Are consumers rational? Shifting the burden of proof

Laurens Cherchye* Thomas Demuynck† Bram De Rock‡
Joshua Lanier§

April 2, 2021

Abstract

We present a statistical test for the hypothesis of rational utility maximization on the basis of nonparametric revealed preference conditions. Our test is conservative for the utility maximization hypothesis. We take as null hypothesis that the consumer behaves randomly, and as alternative hypothesis that she is approximately utility maximizing. Our statistical test uses a permutation method to operationalize the principle of random consumer behavior. We show that the test has an asymptotic power of one against the alternative hypothesis of approximately utility maximizing behavior. We also provide simulated power results and two empirical applications, to experimental and observational data, respectively.
Keywords: utility maximization, revealed preferences, random behavior, permutation test.

1 Introduction

Do consumers act as rational utility maximizers? Despite the huge surge in behavioral economics, the assumption of utility maximization remains a cornerstone of most

*Department of Economics, KU Leuven. E. Sabbelaan 53, B-8500 Kortrijk, Belgium. Laurens Cherchye gratefully acknowledges the Fund of Scientific Research Flanders (FWO-Vlaanderen) and the Research Fund of the KU Leuven for financial support.

†ECARES, Université Libre de Bruxelles. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. Thomas Demuynck acknowledges financial support by the Fonds de la Recherche Scientifique-FNRS under grant nr F.4516.18 and EOS project 30544469.

‡ECARES, Université Libre de Bruxelles, and Department of Economics, University of Leuven (KU Leuven). Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. Bram De Rock gratefully acknowledges FWO and FNRS for their financial support.

§China Center for Behavioral Economics and Finance, Southwestern University of Finance and Economics, Liulin Campus (Main Campus): 555, Liutai Avenue, Wenjiang District, Chengdu, Sichuan, P. R. China, 611130

models in economics. Given its importance, it is crucial to check whether actual consumer behavior is at least close to rationality. Revealed preference theory provides an attractive framework to do so. In his seminal contribution, Afriat (1967) showed that a finite data set on observed prices and consumed bundles is rationalizable by the model of utility maximization if and only if it satisfies the Generalized Axiom of Revealed Preference (GARP).[1] A most attractive conceptual feature of the revealed preference approach is that it is intrinsically nonparametric, meaning that it abstains from imposing any, typically nonverifiable, functional structure on the consumer's utility function. From a practical perspective, it has the additional advantage that it can be meaningfully applied even to small data sets. For example, GARP can be rejected with only two observed consumption bundles. These two features motivate the frequent use of revealed preference methods for testing the hypothesis of utility maximizing consumption behavior.

In applications, revealed preference tests usually start from a finite set of observed consumption decisions (prices and quantities) for a given individual, and then verify whether these observations satisfy some combinatorial condition (like GARP). The result of these deterministic tests is either a 'yes' or a 'no'. A 'yes' means that there exists a utility function that *exactly* rationalizes all observed consumption choices as utility maximizing, while a 'no' indicates the opposite. However, as argued by Varian (1991), exact utility maximization might not be a very interesting hypothesis. What we really want to know is whether consumers exhibit *nearly* optimizing behavior, meaning that the rationality hypothesis provides a useful approximation of their observed behavior. As a response to the sharp nature of the deterministic revealed preference tests, it is nowadays customary to complement the tests with a goodness-of-fit measure that quantifies how close the observed behavior is to passing the strict revealed preference conditions. The most popular measure in the applied literature is Afriat's Critical Cost Efficiency Index (CCEI). This CCEI takes values between 0 and 1, with higher values indicating that behavior is closer to exact utility maximization (see Section 3 for a formal definition). Intuitively, one minus the CCEI equals the fraction that the consumer is allowed to waste in each observed consumption decision while still being labelled as approximately utility maximizing.[2]

[1] To be precise, Afriat (1967) originally derived the empirical equivalence between utility maximization and a "cyclical consistency" condition. Varian (1982) has shown the equivalence between GARP and Afriat's cyclical consistency condition. Afriat (1967) built on earlier work of Samuelson (1938) and Houthakker (1950). See also Diewert (1973) insightful discussions of Afriat's pioneering article, and Chambers and Echenique (2016) for a recent review of the literature.

