Revealed Differences

Marco Castillo*, Mikhail Freer**

*Texas A&M University
**George Mason University
Castillo: marco.castillo@tamu.edu, Freer: mfreer@gmu.edu

March 2017

Keywords: rational behavior, heterogeneity of preferences

Abstract: Two individuals are said to be *revealed different* if their joint decisions are more distant from rationality than either of their individual decisions taken separately. We show that the *revealed different* relation can be used to identify preference types and therefore to evaluate the heterogeneity of preferences in a completely nonparametric way. Using experimental data from a random sample of the Dutch population, we find that 1,182 individuals can be divided into 131 different preference types, men's preferences are more heterogeneous than and different from women's preferences and measured heterogeneity is reduced as stricter definitions of rationality are used.

1. Introduction

We investigate the question of the heterogeneity of preferences from the point of view of revealed preference analysis. Specifically, we propose the *revealed different* relation to study differences in preferences in a completely nonparametric way. Two individuals are said to be *revealed different* if their joint decisions are more distant from rationality than either of their individual decisions taken separately. We can use the *revealed different* relation to answer the question of how heterogeneous preferences are within a group and which groups are more heterogeneous. Our approach can be extended to allow alternative and/or stricter versions of rationality, which allows us to evaluate the cost of different behavioral assumptions in detecting heterogeneity.

Our definition allows a comparison between individuals who themselves violate rationality. This is a necessary requirement since most individual data are not perfectly consistent with rationality. In our analysis, we use the preexisting level of rationality of the individuals being compared as a reference point. This construction allows us to define a *revealed different* relation which is complete[1] and symmetric yet not transitive. The normalization we use is motivated by the observation that the joint decisions of two perfectly rational agents cannot be more rational than the decisions taken separately.

The *revealed different* relation can be used to evaluate the level of heterogeneity in a set of data. For instance, Pattanaik and Xu (2000) use the complementary *similarity* relation to characterize freedom of choice across sets of alternatives. This relation allows the comparison of sets of alternatives in terms of how many different options they provide. The authors suggest using the smallest partition of the similarity relation as a measure of freedom of choice. Because the *similarity* relation is complementary to the *revealed different* relation, the methodology we use to elicit the amount of types in the data corresponds to Pattanaik and Xu's (2000) methodology.

Preference types are defined by two conditions. First, no two members of the same preference type are *revealed different* from each other. Second, for any two different preference types, there is at least one pair containing one person belonging to each type such that they are *revealed different* from each other. While the assignment of people to preference types is not unique, the cardinality of the smallest partition of people into preference types is always well-defined. We use this number as our measure of heterogeneity of preferences.[2]

[1]It is not clear *a priori* that the relation needs to be complete. For instance, two individuals might not be comparable because they face budgets for which rationality has no empirical content.

[2]This is the chromatic number of the undirected graph associated with the *revealed different*

In order to evaluate the empirical content of the *revealed different* relation, some assumptions about how to measure distance from rationality are needed. In this paper we use two measures of distance from rationality: Afriat's (1973) critical cost efficiency index (CCEI) and Echenique et al.'s (2011) money pump index (MPI). The CCEI is appealing because it is monotone in the number of decisions a subject makes. The CCEI never decreases as more decisions are observed. Since the CCEI is bounded between 0 and 1, this implies that the CCEI will converge to its true value as more data are collected, which allows us to assess the extent to which our results are due to sampling error or true differences in preferences. Monotonicity is not a common property among measures of distance from rationality. The MPI, which is not a monotone measure, will provide us with a natural contrast to the CCEI.

Our approach to gathering data is to use a random sample of the Dutch population whose preferences regarding risky assets were elicited using experimental methods in Choi et al. (2014). In that study, participants were presented with 25 randomly generated budgets containing two assets that obtained with equal probability. By defining the *revealed different* relation based on the Generalized Axiom of Revealed Preferences (GARP), we find that the 1,182 individuals in the sample can be partitioned into 131 different preference types. The preference types have an average of 8.03 (s.d. 3.88) members, with the largest type containing 20 subjects and the smallest type containing a single individual. These results are comparable using the MPI to measure distance from rationality instead of the CCEI. In this case, the number of preference types is 101.

As discussed above, the *revealed different* relation can be used to assess differences in heterogeneity across populations and to test the cost for measured heterogeneity of additional behavioral assumptions on preferences. Regarding differences across populations, we find evidence that men's preferences are more heterogeneous than women's preferences, i.e., there are preference types among men that are not found among women. Regarding the cost of behavioral assumptions for measured heterogeneity, we find that if we define rationality to include respecting first-order

relation.

stochastic dominance (FOSD), the number of preference types decreases by 35 percent (from 131 preference types to 85). This illustrates that behavioral assumptions not supported by the data can be costly in detecting heterogeneity of preferences.[3]

To measure the extent to which differences in individual preferences are due to differences in individual characteristics, we conduct a regression analysis of the *revealed different* relation as a function of the differences in demographic variables and individual fixed effects of the individuals being compared. We find that individual differences explain 33 percent of the variance of the *revealed different* relation. Including observable characteristics[4] of the pair adds little to the proportion of the variance explained. This suggests that differences in preferences are largely driven by factors other than differences in individual characteristics.

We are not the first to suggest using revealed preference analysis to identify heterogeneity in preferences. For example, Gross (1995) proposed testing for GARP using cross-sectional data as a way to test for the commonality of preferences. As another example, Heufer (2014) offered a nonparametric way to classify people based on their level of risk aversion.[5] The work of Crawford and Pendakur (2013), who propose finding the smallest partition of people into preference types such that each group's decisions do not violate GARP, is closest to our approach. Using observations of the demand for dairy in a cross-section of the Danish population, they find that there are up to 12 separate preference types in the population. An important difference between their approach and ours is that we explicitly address the possibility of behavior that is not perfectly rational. This is important since, in experimental settings, subjects routinely violate the weak axiom of revealed preferences (Sippel, 1997). It can also be shown that Crawford and Pendakur (2013) provide a lower bound for the degree of heterogeneity, because rational behavior is confounded with the choice sets' inability to reveal violations of rational behavior. Indeed, we use

[3]Expected utility, cumulative prospect theory and disappointment aversion all predict behavior consistent with FOSD in the experiments we analyze.

[4]We consider sex, education, and income level.

[5]Note that this measure is not complete. The method may return people who are not comparable in terms of risk aversion.

Crawford and Pendakur's (2013) method in the random sample of Dutch subjects by treating each separate budget as if it were a unique individual observation. This sample contains 29,550 (25×1182) different observations. We find that these 29,550 budgets can be grouped into only 11 types. Given that we observe that almost 40 percent of the population have a CCEI below 0.9, it is clear that some of the types must include budgets where the GARP cannot fail.[6]

Our research fits in with the larger literature on diversity (Weitzman, 1992; Pattanaik and Xu, 2000; Nehring and Puppe, 2002), which has tasked itself with axiomatizing the measurement of diversity. Weitzman (1992) and Nehring and Puppe (2002) take cardinal measure(s) of difference across individuals or alternatives as a starting point. While the *revealed different* relation can be used to construct a distance function,[7] its appeal lies in the fact that it is an ordinal summary of differences. We consider this to be a necessary first step in exploring diversity as revealed by individual choices.

The remainder of this paper is organized as follows. Section 2 describes the revealed difference approach to heterogeneity. Section 3 describes the data we use and presents our results. Section 4 discusses alternative measures of heterogeneity using revealed preference analysis. Section 5 concludes the paper.

2. Preliminaries

In this section we provide the definitions that are necessary for formally defining the *revealed different* relation and showing when it can be used to elicit differences in the preferences of agents. First, we define the *revealed different* relation in general and provide the conditions under which two subjects's being *revealed different* shows that they do have different preferences. Second, we specify the definitions of rationality that we use in the analysis. Third, we describe the measures we use for the

[6] Our approach takes advantage of a richer data set. Our method can be thought as a natural extension of Crawford and Pendakur's (2013) approach.

