightforward interpretation. The econometric results based on the indices proposed by Varian (1990, 1991) and Houtman and Maks (1985) are presented in Appendix III.[1] In practice, these measures yield qualitatively very similar results. We therefore do not repeat the econometric analysis with the "money pump" measure of Echenique et al. (2011). Figure 2 summarizes the mean Varian (1990, 1991) and Houtman and Maks (1985) scores and 95 percent confidence intervals across selected socioeconomic categories. Table 1 below provides a summary of each consistency score. There is considerable heterogeneity within and across categories for all measures.

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

[Figure 2 here]
[Table 1 here]

III. The power of the GARP tests

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

To generate a benchmark against which to compare these CCEI scores, we use the test designed by Bronars (1987), which builds on Becker (1962) and employs the choices of a hypothetical subject who chooses randomly among all allocations on each budget line as a point of comparison. The mean CCEI score across all subjects in our experiment is 0.881 whereas the mean CCEI score for a random sample of 25,000 simulated subjects is 0.659. More than half of actual subjects have CCEI's above 0.925, while only about five percent of simulated subjects have CCEI's that high.

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

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

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

The histograms in Figure 3 below summarize the distributions of CCEI scores generated by samples of 25,000 simulated subjects who implement the logarithmic von Neumann-Morgenstern utility function $\log x_1 + \log x_2$ with various levels of precision γ. The horizontal axis measures the fractions for different intervals of CCEI scores and the vertical axis measures the percentage of subjects corresponding to each interval. Each of the simulated subjects makes 25 choices from randomly generated budget lines in the same way as the human subjects do. The number above each bar of the histogram represents the percentage of actual subjects corresponding to each interval. The histograms show the extent to which subjects did worse than choosing consistently and the extent to which they did better than choosing randomly. The histograms thus demonstrate that if utility maximization is not in fact the correct model, then our experiment is sufficiently powerful to detect it.

[Figure 3 here]

IV. Additional references

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

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

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

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

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

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

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

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

Table 1. Consistency scores

A. CCEI

	Mean	Std. Dev.	Percentiles 10	25	50	75	90	# of obs.
All	0.881	0.141	0.676	0.808	0.930	0.998	1.000	1182
Female	0.874	0.147	0.666	0.796	0.928	0.998	1.000	537
Age								
16-34	0.920	0.119	0.734	0.881	0.979	1.000	1.000	219
35-49	0.906	0.123	0.708	0.853	0.966	1.000	1.000	309
50-64	0.863	0.142	0.666	0.784	0.901	0.985	1.000	421
65+	0.843	0.164	0.595	0.770	0.882	0.981	1.000	233
Education								
Low	0.863	0.143	0.665	0.782	0.906	0.987	1.000	397
Medium	0.881	0.140	0.689	0.814	0.926	0.998	1.000	351
High	0.899	0.137	0.686	0.842	0.963	1.000	1.000	430
Household monthly income								
€0-2499	0.856	0.154	0.617	0.769	0.911	0.983	1.000	269
€2500-3499	0.885	0.133	0.705	0.809	0.925	0.999	1.000	302
€3500-4999	0.882	0.141	0.649	0.817	0.932	0.999	1.000	345
€5000+	0.901	0.131	0.729	0.836	0.968	1.000	1.000	266
Occupation								
Paid work	0.896	0.131	0.705	0.833	0.950	1.000	1.000	628
House work	0.873	0.151	0.649	0.795	0.937	0.999	1.000	137
Retired	0.839	0.158	0.597	0.767	0.876	0.971	1.000	247
Others	0.891	0.129	0.712	0.809	0.936	0.998	1.000	170
Household composition								
Partnered	0.878	0.142	0.673	0.802	0.927	0.998	1.000	956
Children	0.899	0.128	0.704	0.835	0.959	1.000	1.000	490