[2] See Choi, Kariv, Müller, and Silverman (2014) and Dziewulski (2020) for more discussion.

Our contribution. Despite the popularity of the CCEI in applied work, there does not exist a method that determines at which CCEI value we can reasonably conclude that the model of (approximate) utility maximization provides a good description of the observed behavior. The current paper aims to fill this gap, by providing a statistical test of individual utility maximization. More specifically, we propose to use the CCEI as a statistic for testing the null hypothesis of irrational, random consumption behavior against the alternative hypothesis of approximate utility maximization.[3] As a consequence, our testing method allows for calculating critical CCEI values to determine the statistical support for the rationality hypothesis.

Our method shifts the burden of proof for the utility maximization hypothesis: we only reject irrational/random consumer behavior if there is substantially strong evidence favoring approximate utility maximization. Our default hypothesis is not that the consumer is utility maximizing but, instead, that she is irrational. To be more precise, the null hypothesis of our test specifies that the consumer's purchasing decisions cannot be distinguished from random behavior. As we motivate in more detail in Section 4, we model irrational behavior by assuming that the consumer randomly draws consumption rays from some distribution that is independent from the budget (i.e. prices and income). Our alternative hypothesis is that the consumer is approximately utility maximizing (as characterized by a specific CCEI value). This means that our framework is conservative for the utility maximization hypothesis. The underlying argument is that, *if a data set cannot be distinguished from random behavior, then it should not be treated as arising from the process of utility maximization.*

Our testing procedure relies on a permutation approach to operationalize the principle of irrational, random choice behavior.[4] The idea of the test is fairly straightforward. For a given data set on prices and quantities, we consider the population of data sets that is obtained by fixing the budgets but permuting the consumption rays over the different observations. If the consumer is really randomizing, then the CCEI of the observed data set is equally likely to be realized as any CCEI of these permuted data

[3] We focus on the CCEI as our test statistic as this measure is well-known and easily computable. Importantly, however, the use of our testing method is not restricted to the CCEI. One may equally well use other goodness-of-fit measures that have been proposed in the revealed preference literature. Examples include the Houtman-Maks index (Houtman and Maks, 1985), the Varian index (Varian, 1991), the money pump index (Echenique, Lee, and Shum, 2011), the swaps index (Apesteguia and Ballester, 2015) or the minimum cost index (Dean and Martin, 2016). Further, our permutation test framework could be adapted to test any model of consumer behavior (rational, irrational, or anything in between) provided this model is taken as the alternative hypothesis and tested against a null of random behavior. Such an exercise would establish if a particular model provides a better account of the data than pure random choice. The key to adapting the method is to find a suitable goodness-of-fit measure analogous to the CCEI for the model considered.

[4] See, for example, Pesarin and Salmaso (2010) for a review of the permutation testing approach.

sets. As such, the distribution of the CCEIs over the permuted data sets provides the distribution for the CCEI of the true data set under the null hypothesis, conditional on the realized observations of prices and quantities.

As our test belongs to the family of permutation tests, it has the specific advantage that it is exact for any sample size. This is particularly convenient in the current setting, as individual revealed preference tests are usually conducted for a small number of observations. For example, our own empirical exercises consider real-life panel data with 26 observations per subject and experimental data with 50 choice observations per subject. In what follows, we will characterize the asymptotic power of our permutation test as well as provide simulation evidence of its power in finite samples.

Outline. The remainder of this paper unfolds as follows. Section 2 discusses related concepts that appeared in the literature on empirical revealed preference analysis. Section 3 introduces some preliminary notions, and motivates our testing procedure through numerical examples. Section 4 formally presents our statistical test and establishes its theoretical properties. Section 5 discusses simulated power results and provides two empirical applications (on experimental and observational data, respectively). Section 6 contains our conclusion. All our proofs are in the Appendix.

2 Related concepts in empirical revealed preference analysis

We start by discussing the relationship between our statistical test and some other concepts in the revealed preference literature. First, we consider the notions of power and predictive success of revealed preference tests. Next, we relate our test to other statistical tests based on revealed preference conditions for rational consumer behavior.