[7] The distance will satisfy positivity and symmetry, but it will fail the triangle inequality. It will also fail to be an ultrametric as defined in Weitzman (1992). The same issues arise if we use the change in distance from rationality to measure differences in preferences.

distance from rationality. Finally, we discuss how to classify people into preference types using the *revealed different* relation.

Let $E = \{x^i, p^i\}_{i=1}^N$ be a **consumption experiment**, where $x^i \in \mathbb{R}^T_+$ are consumption bundles chosen at prices $p^i \in \mathbb{R}^T_+$, and let $\mathcal{E}$ denote the set of all consumption experiments. Let us assume that every E is generated by a complete preference relation R over $\mathbb{R}^T_+$. Let $\rho : \mathcal{E} \to [0, 1]$ be a **distance from rationality**, where $\rho(E) = 0$ if and only if there is a complete and transitive relation that rationalizes E.[8] Let $D = \{E_1, E_2, \ldots, E_n\}, E_i \in \mathcal{E}$ be a set of consumption experiments. Whenever it is clear from the context, we will refer to a set of consumption experiments as a consumption experiment.

2.1. The Revealed Different Relation

In this section we define what we mean when we say that two consumption experiments are revealed different. Additionally, we show the conditions under which two consumption experiments will be revealed different only if they are generated by two different underlying preference relations.

Two consumption experiments are *revealed different* if jointly they are further from rationality than each of them taken separately:

$$RD_\varepsilon(E_r, E_s) = \begin{cases} 1, \text{ if } \rho(E_r \cup E_s) - \max\{\rho(E_r); \rho(E_s)\} > \varepsilon \\ 0, \ \textit{otherwise.} \end{cases}$$

Note that due to the partial observability of individual preferences, the distance from rationality for the consumption experiment and for the underlying preference relation can be different. Hence, we need a precision parameter (ε) to account for measurement error. Note as well that RD is an indicator function, i.e., it only shows whether consumption experiment E_r is different from consumption experiment E_s, but not how different people are from each other.

Observation 1. *For every $E_r, E_s \in \mathcal{E}$, $RD_\varepsilon(E_r, E_s)$ satisfies the following properties:*

[8] Generally, the assumption of a complete and transitive relation can be replaced by any other notion of rationality.

– $RD_\varepsilon(E_r, E_r) = 0$

– $RD_\varepsilon(E_r, E_s) \geq 0$

– $RD_\varepsilon(E_r, E_s) = RD_\varepsilon(E_s, E_r)$.

Observation 1 shows that RD_ε is a distance function. This is important since Weitzman (1992) shows that if differences between species can be represented by a distance function, it is always possible to construct a value function that represents the costs of losing a species. Note that in the current paper we do not address the issue of cardinal measures of diversity. However, Observation 1 shows that it is possible to use RD_ε to construct a cardinal measure of diversity.

Let us now show that if two consumption experiments are revealed different, then they cannot be generated by the same preference relation. In showing this, we prove that RD_ε correctly captures differences in preferences. Let $\mathcal{E}(R) \subseteq \mathcal{E}$ denote the space of consumption experiments generated by R.[9] Let E^n be a consumption experiment generated by R over n distinct budgets[10] (i.e., $|E^n| = n$). Let $\{E^n\}_{n=1}^\infty$ be a **sequence of consumption experiments** if $E^n \subset E^{n+1}$. We assume that $\rho(E)$ is defined over the metric space $([0,1], d_1)$, where $d_1(x,y) = |x-y|$ for every $x, y \in [0,1]$. Hence, we will assume convergence with respect to d_1. For every sequence of consumption experiments $\{E^n\}_{n=1}^\infty \subseteq \mathcal{E}(R)$, let $\rho(R) = \lim_{n\to\infty} \rho(E^n)$. The existence of this limit is not guaranteed. However, if we assume that distance from rationality, ρ, is **monotone**, i.e. it can only increase while adding new budgets to the consumption experiment, then the existence of the limit is guaranteed by the Monotone Convergence Theorem; this is because the sequence $\{\rho(E^n)\}_{n=1}^\infty$ is monotone and bounded.

Observation 2. *For every $E_r, E_s \in \mathcal{E}$, there is a ε such that $RD_\varepsilon(E_r, E_s) = 1$ only if E_r and E_s are generated by different preference relations.*

Proof. Let R_r be the complete preference relation. Let E_r be a finite consumption experiment generated by R_r, i.e. $E_r \in \mathcal{E}(R_r)$. Let $\varepsilon_r = \sup_{E_r \subset E \in \mathcal{E}(R_r)} |\rho(E) - \rho(E_r)|$.

[9] Recall that any consumption experiment includes a choice set and a chosen option.

[10] Henceforth budget means any collection of points from which a subject can choose.

Since ρ is bounded, there is a finite supremum $\varepsilon_r \leq 1$. Similarly we can define ε_s. Let $\varepsilon = \max\{\varepsilon_r, \varepsilon_s\}$. Assume on the contrary that E_r and E_s are generated by similar preference relation and $RD_\varepsilon(E_r, E_s) = 1$. This implies that $\rho(E_r \cup E_s) - \min\{\rho(E_r), \rho(E_s)\} > \varepsilon$. Without loss of generality assume that $\rho(E_r) \leq \rho(E_s)$, then $\rho(E_r \cup E_s) - \rho(E_r) > \varepsilon_r$. This implies a contradiction, because $\varepsilon_r \geq \rho(E_r \cup E_s) - \rho(E_r)$ (implied by $E_r \cup E_s \in \mathcal{E}(R_r)$). □

Observation 2 shows that it is possible to find a precision level ε such that two consumption experiments are revealed different only if they are generated by difference preference relations. Note that the observation establishes that preferences are revealed different by the data, only if they are generated by different preference relations; it does not claim the reverse. This implies that RD_ε tends to underestimate heterogeneity.

The following lemma allows us to use RD_0 instead of RD_ε:

Lemma 1. *If for every $E \in D$, $\rho(E) = \rho(R)$, then $RD_0(E_r, E_s) = 1$ only if E_r and E_s are generated by different preference relations.*

We have to assume that the observed distance from rationality is equal to the underlying preference relation's distance from rationality. This assumption implies that the consumption experiment is powerful enough to elicit the real distance from rationality.

2.2. Axioms of Revealed Preferences

Because the *revealed different* relation is defined with respect to a distance from rationality and distance from rationality in turn depends on the definition of rationality, we now state the axioms that specify the definitions of rationality that we use for the analysis.

For a given $e \in [0, 1]$, define the revealed preference relation as $x^i R(e) x^j$ if $p^i x^j \leq e p^i x^i$ and the strict preference relation as $x^i P(e) x^j$ if $p^i x^j < e p^i x^i$. Let $T(e)$ denote the transitive closure of $R(e)$.

Definition 1. *A consumption experiment $E = \{x^i, p^i\}_{i=1}^N$ satisfies the* ***Generalized Axiom of Revealed Preference (GARP)*** *if $x^i T(1) x^j$ implies not $x^j P(1) x^i$.*

Afriat (1967) showed that a consumption experiment satisfies GARP[11] if and only if there exists a nonsatiated and continuous utility function that represents $R(1)$. Hence, GARP is a necessary and sufficient condition for the existence of a complete and transitive preference relation that rationalizes the consumption experiment. This justifies using GARP as our main definition of rationality.

Previewing the analysis below, we will augment the definition of rationality to include the satisfaction of FOSD. When subjects choose between contingent assets (lotteries), each bundle can be described by a cumulative probability distribution function F_x. Then lottery x **FOSD** lottery y if $\forall z : F_y(z) \leq F_x(z)$. Note that FOSD is also a preference relation, therefore, we can define $R_F(e) = R(e) \cup \mathbf{FOSD}$, that is $xR_F(e)y$ if and only if $xR(e)y$ or $x\mathbf{FOSD}y$. Denote by $P_F(e)$ the strict part of $R_F(e)$ and by $T_F(e)$ the transitive closure of $R_F(e)$.