B. Varian (1990, 1991)

	Mean	Std. Dev.	Percentiles 10	25	50	75	90	# of obs.
All	0.736	0.262	0.330	0.515	0.820	0.991	1.000	1182
Female	0.724	0.268	0.325	0.484	0.804	0.989	1.000	537
Age								
16-34	0.818	0.236	0.418	0.670	0.945	1.000	1.000	219
35-49	0.782	0.246	0.398	0.590	0.882	1.000	1.000	309
50-64	0.699	0.263	0.296	0.479	0.764	0.949	1.000	421
65+	0.664	0.272	0.293	0.427	0.687	0.941	1.000	233
Education								
Low	0.696	0.268	0.301	0.452	0.760	0.961	1.000	397
Medium	0.734	0.253	0.380	0.515	0.787	0.990	1.000	351
High	0.776	0.256	0.341	0.600	0.891	1.000	1.000	430
Household monthly income								
€0-2499	0.687	0.263	0.302	0.452	0.726	0.949	1.000	269
€2500-3499	0.739	0.255	0.380	0.520	0.801	0.994	1.000	302
€3500-4999	0.739	0.269	0.315	0.479	0.838	0.993	1.000	345
€5000+	0.778	0.252	0.370	0.583	0.899	1.000	1.000	266
Occupation								
Paid work	0.761	0.255	0.350	0.553	0.863	1.000	1.000	628
House work	0.719	0.281	0.277	0.439	0.821	0.989	1.000	137
Retired	0.663	0.262	0.293	0.437	0.686	0.920	1.000	247
Others	0.760	0.252	0.373	0.536	0.853	0.991	1.000	170
Household composition								
Partnered	0.732	0.263	0.330	0.512	0.818	0.989	1.000	956
Children	0.773	0.252	0.372	0.558	0.883	1.000	1.000	490

C. Houtman and Maks (1985)

	Mean	Std. Dev.	Percentiles 10	25	50	75	90	# of obs.
All	22.361	2.259	19	21	23	24	25	1182
Female	22.289	2.306	19	21	23	24	25	537
Age								
16-34	22.950	2.147	19	22	24	25	25	219
35-49	22.773	2.176	19	21	23	25	25	309
50-64	22.057	2.185	19	21	22	24	25	421
65+	21.811	2.387	19	20	22	24	25	233
Education								
Low	21.990	2.360	19	20	22	24	25	397
Medium	22.342	2.249	19	21	23	24	25	351
High	22.737	2.113	19	21	23	25	25	430
Household monthly income								
€0-2499	22.086	2.263	19	20	22	24	25	269
€2500-3499	22.421	2.187	19	21	23	24	25	302
€3500-4999	22.330	2.384	19	21	23	24	25	345
€5000+	22.613	2.147	19	21	23	24	25	266
Occupation								
Paid work	22.584	2.191	19	21	23	24	25	628
House work	22.307	2.451	19	21	23	24	25	137
Retired	21.672	2.320	18	20	22	24	25	247
Others	22.582	2.063	19	21	23	24	25	170
Household composition								
Partnered	22.304	2.279	19	21	23	24	25	956
Children	22.645	2.189	19	21	23	24	25	490