Power of revealed preference tests. Our approach shares some resemblance with Bronars' (1987) procedure for measuring the power of revealed preference tests. Similar to our procedure, Bronars' power index starts from the idea that irrational behavior can be modeled as random behavior.[5] Computing this index starts by generating a large number of random data sets, and the index is calculated as the fraction of these random data sets that violate (approximate) utility maximization. In the operationalization of Bronars' procedure, random behavior is usually simulated by drawing consumption bundles at random from the budget hyperplane. This, however, implies an ad hoc

[5]This idea goes back to Becker (1962).

reliance on some distribution to simulate random behavior, and different distributions may generate different power results. In addition, the chosen distribution may bear little resemblance to the actual distribution of consumption, even if the consumer is truly drawing consumption bundles at random. By contrast, our notion of irrational behavior allows subjects to draw consumption rays at random from any distribution.

From this perspective, our permutation method is more closely related to the 'bootstrap' method that has also been used for measuring the power of revealed preference tests (see, for example, Andreoni and Miller (2002)). Although this bootstrapping approach does away with the reliance on some arbitrary distribution, it has –to our knowledge– no theoretical grounding. Another main difference with our procedure is that these power measures are designed to produce an index for the strictness of a deterministic revealed preference test. Lower index values then indicate a rather weak revealed preference tests and, therefore, do not allow for strong statements favoring the rationality hypothesis. Unlike our method, however, the Bronars or bootstrap index cannot be used directly to test whether or not a particular consumer is a utility maximizer.

Predictive success. Another popular measure in empirical revealed preference analysis is Beatty and Crawford (2011)'s predictive success measure, which is based on an original idea of Selten (1991). This measure is computed as the difference between the pass rate of a revealed preference test (from a population of data sets) and one minus Bronars' power index of this test. Predictive success values close to zero imply that the pass rate for the observed data sets is close to the pass rates for the population of randomly generated data sets. By contrast, values close to one point out that (almost) all observed data pass the revealed preference tests, while the opposite holds for random data. Finally, values below zero indicate that random behavior performs better than actual behavior on the revealed preference tests.

Summarizing, Beatty and Crawford's predictive success measure tells us how well a revealed preference test can distinguish between actual behavior and random behavior. However, it remains silent about whether a particular individual behaves according to the utility maximization model or what critical values are to be used to reach that conclusion. This is exactly the distinguishing feature of our procedure. In this sense, we see the two procedures as complementary, each one highlighting a different aspect of the data.

Other statistical tests. Finally, there exist a number of statistical tests for rational consumer behavior that are based on nonparametric revealed preference conditions.

For example, some authors have developed statistical tests of utility maximization for populations of individuals on the basis of revealed preference theory (see the recent paper of Kitamira and Stoye (2018) and references therein). In the current study, however, we focus on individuals rather than populations of individuals.

We are aware of three other papers in which stochastics are introduced to the revealed preference analysis of an individual consumer. Varian (1985) and Epstein and Yatchew (1985) consider a perfectly utility maximizing consumer whose demand is observed with identically and independently normally distributed measurement error. Under this condition, they provide a lower bound on a chi-squared distributed test statistic that can be calculated given knowledge of the variance of the measurement error. A paper by Echenique, Lee, and Shum (2011) follows a similar approach. However, instead of introducing measurement error into demand, these authors introduce measurement error into prices. The Echenique, Lee, and Shum (2011) approach also requires that the marginal utility of income is constant at the observed price levels. All three studies require the analyst to know (or guess) the standard deviation of the error term. Comfortingly, our test procedure avoids such issues. Also, these other studies assume the same standard deviation for all goods (or prices), which implies that the units of analysis matter.

The most crucial difference between these tests and ours, however, pertains to the specification of the null hypothesis. These existing studies take as null the assumption that observed behavior results from rational utility maximizing behavior (with normally distributed measurement error). By contrast, the null hypothesis of our test states that a consumer's decisions follow random behavior.