Definition 2. *A consumption experiment $E = \{x^i, p^i\}_{i=1}^N$ satisfies GARP and FOSD (**FGARP**) if $x^i T_F(1) x^j$ implies not $x^j P_F(1) x^i$.*

Satisfaction of FGARP is a necessary and sufficient condition for the existence of a complete and transitive preference relation, that is monotone with respect to FOSD and rationalizes the consumption experiment.[12]

2.3. Measures of Rationality

Recall that to define RD_ε, we need to specify the distance from rationality. We introduce two alternative measures of rationality: the CCEI (Afriat, 1973) and the MPI (Echenique et al., 2011).

Definition 3. ***The Critical Cost to Efficiency Index*** *(CCEI) is the maximum $e \in [0,1]$ such that $R(e)$ is consistent with GARP (FGARP).*

The value 1 - CCEI can be interpreted as the share of income a person is willing to waste to behave irrationally. Note that a subject who is rational has a CCEI of 1.

[11] GARP is introduced by Varian (1982), while Afriat (1967) refers to an equivalent condition called cyclical consistency.

[12] This follows from Proposition 10 in Nishimura et al. (2017). The statement of FGARP we use is equivalent to the one in Polisson et al. (2015).

Importantly, Afriat (1973) showed that e is the CCEI of a consumption experiment if and only if there exists a nonsatiated and continuous utility function that represents $R(e)$. In this setup, $\rho = w = 1 - e$, and therefore larger values of w mean greater distances from rationality. The CCEI is monotone in the sense that it cannot increase as budgets are added to the consumption experiment. Hence, the corresponding distance from rationality (w) is monotone (but increasing) as well.

The second distance function we use is the MPI. Let $x^{i_1}, x^{i_2}, \ldots, x^{i_k}$ be the sequence that violates GARP, i.e., $x^{i_j} R(1) x^{i_{j+1}}$ for any $j \in \{1, \ldots, k-1\}$ and $x^{i_k} P(1) x^{i_1}$. The *money pump* of the sequence is defined as $mp = \sum_{j=1:k} p^{i_j}(x^{i_j} - x^{i_{j+1}})$, where $i_{k+1} = i_1$. The *relative money pump* is defined as the money pump divided by the total income in the chain ($rmp = \frac{\sum_{j=1:k} p^{i_j}(x^{i_j} - x^{i_{j+1}})}{\sum_{j=1:k} p^{i_j} x^{i_j}}$).

Definition 4. *The* ***Money Pump Index*** *(MPI) is the median (or the mean) relative money pump over all sequences that violate GARP.*

Note that the MPI (unlike the CCEI) is not defined for the case of FGARP. Indeed, it is not clear what the definition of the money pump is in this case. Moreover, the MPI is not a monotone measure, and therefore the existence of a limit analogous to that in Lemma 2 cannot be guaranteed.

There are two measures of distance from rationality that we do not consider in our analysis. Apesteguia and Ballester's (2015) *swaps index* can only be defined over a finite set of alternatives and is therefore not feasible in our context. Houtman and Maks (1985) proposed finding the maximum subset of choices that is consistent with a given behavioral axiom. However, this index does not naturally lend itself to comparison across individuals.[13]

2.4. How should heterogeneity be measured?

We now discuss how to measure heterogeneity using the *revealed different* relation.

Pattanaik and Xu (2000) proposed a measure of freedom of choice based on the "similarity" of the alternatives available in different choice sets. The *revealed*

[13] This index can be normalized by the number of budgets available to a subject. Note that while the number of preference types consistent with a consumption experiment increases in the number of budgets, the normalized index is not necessarily monotone.

similar and *revealed different* relations are complements of each other. According to Pattanaik and Xu (2000), the diversity of a set is determined by the smallest partition of alternatives into homogeneous disjoint subsets.[14] A set is said to be homogeneous if none of its elements are different from each other (or, alternatively, any two elements of the set are similar). Pattanaik and Xu's (2000) definition of diversity has a natural appeal for classifying people into preference types.

We remark that given this definition, there might be two individuals who are not *revealed different* from each other but are nevertheless assigned to different preference types. This can occur if one of these individuals, a, is not *revealed different* from any other individual, while the other, b, is *revealed different* from a third individual, c. In this case, b and c belong to different preference types, and a can freely be assigned to either b's preference type or c's preference type. This makes it clear that measurement of the heterogeneity of preferences is akin to the well-studied problem of graph coloring in graph theory. We will now provide the necessary definitions to show the equivalence of the two problems and to explain how we group preferences into types.

A graph is a tuple $G = (V, E)$ in which V is the set of vertices and $E \subseteq \{V, V\}$, where $\{V, V\}$ is the set of all 2-element subsets of V. Let the vertexes in G be the individuals i, and let $\{i, j\} \in E$ if and only if $RD_\varepsilon(i, j) = 1$. Then we have

Definition 5 (Correct Graph Coloring)**.** *Two vertexes of graph G (individuals) cannot have the same color (belong to the same type) if they are adjacent (revealed different).*

The chromatic number of a graph is the minimum number of colors required for a correct coloring. Graph coloring is equivalent to the problem of covering a graph by the minimum number of independent sets.[15] Therefore, the coloring of a graph is an

[14] This measure is the only one satisfying the following properties. First, all singleton sets have the same level of diversity. Second, if adding a new element to a homogeneous set does not change the level of diversity of the set, then the element is similar to all elements of the set. Third, if set A is more diverse than set B, then A remains more diverse than B if an x that is not an element of A is added to both sets.

[15] An independent set is a subset of V such that none of the vertexes in the independent set are connected.

appropriate technique to reveal the number of types borne by the *revealed different* relation. Note that, by definition, the correct coloring will correspond to a division of preferences into groups as suggested by Pattanaik and Xu (2000). Importantly, the chromatic number is unique, and therefore the measure of diversity is always well-defined. We present the algorithms to compute the chromatic number in the appendix.

It is important to note that while the chromatic number is unique, the actual coloring of a graph need not be. This can be shown using a simple example. Assume there is a vertex that is not connected to any other vertex (an isolated vertex); then this can be colored with an arbitrary color. Hence an isolated vertex can belong to any type, and therefore, coloring is not unique. Thus, while we can answer the question of how many preference types there are, we cannot determine the composition of the types. This is also an important methodological point since, in general, we will not be able to use preferences types for applied analysis.

3. Results

In this section, we provide an empirical analysis of heterogeneity using the *revealed different* relation. We first discuss the data we use. Second, we determine ε for RD_ε according to Lemma 2 and Lemma 1. Third, we use RD_ε to determine the number of preference types in the data.

3.1. Data

The analysis is based on the following data sets:

– Choi et al. (2007). The participants are 93 undergraduates and staff members from the University of California – Berkeley. Each subject faced 50 different budgets with two risky assets that obtained with probabilities $\frac{1}{2}$, $\frac{1}{3}$, or $\frac{2}{3}$. The prices of the assets were assigned randomly.

– Choi et al. (2014). The participants are a random sample of 1182 Dutch adults. Each participant faced 25 budgets with two risky assets that obtained with equal probability. The prices of the assets were assigned randomly.

3.2. Testing Assumptions

For the *revealed different* relation to elicit real differences in preferences we have to make two assumptions: (i) distance from rationality converges to some "true" distance to rationality, and (ii) observed distances from rationality are "close enough" to the "real" distance to rationality.

While assumption (i) has to be taken for granted, assumption (ii) can be tested. Recall that the data in Choi et al. (2014) has 25 budgets per subject, and for the unions of two consumption experiments we obtain 50 budgets per pair. Therefore, using Choi et al.'s (2007) data (which has 50 budgets per subject), we can test the perturbations in the distance from rationality that are caused by increasing the amount of budgets per person in order to determine the precision parameter (ε).

This is necessary to control for possible false positive results in which two consumption experiments generated by the same underlying preference relation appear to be *revealed different*.

For this purpose, we bootstrap 100 samples of 25 budgets for each subject and measure the distance from rationality using these subexperiments. Then we compare the distances from rationality obtained from the bootstrapped subexperiments using 25 budgets to the distances from rationality obtained from the consumption experiment with 50 budgets. Since the consumption experiments are generated using the data from the same subjects, they are generated by the same underlying preference relation. Hence, the distance from rationality in the experiment with 50 budgets minus the distance from rationality in the experiment with 25 budgets characterizes the value of ε that is needed to eliminate the possibility of a subject being *revealed different* from him- or herself.