Figure 1: The construction of the CCEI for a simple violation of GARP

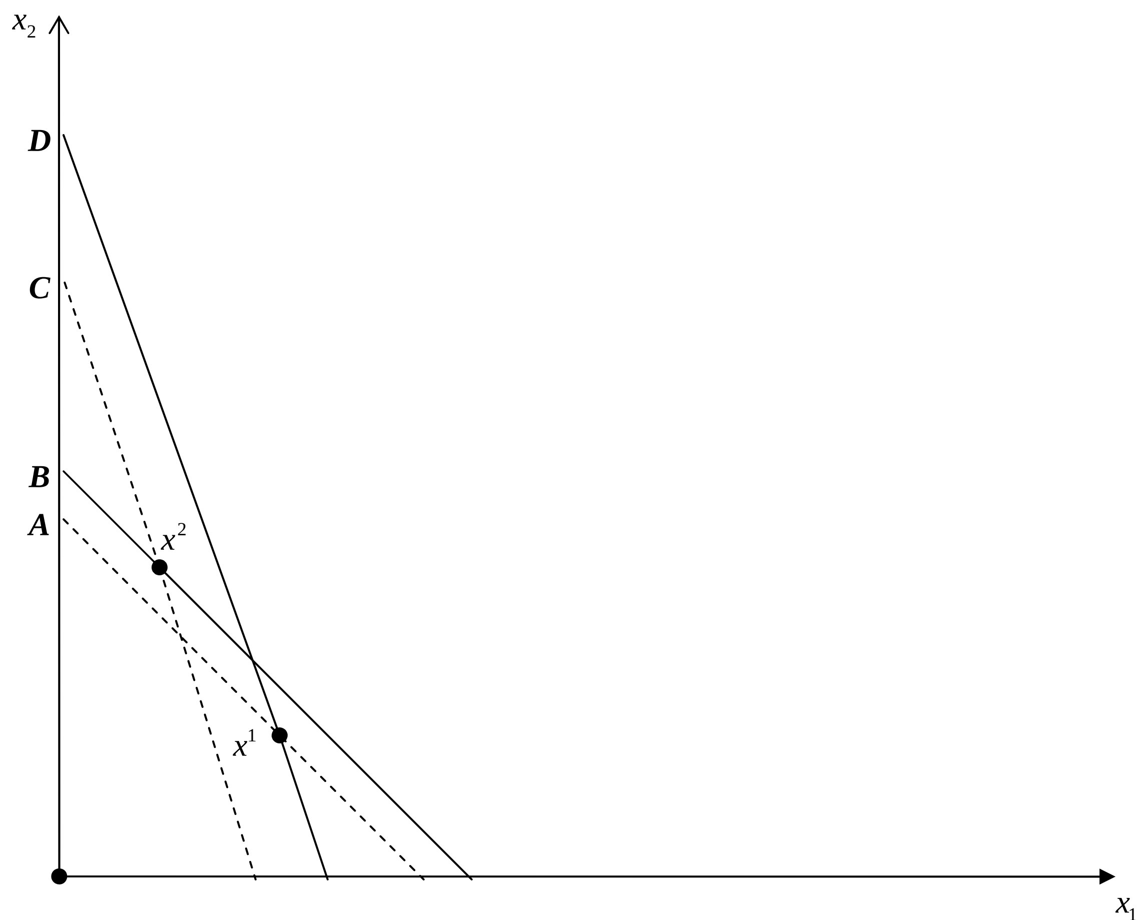

Figure 2A. Varian (1990, 1991) scores