3 Basic concepts

We begin our formal exposition by briefly introducing some necessary concepts and notation. Throughout, we will consider a consumption setting with $L \geq 2$ goods. A revealed preference analysis usually departs from a finite set of T observed prices $p^t = [p_1^t, \ldots, p_L^t] \in \mathbb{R}^L_{++}$ and associated quantities $q^t = [q_1^t, \ldots, q_L^t] \in \mathbb{R}^L_+$ $(1 \leq t \leq T)$. The idea is that, at each observation t, the consumer purchased the bundle q^t under the prevailing prices p^t. Let $m^t = p^t \cdot q^t$ denote the total amount of money spent. A data set is defined as an ordered collection of triples (q^t, p^t, m^t) and denoted by $D^T = (q^t, p^t, m^t)_{t \leq T}$.[6]

[6]In the literature, one often refers to a data set as the collection of pairs $D^T = (q^t, p^t)_{t \leq T}$. Indeed, the inclusion of expenditure levels is redundant as they can be calculated from the pairs (q^t, p^t). We include expenditures as it greatly simplifies our exposition when introducing the stochastic structure to the analysis.

GARP and CCEI. We say that the bundle q^t is revealed preferred to the bundle q^v if $m^t \geq p^t \cdot q^v$. We denote this by $q^t \, R \, q^v$. In words, the bundle q^t was chosen at observation t while q^v was also attainable (for the given expenditure m^t and prices p^t). Similarly, a bundle q^t is strictly revealed preferred to q^v if $m^t > p^t \cdot q^v$, which we denote by $q^t \, P \, q^v$. Intuitively, q^t was chosen although q^v was equally affordable together with some additional money left for the consumer.

A data set D^T satisfies the Generalized Axiom of Revealed Preference (GARP) if there is no 'strict' cycle in the revealed preference relation: for any sequence of observations $t_1, \ldots, t_M \leq T$:

$$q^{t_1} \, R \, q^{t_2} \, R \, \ldots \, R \, q^{t_M} \text{ implies not } q^{t_M} \, P \, q^{t_1}.$$

Afriat (1967) has shown that the observed behavior (captured by the data set D^T) can be rationalized as maximizing a well-behaved (i.e. continuous, increasing and quasi-concave) utility function if and only if the data set D^T satisfies GARP.

If a data set does not satisfy GARP, we may consider a weakening of the sharp GARP condition. As indicated in the Introduction, a popular way to do so makes use of the Critical Cost Efficiency Index (CCEI). To formally define the CCEI, we consider the relations $q^t \, R^e \, q^v$ if $e \, m^t \geq p^t \cdot q^v$ and $q^t \, P^e \, q^v$ if $e \, m^t > p^t \cdot q^v$, which make use of a prespecified 'efficiency' value $e \in [0, 1]$. Intuitively, the revealed preference relations R^e and P^e imply a weakening of the relations R and P, as q^t is now said to be (strictly) revealed preferred to q^v only if q^v was available when the budget at observation t was decreased by a fraction $(1 - e)$. We say that a data set D^T satisfies e-GARP if, for all sequences of observations $t_1, \ldots, t_M \leq T$:

$$q^{t_1} \, R^e \, q^{t_2} \, R^e \, \ldots \, R^e \, q^{t_M} \text{ implies not } q^{t_M} \, P^e \, q^{t_1}.$$

Obviously, for $e = 1$ we have that e-GARP coincides with GARP. Moreover, any data set satisfies e-GARP for $e = 0$. More generally, if a data set D^T satisfies e-GARP, then it will satisfy e'-GARP for any value $e' \leq e$. This calls for defining the highest value of e such that a data set still satisfies e-GARP. This value gives us the CCEI, which we denote by $\tau(D^T)$ for a data set D^T:

$$\tau(D^T) = \sup\{e \in [0, 1] : D^T \text{ satisfies } e\text{-GARP}\}.$$

Varian (1990) proposed the CCEI as a goodness-of-fit measure in empirical revealed preference analysis. The higher the value of the CCEI, the closer the observed data set is to satisfying GARP. As is clear from the definitions above, one minus the CCEI equals the fraction that the consumer is allowed to waste at each observed consumption decision, while still being labeled as approximately utility maximizing.