Figure 1 shows the mean changes in distance from rationality for the subjects for both the CCEI and the median MPI. We construct RD_ε relations with ε equal to 0, to the median change in the distance from rationality, and to the 95^{th} percentile of the empirical distribution of the change in the distance from rationality.

Table 1 shows the values of ε we use in our further analysis. To simplify notation, we use RD_0 to denote the relation that does not allow for error, RD_M to denote the relation that uses the median change in distance from rationality, and RD_{95CI} to denote the relation that uses the 95^{th} percentile of the empirical distribution of the change in distance from rationality.

3.3. The Revealed Different Relation

This section presents basic information about the *revealed different* relation. The richness of the *revealed different* relation can be appreciated through Figure 2. The top panel shows the density of the relation for a random sample of 100 subjects. The bottom panel shows the same sample after coloring.

Table 2 presents the basic statistics for the *revealed different* relation. The average number of people a subject is different from is 695 (s.d. 237), with a minimum of 20 and a maximum of 1176.

3.4. Revealing Differences

We want to answer the following questions with our analysis: first, whether heterogeneity in terms of preferences is present in the data, and how many types there are; second, whether heterogeneity can be explained by some observable characteristics (demographic data) or is mostly idiosyncratic; and third, how our method should be used to reveal differences across groups – in particular, differences by gender. Note that the method we construct allows us not only to answer the question of whether groups have different types of preferences, but also to compare the groups on the basis of how heterogenous they are.

As mentioned above, we use the data from Choi et al. (2014) for this purpose. Table 3 provides descriptive statistics for the sample. The sample has a majority of men, people who are older than 50 years of age, people who are salaried, and people who are in a stable relationship.

3.4.1. Are Preferences Different?

Figure 3 shows the number of different preference types as a function of the population size. These estimates are obtained by random sampling subsets of individuals and calculating the chromatic numbers for those sets based on the RD_ε relation. The RD_ε relation is defined in terms of the CCEI and the MPI. Panel (a) shows the results using the MPI, and Panel (b) shows the results using the CCEI. Solid lines are the (mean) number of types for the corresponding group, and dashed lines show the 95% confidence intervals for the number of colors given a population size.[16] In both cases, the higher we set ε, the fewer the types that are revealed.

We observe that the (mean) number of preference types in a population of 1000 individuals is 90 according to the CCEI and 70 according to the MPI. In the population at large (1,182 people), we find that there are 131 types if we use the CCEI to define RD_0 and there are 101 types if we use the MPI to define RD_0. If we restrict the definition of revealed difference to RD_M, then the number of colors is 90 for the CCEI and 101 for the MPI. If we restrict the definition of revealed difference to RD_{95CI}, then the number of colors is 35 for the CCEI and 23 for the MPI. This shows that there is a large degree of heterogeneity in the data regardless of the measure used for distance from rationality. Note that there are fewer types according to the MPI compared to the CCEI. This can be explained by the fact that the MPI is not a monotone measure of distance from rationality.

Figure 3 suggests that the total number of groups in the population is finite. We do not observe a linear relationship between population size and the number of groups (colors). Given that the minimum size of a group is one, we would expect that the number of groups eventually converges. Note, however, that the answer to the question of how many types there are is complicated by measurement error. Figure 3 demonstrates that making the *revealed different* relation stricter leads to a reduction in the number of types in the population. One implication is that heterogeneity is likely to be underestimated due to the fact that all consumption experiments are finite and therefore unable to detect some differences (see Observation 2). More importantly, erroneous assumptions regarding rationality are likely to reduce the measured level of heterogeneity. We will discuss this issue more extensively below.

Figure 1 shows a significant portion of the population undergoing very large changes in their CCEIs as we move from 25 to 50 budgets. This suggests that estimates of ε based on means will likely be biased upwards. That is, the ε needed to statistically reject the null hypothesis of no difference in preferences is likely over-

[16]We use 400 bootstrapped samples per population size.

estimated. Figure 3 illustrates this. The estimated number of groups based on the median change in CCEI is much larger than the estimates based on the 95% percentile of the change in CCEI. However, using this very strict threshold, we still find a large number of types in the population. We can then conclude that there is a high degree of heterogeneity in the data.

3.4.2. What Can Explain Heterogeneity?

To investigate whether the observed level of heterogeneity in the data is due to observed individual characteristics, we conduct regression analysis on the *revealed different* relation on the characteristics of the individuals in a pair. Table 4 presents a linear probability model on a variable that equals 1 if individuals i and j are *revealed different* and 0 if they are not. The main statistic of interest is the portion of the variance of the *revealed different* relation that can be explained by differences in individual characteristics. Given that there are 1,182 subjects, we have 1,397,124 such comparisons. All the regressions cluster errors at the i level and all regressions include fixed effects on the i and j subject.

The first four columns in Table 4 show estimates for the case in which the *revealed different* relation is calculated using Afriat's CCEI. The first column shows a regression in which only fixed effects for each member of the pair are included. Fixed effects capture any individual differences across individuals. They capture about 33 percent of the variance in the *revealed different* relation. This means that two-thirds of the variance in the relation cannot be explained by individual differences. The next regressions then include characteristics of the pair. We observe in the second column that adding characteristics of the pair adds little to the proportion of the variance explained. This finding holds regardless of how strict the definition of the *revealed different* relation is. This suggests that there other factors that are particular to the pair that explain the differences in preferences. The fifth and sixth column reproduce the analysis using the MPI. We find similar results using this definition. The last column presents the results when satisfaction of FOSD is added to GARP. We observe that the proportion of the variance explained decreases as the definition of rationality is made stricter.

Table 4 allows us to take a look at differences in preferences across and within subpopulations. We observe that a pair of males are more likely to be different than a mixed gender pair and that a pair of females are less likely to be different than a mixed gender pair. This result is consistent regardless of how we calculate the *revealed different* relation. Regarding age differences, we do not observe a consistent pattern across definitions and assumptions. Finally, differences across other characteristics are relatively small and not always consistent.

3.4.3. Which Groups Are Different?

We have already shown evidence regarding the amount of heterogeneity within and between subpopulations. We now investigate whether there are differences across groups and whether there are groups with more diverse preferences.

To answer the first question, we check whether the number of groups in the population as a whole is larger than the number of groups in its constituent subpopulations. We do this because for an additional group to appear in the population as a whole, there must exist a member of one of the subpopulations who is *revealed different* than some person in each of the groups of the other subpopulations. That is, there must be a person in one subpopulation who cannot be included in any of the existing groups in other subpopulations. To answer the second question, we simply look at the number of groups of different subpopulations. Populations that are more heterogeneous are expected to be divided into more groups.

Since the number of groups tends to be larger in larger populations, we need to study the distribution of types across populations holding constant the sample size of individuals used in the estimation. This can be done by taking random samples of equal size from the population of interest (males, females, mixed gender). We present results for samples sizes ranging from 50 to 500.

Figure 4 shows the number of types for males, females, and the population at large. The graphs are presented by the measure of rationality we use (MPI or CCEI) and the value of ε. Solid lines are the (mean) number of types for the corresponding group and dashed lines show the 95% confidence intervals for the number of types given a population size.[17]

Figure 4 shows two main findings. First, males and females have different preferences. We can conclude this from the fact that the number of types in the population at large is larger than the number of types in the two mono-gender samples (54.6 for the mixed gender group, 48.5 for the males-only group, and 45.8 for the females-only group). In other words, there are types of preferences that are specific to males and females. Second, males are more heterogeneous than females. We can infer this from the fact that the number of types in the males-only group is greater than the number of types in the females-only group (48.5 for the males-only group and 45.8 for the females-only group).[18]

We verify the existence of gender differences in risk preferences by comparing their propensity to deviate from asset allocations that give equal payments in all of the states of the world. Figure 5 shows the average absolute deviation from equal payments for male and female subjects.[19] The average deviation of males is significantly different from that of females (0.1979 vs. 0.1763, t-test = 3.1138, p-value = 0.0019). This comparison is valid since budgets were randomly assigned to subjects. Any systematic difference in asset allocation must be due to difference in risk attitudes. While reassuring, this exercise also illustrates an advantage of our method. The observed difference in asset allocations might be due to the existence of different preference types among male and female subjects or to the fact that some preference types are over-represented in these populations. Our method shows that there are different preference types in these populations.