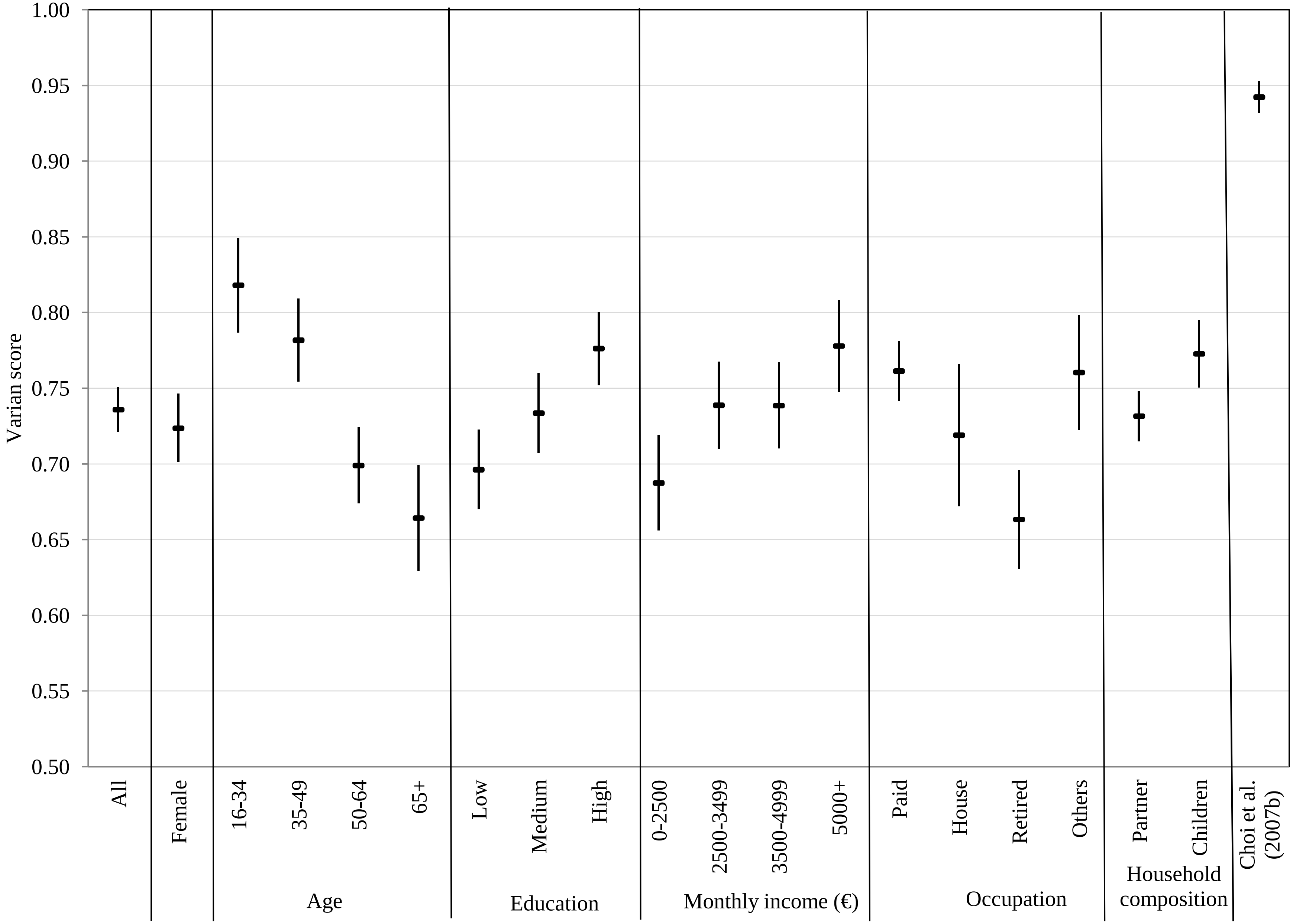

Figure 2B. Houtman and Maks (1985) scores

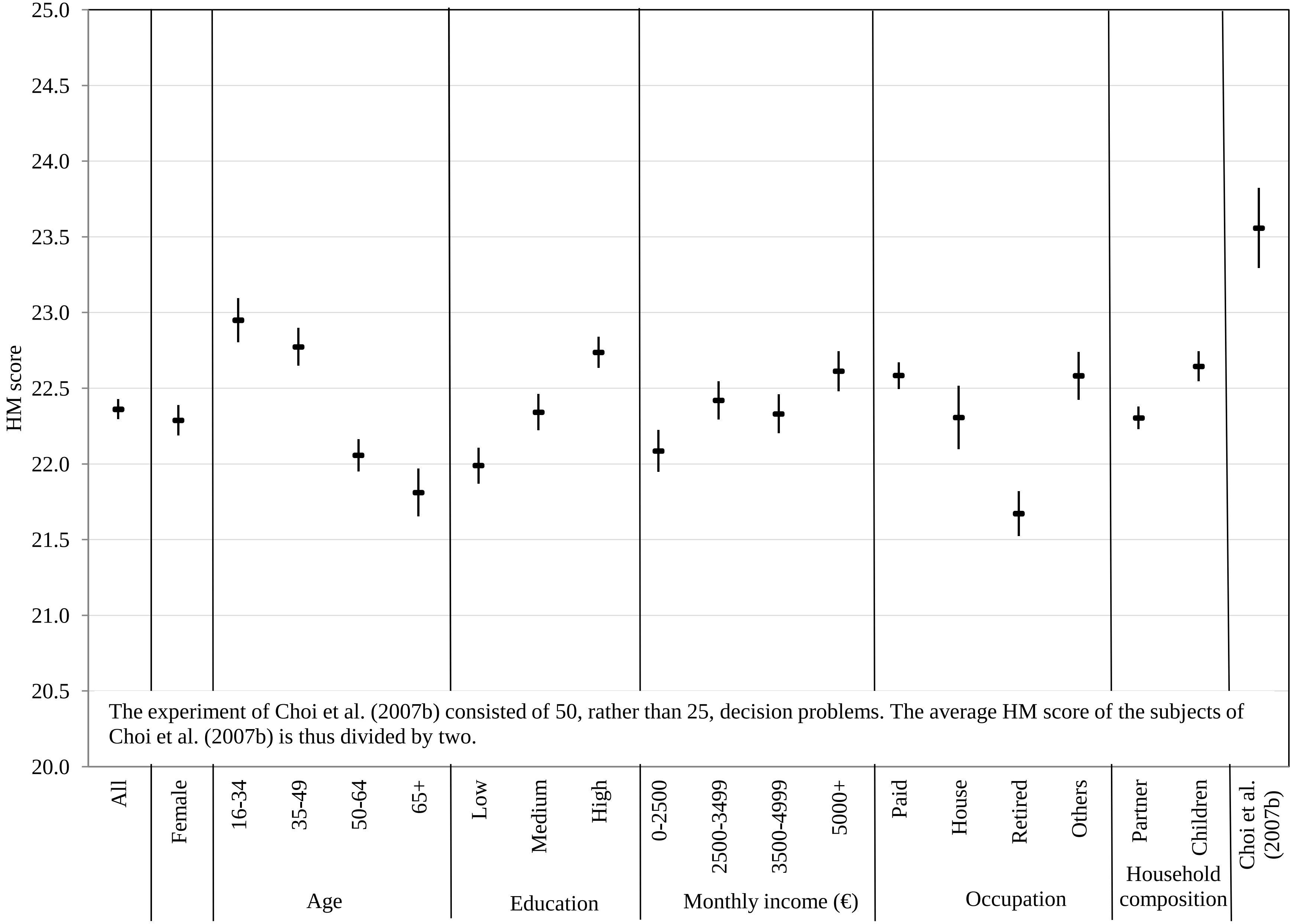

Figure 3.The distributions of CCEI scores of simulated subjects