Critical CCEI value. A natural question is whether an observed CCEI value is sufficiently high to conclude that the decision maker is approximately utility maximizing and not just picking consumption bundles at random. In the literature, there is no consensus on what value the CCEI should minimally attain to conclude that behavior is (approximately) rational. Varian (1991) mentions the critical value of 0.95, but admits that this is mainly out of sentimental reasons. Choi, Fisman, Gale, and Kariv (2007) use 0.90 based on their results for Bronars' power procedure. Particularly, for their application these authors find that the CCEI is below 0.90 for most randomly generated data sets (using a uniform distribution to simulate random behavior). Most of the other papers in the literature tend to use cut-offs of 0.90, 0.95 or 0.99 (see, for example, Polisson, Quah, and Renou (2020)).

In the current paper, we set out a framework to define an individual-specific cut-off that is determined as the critical value of a statistical test. To determine this cut-off, we consider a data set that is obtained by random choice. Random choice is modeled by fixing the various budgets but permuting the consumption rays over the different observations. Figure 1 provides an illustration for three observations and two goods. The budgets are given by the solid lines and the bundles are represented by squares. Assume that we observe a consumer who picks the consumption bundles according to the top left panel of the figure. The observed data set then produces three consumption rays depicted by the dashed lines through the origin. If the consumer were irrational and picked her consumption rays from some random distribution of rays, the observed consumption pattern would be equally likely as any consumption pattern in the other 5 panels of Figure 1 that are obtained by permuting the three observed consumption rays over the observed budgets.

For a permutation σ, we denote the permuted data set by D^T_σ (see Section 4 for a formal definition), and the corresponding CCEI value by $\tau(D^T_\sigma)$. If the individual chooses her consumption bundles by randomly picking rays from some distribution then, conditional on the observed rays and budgets, the probability of observing the data set D^T must have the same likelihood as observing the data set D^T_σ. This is the main idea behind our permutation test. To put this into practice, we compute the CCEI values for all possible data sets that are obtained by permuting the consumption rays over the observations. If actual behavior picked consumption rays at random, then the CCEI value for the true data set would be a random draw from the distribution of all these CCEI values. Thus, we can reject the hypothesis of random behavior at the significance