[17] We use 400 bootstrapped samples per population size.

[18] The numbers presented in this paragraph are for RD_0 based on the CCEI at the group size 500. However, the difference is robust to different values of ε and to using the MPI instead of the CCEI.

[19] The formula is $\frac{1}{K}\sum_{i=1}^{i=K}|\frac{x}{x+y}-\frac{1}{2}|$ where x and y is the allocations in each one of the assets and K is the number of budgets faced by the subject.

4. Discussion

4.1. Additional axioms for preferences

In the context of decisions regarding Arrow-Debreu securities, the requirement that subjects satisfy FOSD is desirable since it is implied by both expected utility and cumulative prospect theory. As we mentioned earlier, we can use alternative definitions of rationality. In particular, we can modify the *revealed different* relation to consider consistency with FOSD (FGARP).

We now show that adding assumptions that are not supported by the data can produce a severe underestimation of the heterogeneity of preferences. This is because when people are further from rationality, it is more difficult to detect differences between them. Intuitively, adding assumptions not supported by the data increases the number of patterns of behavior that a person can have without increasing his or her distance from rationality.

Figure 6 shows the number of types according to GARP and FGARP for the random samples of size ranging from 50 to 1000. We observe that the number of types according to FGARP is about half of the number of types according to GARP. This is consistent with the fact that the CCEI associated with FGARP is smaller than the CCEI associated with GARP (the mean CCEIs are .81 vs .88). Equivalently, the distance from rationality with respect to FGARP is larger than with respect to GARP.

This exercise shows that the measurement of the heterogeneity of preferences is intimately related to the assumptions made about behavior. Wrong, parametric or nonparametric assumptions about preferences will lead to an underestimation of the number of preference types.

4.2. Alternative ways to group people

Recently, Crawford and Pendakur (2013) have proposed using revealed preference methods to determine the number of different preference types in a population. The intuition behind their approach is similar to the intuition behind ours, with the difference that they use cross-sectional data with one observation per person. A potential shortcoming of using cross-sectional data to group people into preference types is that the method cannot distinguish between being mutually consistent and not being comparable. That is, a person whose budget sets are contained in the budget sets of another person cannot be revealed different from this person without invoking stronger preference assumptions such as homotheticity. This could generate a serious downward bias on the amount of heterogeneity whenever the number of goods a person considers is small.

We test whether Crawford and Pendakur's (2013) estimate of heterogeneity might be biased downwards when the set of goods is small. To do this, we consider each budget in our data as if it were a unique individual observation and calculate the number of types using Crawford and Pendakur's (2013) method.[20] We find that a total sample of $29,550$ (25×1182) can be classified into 11 different types according to their method. Given that a significant portion of the subjects in the sample were themselves not close to rationality, this suggests that the method might underestimate the level of heterogeneity in the data. We should remark, however, that as more goods are considered, the bias should diminish. However, the method requires restricting the analysis to situations that make decisions across individuals comparable.

5. Conclusions

We use the revealed preference approach to analyze the heterogeneity of preferences. We define two individuals to be *revealed different* if their joint decisions are further from rationality than each of their individual decisions taken separately. The method is useful for analyzing the structure of preferences and assessing the effect of stricter definitions of rationality on the measured heterogeneity of preferences. Using a random sample of Dutch adults, we find that men possess more heterogeneous preferences than women. We also find that stricter definitions of rationality decrease the level of heterogeneity observed in the data. This suggests there is a trade-off

[20]We improve the method in Crawford and Pendakur (2013) by providing an alternative algorithm to calculate the number of types. Our algorithm implements the original procedure for the special case of two goods and linear budgets. It exploits the fact that in this case, every violation of GARP is also a violation of WARP.

between precision in estimates and revealed heterogeneity in preferences in the parametric approach. We show that our results are robust to using alternative definitions of distance from rationality. Our method shows that revealed preference analysis can be a useful tool for analyzing differences in preferences across individuals.

Appendix

Appendix A: Graph Coloring

For an arbitrary set A, let $\{A, A\} = \{\{a, b\} : \ a, b \in A\}$.

Definition 6. *A **graph** is a tuple $G = (V, E)$, where V is a (finite) set of vertexes and $E \subseteq \{V, V\}$ is a set of edges.*

A vertex $u \in V(G)$ is said to be **adjacent** to a vertex $v \in V(G)$ if there is an edge $e = \{u, v\} \in E(G)$.

Definition 7. *For a graph $G = (V, E)$, a **coloring** is a function $c : V \to \mathbb{N}$ such that for any $e \in E$, if $e = \{u, v\}$ for some $u, v \in V$ then $c(v) \neq c(u)$.*

So a function is a coloring if it assigns different values to adjacent vertices.

Definition 8. *A graph $G = (V, E)$ is said to be k-**colorable** if there is a coloring function $c : V \to \{1, 2, ..., k\}$.*

Definition 9. *A number $\chi(G)$ is said to be a **chromatic number** of a graph $G = (V, E)$ if G is $\chi(G)$-colorable and there is no $k < \chi(G)$ s.t. G is k-colorable.*

Remark: Note that the chromatic number of a graph is well-defined, while the coloring is not necessarily unique. Hence, a graph can have an entire (finite) class of coloring functions $c^i : V(G) \to \{1, ..., \chi(G)\}$. Let us consider a simple $G = (\{u_1, u_2, u_3\}, \{\{u_1, u_3\}\})$. It is clear that $\chi(G) = 2$, but there are 2 correct colorings: $V_1 = \{u_1, u_2\}$ and $V_2 = \{u_3\}$, and $U_1 = \{u_1\}$ and $U_2 = \{u_2, u_3\}$. Here we use V_i and U_i to denote the partitions that correspond to the correct coloring of G.

Finding an exact chromatic number is not a feasible problem (it is NP-hard); thus we have to use an approximation algorithm called *Greedy Graph Coloring*. For this we need to define the degree of a vertex: for any $u \in V(G)$, $d_G(u) = |\{e \in E : \ u \in e\}|$, i.e., it is the number of edges that have vertex u as an end vertex.

Algorithm 1 Greedy Graph Coloring

function GGC(V, E)
    $I = V$
    $\delta = 0$
    while $I \neq \emptyset$ **do**
        Find $u_i \in I$ s.t. $d_G(u_i) = \max_{v \in I}(d_G(v))$
        $U_i = \{u_i\} \cup \{v \in I : \ \{u, v\} \notin E(G)\}$
        $I = I \setminus U_i$
        $\delta = \delta + 1$
        $c(U_i) = \delta$
    end while
    return δ, c
end function

Theorem 1 (See, e.g., Agnarsson and Greenlaw (2006))**.** *Let* $G = (V, E)$ *be a graph and let* $\delta = GGC(V, E)$*. Then* $\chi(G) \leq \delta$*.*

Let us also note that δ is a tight bound, i.e., we cannot find a smaller upper bound without imposing additional restrictions on the structure of the graph. Let K_n denote $(V, \{V, V\})$, $|V| = n$, the complete graph on n vertexes, i.e., any two vertexes are adjacent. $\chi(K_n) = n$, since any two vertexes are adjacent, and $\delta = n$ as well. Moreover, if $\delta \leq 3$, then $\chi(G) = \delta$.

Appendix B: Additional Robustness Checks

In this appendix, we show results with respect to an additional level for ε. As an alternative way to compute the 95% confidence interval, we can use the following logic. We can determine the value of ε such that according to the t-test, the hypothesis of the equality of the distance from rationality for the consumption experiment with 25 budgets and the distance from rationality for the consumption experiment with 50 budgets cannot be rejected at $p < 0.05$. In this case, we obtain $\varepsilon = .005$ using the CCEI as the distance from rationality and $\varepsilon = .0011$ using the MPI as the distance from rationality. The total number of types in this case is 108 using the CCEI as the distance from rationality and 96 using the MPI as the distance from rationality.