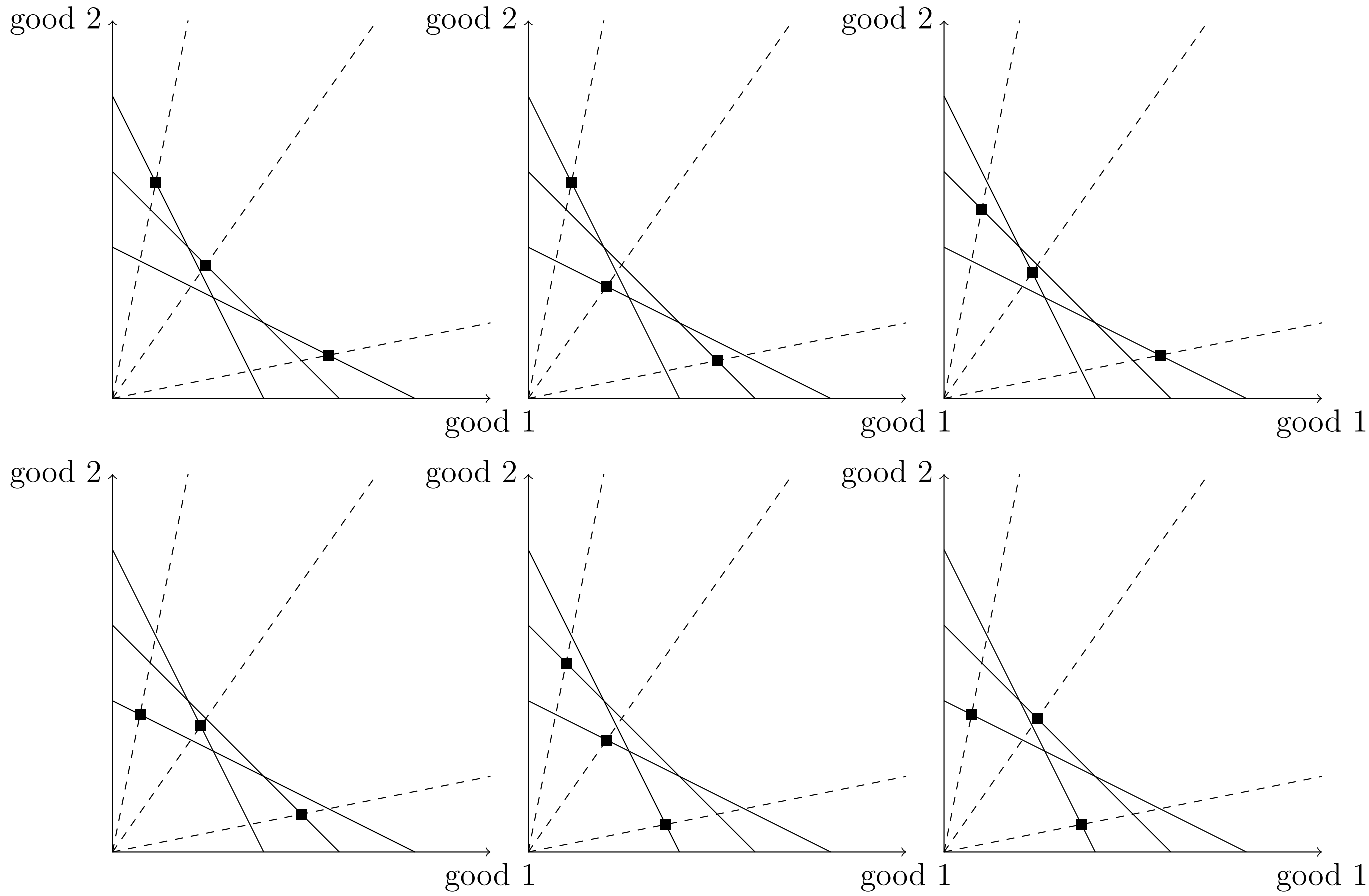

Figure 1: Permuting consumption rays

level α if at most a fraction α of all the permuted data sets have a CCEI value above or equal to $\tau(D^T)$.

We end this section with an illustrative example on how our methods can be used in practice.

Example 1. *Figure 2 shows an artificial data set with nine observations and two goods. This data set violates GARP but only to a small degree. In particular, we have that* $\tau(D^T) = 0.984$, *indicating that the CCEI is quite close to one. When computing the CCEI for all 362 880 permuted data sets, we find that 2.49% of these data sets have a CCEI that is at least as high as* 0.984. *In other words, the p-value for the null hypothesis that the consumer determines her consumption by randomly picking rays is 0.0249. We conclude that the hypothesis of random consumer behavior cannot be rejected at a significance level of 1%, while it is rejected at the 5% or 10% level.*

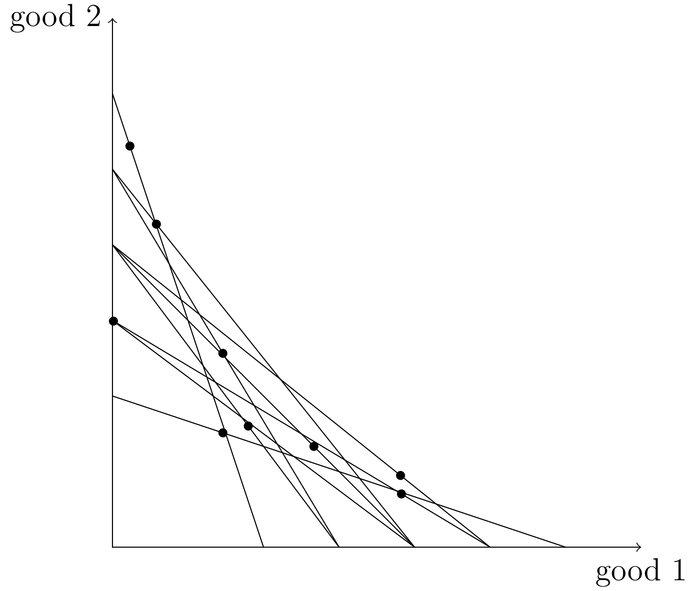

Figure 2: Data set for Example 1

4 Hypothesis Test

We aim to develop a statistical test that distinguishes between the following two hypotheses:

H_0 : The observed consumption data is generated by a random consumer.

H_1 : The observed consumption data is generated by an approximate utility maximizer.

Naturally, we have to precisely define what it means to be a random consumer as well as to be an approximate utility maximizer. We also need to show that the two hypotheses are exclusive. To do so, it will be useful to express chosen consumption bundles in terms of consumption rays. More specifically, any consumption bundle, price vector, and expenditure level triple (q, p, m) has an equivalent dual representation as a ray, price vector, and expenditure level triple (r, p, m) via the transformations:

$$r_\ell = \frac{q_\ell}{\sum_{i=1}^{L} q_i}, \qquad \ell = 1, \ldots, L, \tag{1}$$

and

$$q_\ell = m \frac{r_\ell}{p \cdot r}, \qquad \ell = 1, \ldots, L. \tag{2}$$

Given this, we can express any data set D^T either in *bundle form* $D^T = (q^t, p^t, m^t)_{t \leq T}$ or in *ray form* $D^t = (r^t, p^t, m^t)_{t \leq T}$. We will freely switch between these two representations.

In view of what follows, we note that any *ray* r belongs to the $L-1$ dimensional unit simplex:

$$\Delta^{L-1} = \left\{ r \in \mathbb{R}^L_+ : \sum_{\ell=1}^{L} r_\ell = 1 \right\}.$$

Further, we will write random variables and vectors in bold. For example, $\mathbf{r}$ is a random ray while r is a deterministic ray. A random data set in ray form and bundle form is expressed as, respectively:

$$\mathbf{D}^T = (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)_{t \leq T} \text{ and } \mathbf{D}^T = (\mathbf{q}^t, \mathbf{p}^t, \mathbf{m}^t)_{t \leq T}.$$

Next, the expression $P(E)$ will denote the probability of some event E. For example, $P(\mathbf{q}^t \in Q)$ denotes the probability that the consumer selects a consumption bundle in some set Q at observation t.[7]

Random consumer. We define the random consumer as a consumer who makes choices from budget sets by selecting an affordable bundle along a random consumption ray without regard to the budget being faced. In particular, a random consumer can be identified by a distribution R over the set of rays Δ^{L-1}.[8]

Definition 1. A random data set $\mathbf{D}^T = (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)_{t \leq T}$ is *generated by a random consumer* if there is a distribution R on Δ^{L-1} such that the rays $\mathbf{r}^1, \ldots, \mathbf{r}^T$ are independent draws from R that are also independent of the budgets $(\mathbf{p}^1, \mathbf{m}^1), \ldots, (\mathbf{p}^T, \mathbf{m}^T)$.

Definition 1 specifies that a consumer behaves randomly when her consumption choices are defined by randomly picking consumption rays from a particular distribution, irrespective of the observed prices and income levels. At this point, we remark that our null hypothesis does not exclude all utility maximizing behavior. For example, a consumer with a Leontief utility function (i.e. goods are perfect complements) will always consume on a fixed ray. We will rule out this extreme case by requiring utility functions to be strictly increasing (see Assumption 1).

We acknowledge that the random rays null hypothesis that we use represents only one of many possible ways to model random behavior. For example, Becker (1962), Bronars (1987) and Beatty and Crawford (2011) equate random behavior as picking budget shares instead of rays. Our test can easily be altered to allow for a null hypothesis with the consumer selecting random budget shares instead of random consumption rays. Importantly, however, an advantage of our random rays approach is elucidated via

[7]More technically, there is some underlying probability space $(\Omega, \mathcal{F}, P)$. An observation $(\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)$ is a triple of random vectors, each measurable in the appropriate Borel sigma algebra.

[8]That is, R is a probability measure on the Borel sigma algebra on the unit simplex Δ^{L-1}.

Theorems 3 and 4 that we present below. These results lean heavily on the random rays null hypothesis and it is not straightforward to see how they can be altered to accommodate a random shares hypothesis. Moreover, Example 2 provides an example of one (of many) specific intuition of our random consumer model as a special case of a random utility model (RUM). In particular, any data set generated by a random consumer (according to our definition) can be thought of as being generated by a random utility maximizer with Leontief preferences.