Figure 7 presents the results of the analysis done for gender differences. Consistent with the major part of our analysis, we obtain the result that males and females are different in terms of preferences and that males have more different types of preferences than females.

References

Afriat, S.N., 1967. The construction of utility functions from expenditure data. International economic review 8, 67–77.

Afriat, S.N., 1973. On a system of inequalities in demand analysis: an extension of the classical method. International economic review , 460–472.

Agnarsson, G., Greenlaw, R., 2006. Graph Theory: Modeling, applications, and algorithms. Prentice-Hall, Inc.

Apesteguia, J., Ballester, M.A., 2015. A measure of rationality and welfare. Journal of Political Economy 123, 1278–1310.

Choi, S., Fisman, R., Gale, D., Kariv, S., 2007. Consistency and heterogeneity of individual behavior under uncertainty. American Economic Review 97, 1921–1938.

Choi, S., Kariv, S., Mueller, W., Silverman, D., 2014. Who Is (More) Rational? American Economic Review 104, 1518–1550.

Crawford, I., Pendakur, K., 2013. How many types are there? Economic Journal 123, 77–95.

Echenique, F., Lee, S., Shum, M., 2011. The Money Pump as a Measure of Revealed Preference Violations. Journal of Political Economy 119, 1201–1223.

Gross, J., 1995. Testing data for consistency with revealed preference. Review of Economics and Statistics 77, 701–710.

Heufer, J., 2014. Nonparametric comparative revealed risk aversion. Journal of Economic Theory 153, 569–616.

Houtman, M., Maks, J., 1985. Determining all maximal data subsets consistent with revealed preference. Kwantitatieve methoden 19, 89–104.

Nehring, K., Puppe, C., 2002. A theory of diversity. Econometrica 70, 1155–1198.

Nishimura, H., Ok, E.A., Quah, J.K.H., 2017. A comprehensive approach to revealed preference theory. American Economic Review forthcoming.

Pattanaik, P.K., Xu, Y., 2000. On diversity and freedom of choice. Mathematical Social Sciences 40, 123–130.

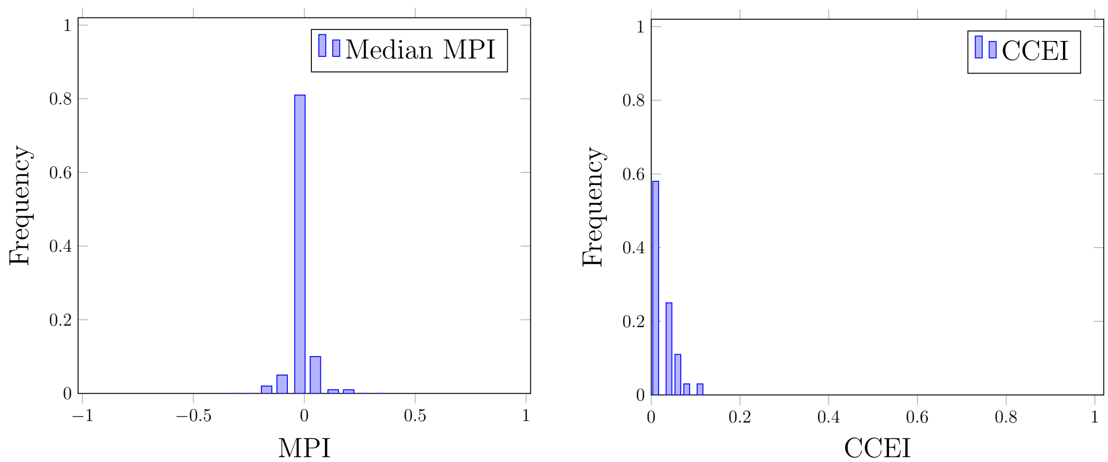
Polisson, M., Quah, J.K.H., Renou, L., 2015. Revealed preferences over risk and uncertainty. Technical Report. IFS Working Papers.

Sippel, R., 1997. Experiment on the pure theory of consumer's behaviour. Economic Journal 107, 1431–1444.

Varian, H.R., 1982. The nonparametric approach to demand analysis. Econometrica: Journal of the Econometric Society , 945–973.

Weitzman, M.L., 1992. On diversity. Quarterly Journal of Economics 107, 363–405.

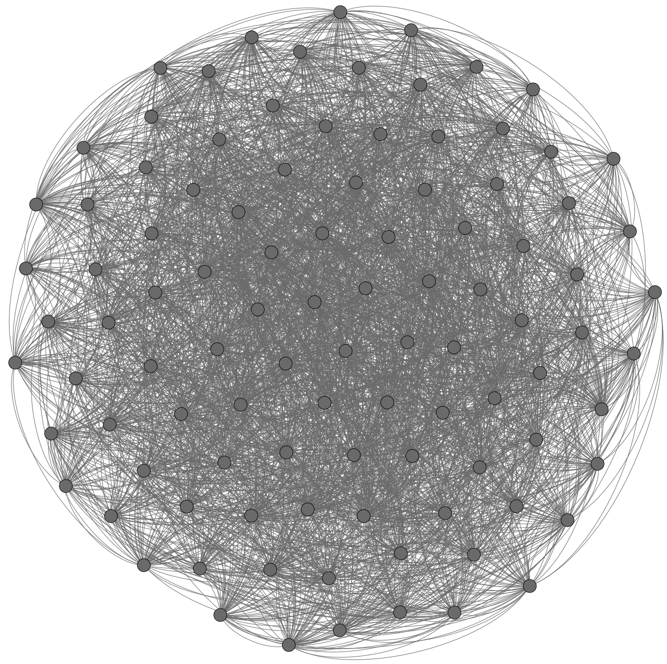
Figures



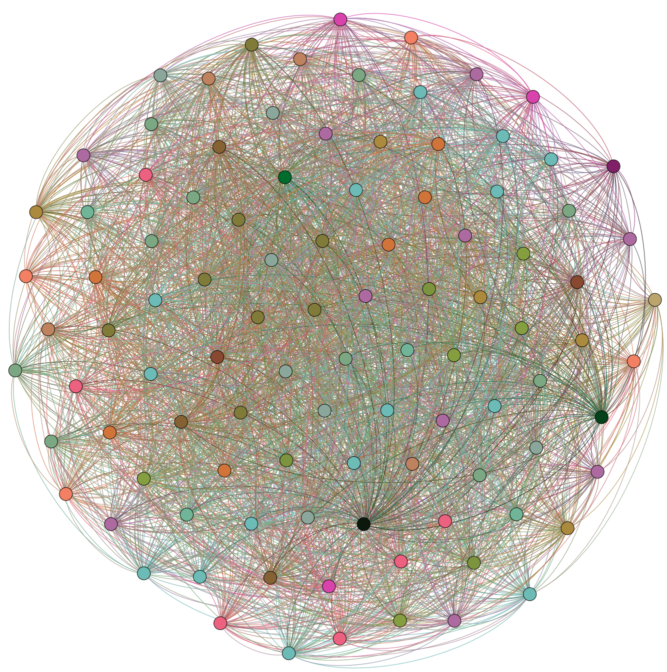
(a) Money Pump Index

(b) Critical Cost Efficiency Index

Figure 1: CHANGE IN DISTANCE FROM RATIONALITY: CHANGING THE NUMBER OF BUDGETS FROM 25 TO 50.

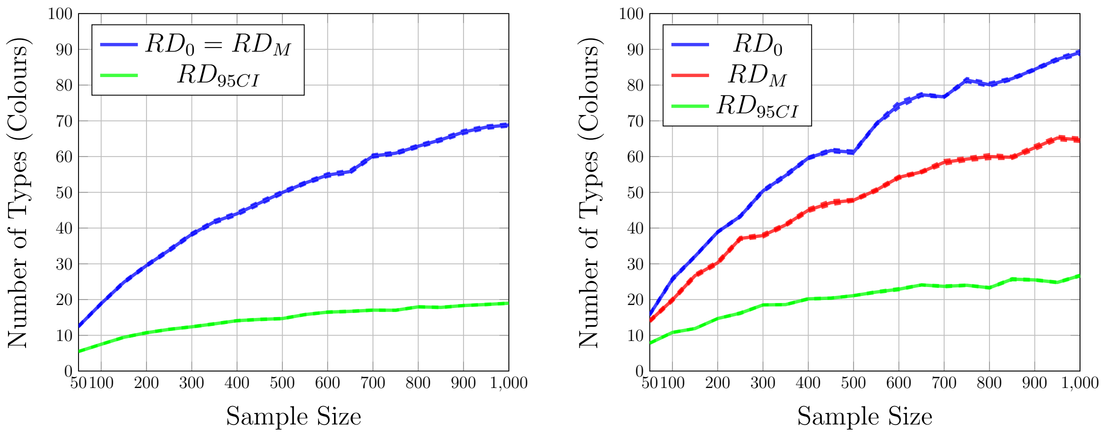


Uncolored



Colored

Figure 2: GRAPH OF THE *revealed different* RELATION



(a) Money Pump Index

(b) Critical Cost Efficiency Index

Figure 3: NUMBER OF TYPES

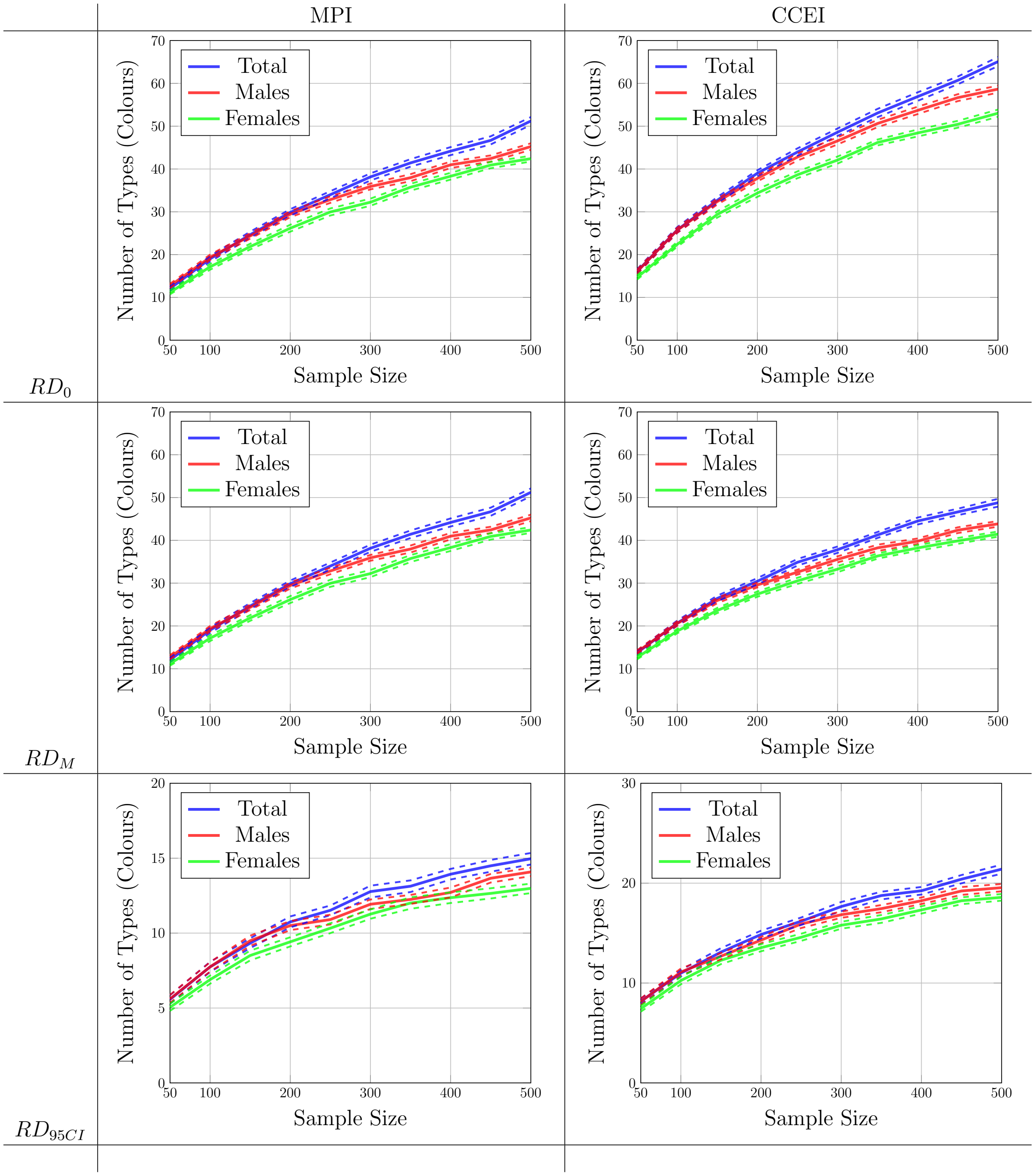


Figure 4: Number of Types for Males and Females

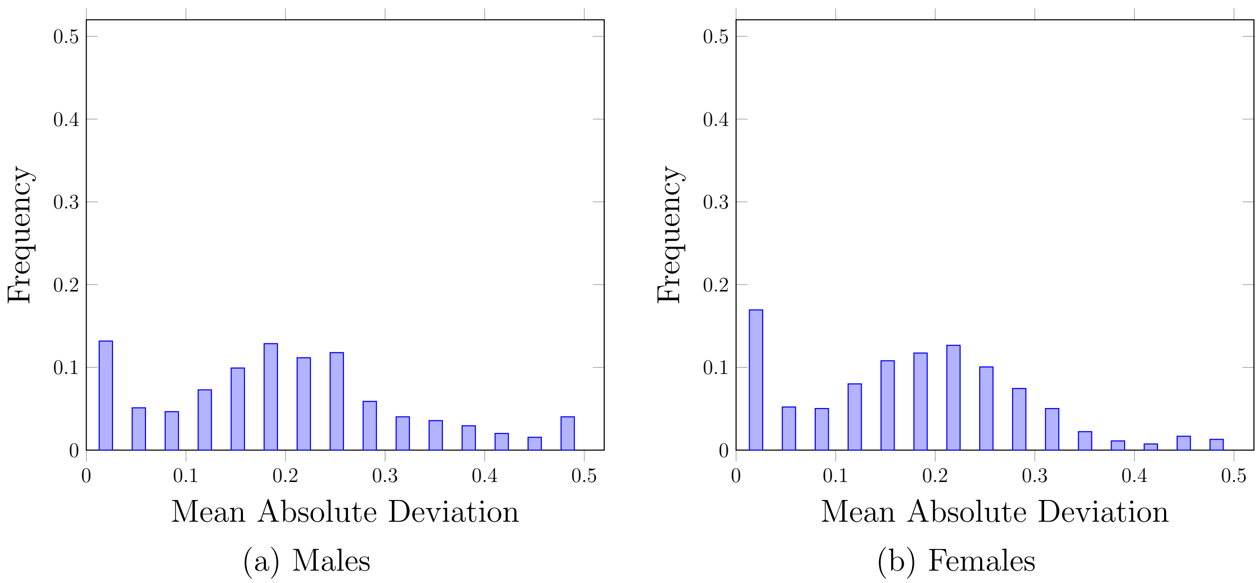


Figure 5: Distributions of Mean Absolute Deviations from 50-50 Allocation.

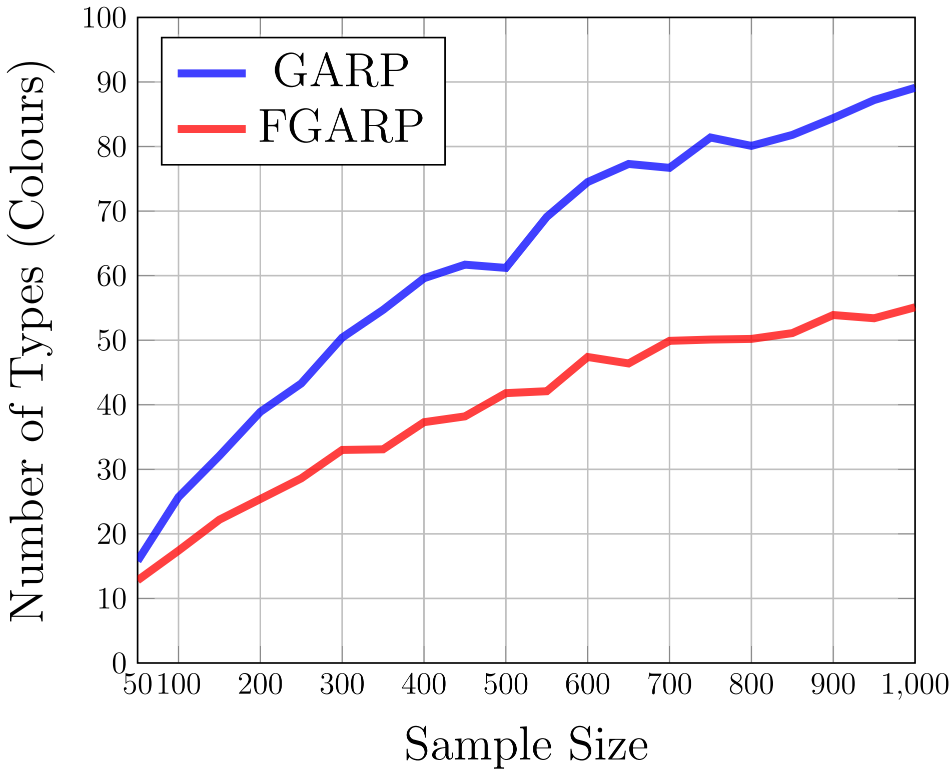
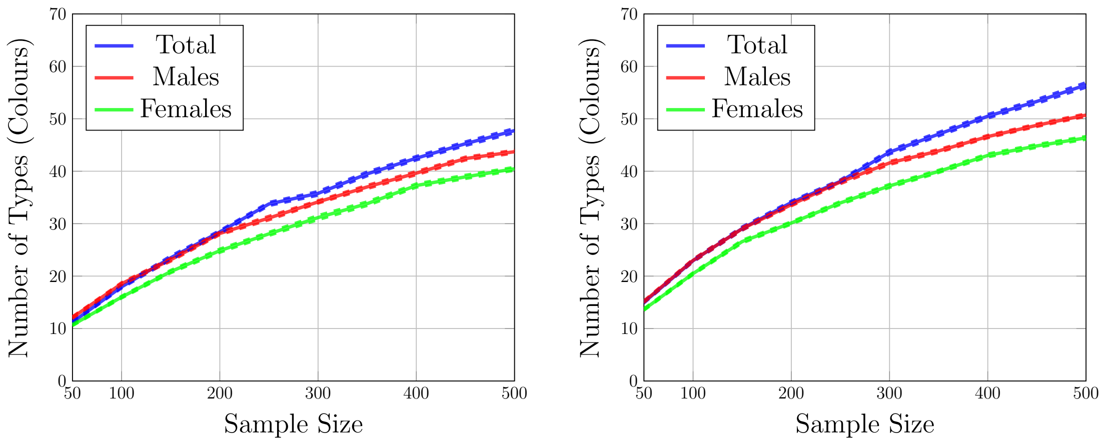


Figure 6: Number of Types: GARP and FGARP



(a) MPI (b) CCEI

Figure 7: Number of Types for Males and Females

Tables

	Median	95^{th} percentile
Median MPI	0	.04
CCEI	.01	.07

Table 1: RD_ε for Median and 95th percentile in Distance From Rationality

GARP CCEI				
Type of RD	Mean	SD	Min	Max
RD_0	0.5879	0.2006	0.1690	0.9949
RD_M	0.5402	0.2014	0.0093	0.9873
RD_{95CI}	0.3283	0.1988	0.0000	0.9501
GARP MPI				
Type of RD	Mean	SD	Min	Max
$RD_0 = RD_M$	0.3978	0.2216	0.0000	0.9679
RD_{95CI}	0.1472	0.1620	0.0000	0.8672
FGARP CCEI				
Type of RD	Mean	SD	Min	Max
RD_0	0.4761	0.1893	0.0034	0.9239

Table 2: PROBABILITY OF BEING *revealed different*

	Mean	Std. Dev.
Females	0.454	0.498
Age :		
16-34	0.185	0.389
35-49	0.261	0.440
50-64	0.356	0.479
65+	0.197	0.397
Education:		
Low	0.336	0.472
Medium	0.297	0.457
High	0.364	0.481
Occupation:		
Paid work	0.531	0.499
Housework	0.116	0.320
Retired	0.209	0.407
Others	0.144	0.351
Household Composition:		
Partner	0.809	0.393
Number of children	0.843	1.126
Income :		
0-2499	0.228	0.419
2500-3499	0.255	0.436
3500-4999	0.292	0.455
5000+	0.225	

Table 3: Descriptive Statistics for our Sample

		GARP×CCEI			GARP×MPI		FGARP×
VARIABLES	RD_0	RD_0	RD_M	RD_{95CI}	$RD_0 = RD_M$	RD_{95CI}	CCEI
Female$_i$-Female$_j$		-0.580***	-0.274***	-0.340***	-0.080***	-0.094***	-0.575***
		[0.015]	[0.013]	[0.014]	[0.015]	[0.009]	[0.017]
Male$_i$-Male$_j$		0.572***	0.271***	0.337***	0.073***	0.087***	0.571***
		[0.015]	[0.013]	[0.014]	[0.015]	[0.009]	[0.017]
Age$_i$ $>$ 49 & Age$_j$ $>$ 49		0.540***	0.238***	0.325***	-0.167***	-0.005	0.045***
		[0.015]	[0.012]	[0.014]	[0.012]	[0.005]	[0.012]
Age$_i$ $\leq$ 49 & Age$_j$ $\leq$ 49		-0.509***	-0.226***	-0.310***	0.191***	0.022***	-0.006
		[0.015]	[0.012]	[0.014]	[0.012]	[0.005]	[0.012]
One-level diff. in educ.+		-0.005**	0.000	-0.002	-0.002	-0.001	-0.006**
		[0.002]	[0.001]	[0.002]	[0.002]	[0.001]	[0.002]
Two-level diff. in educ.+		-0.015***	-0.001	-0.005**	-0.005	-0.002	-0.016***
		[0.004]	[0.002]	[0.003]	[0.003]	[0.002]	[0.004]
One-level diff. in income+		0.002	0.001	0.002	0.000	-0.001	0.002
		[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]
Two-level diff. in income+		0.003	0.002	0.003	0.001	-0.002	0.003
		[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Three-level diff. in income+		0.000	-0.001	0.000	-0.007	-0.006*	-0.000
		[0.004]	[0.003]	[0.003]	[0.004]	[0.003]	[0.004]
Constant	0.712***	0.755***	0.271***	0.381***	-0.066***	-0.019***	0.396***
	[0.014]	[0.012]	[0.010]	[0.012]	[0.013]	[0.007]	[0.011]
i's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
j's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1,397,124	1,397,124	1,397,124	1,397,124	1,397,124	1,397,124	1,397,124
R-squared	0.332	0.332	0.370	0.358	0.410	0.418	0.287

Robust standard errors in brackets, clustered at the individual level

*** p<0.01, ** p<0.05, * p<0.10

+ We use variables as described in Table 3. For instance, "three-level difference in income" equals 1 if one subject has an income in the bracket [0-2499] and the other subject has an income in the bracket [5000+]

Table 4: EXPLAINING DIFFERENCES: $y_{i,j} = 1$ IF i IS REVEALED DIFFERENT THAN j, 0 OTHERWISE