Example 2. *The following household acts as a whole as a random consumer. There are two goods; cereal and milk. Ann and Bob live together. Once a week, either Ann or Bob goes shopping. The probability that Ann goes is 50 percent and the probability that Bob goes is also 50 percent. If Ann goes shopping then she always buys two units of cereal and one unit of milk. If Bob goes then he always buys one unit of cereal and two units of milk.*

Permutation test. We are interested in comparing the CCEI of a random data set $\mathbf{D}^T$ with that of a permuted data set $\mathbf{D}^T_\sigma$. From our definition of permuted data sets, it will be clear that if $\mathbf{D}^T$ is generated by a random consumer, then so is $\mathbf{D}^T_\sigma$, from which it follows that the CCEI values $\tau(\mathbf{D}^T)$ and $\tau(\mathbf{D}^T_\sigma)$ have the same distribution. This will be instrumental in defining bounds on the Type-1 errors in Theorem 1 below.

Let Π be the set of all permutations on $\{1,\ldots,T\}$.Then for $\sigma \in \Pi$, we define:

$$\mathbf{D}^T_\sigma = (\mathbf{r}^{\sigma(t)}, \mathbf{p}^t, \mathbf{m}^t)_{t\leq T}$$

to be a *permuted version* of $\mathbf{D}^T = (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)_{t\leq T}$. For a data set $\mathbf{D}^T$ and a significance level $\alpha \in (0,1)$, the binary random variable $\phi_\alpha(\mathbf{D}^T)$ returns a value of 1 if the fraction of permuted data sets that have higher CCEI levels than the data set $\mathbf{D}^T$ is less than α, while it returns a value of 0 otherwise. Formally:

$$\phi_\alpha(\mathbf{D}^T) = \mathbb{1}\left[\frac{\left|\left\{\sigma \in \Pi : \tau(\mathbf{D}^T_\sigma) \geq \tau(\mathbf{D}^T)\right\}\right|}{T!} \leq \alpha\right], \tag{3}$$

where $\mathbb{1}[.]$ is the indicator function that equals 1 if the expression between brackets is true and 0 otherwise. Using this, we can describe our permutation test as follows.

Procedure 1. *Let $\alpha \in (0,1)$. Reject H_0 in favor of H_1 at significance level α if the proportion of permutations $\sigma \in \Pi$ that satisfy $\tau(\mathbf{D}^T_\sigma) \geq \tau(\mathbf{D}^T)$ is weakly smaller than α, that is, if:*

$$\phi_\alpha(\mathbf{D}^T) = 1.$$

The next theorem motivates the theoretical validity of Procedure 1, by showing that the probability of making Type-1 errors is at most α.[9] We note that this result holds for a data set of any size T.

Theorem 1. *Let $\alpha \in (0,1)$ and suppose that the random data set $\mathbf{D}^T = (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)_{t \leq T}$ is generated by a random consumer. Then Procedure 1 rejects H_0 with probability less than or equal to α. That is, $P\left(\phi_\alpha(\mathbf{D}^T) = 1\right) \leq \alpha$.*

Approximate utility maximization. We next present the details of our alternative hypothesis, which states that "the observed consumption is generated by an approximate utility maximizer". We will characterize an approximate utility maximizer by a pair (U, e) where $e \in (0, 1]$ is a measure of rationality and U is well-behaved in the following sense.[10]

Assumption 1 (Well-Behaved)**.** *The utility function $U : \mathbb{R}^L_+ \to \mathbb{R}$ is continuous, strictly increasing, and quasi-concave.*

Let $B(p, m) = \{q \in \mathbb{R}^L_+ : p \cdot q = m\}$ denote the budget hyperplane given a price vector and expenditure pair $(p, m) \in \mathbb{R}^L_{++} \times \mathbb{R}_{++}$. Then we use $x(p, m)$ for the *(optimal) demand correspondence* generated by a well-behaved utility function U, that is:

$$x(p, m) = \arg\max U(q) \text{ subject to } q \in B(p, m). \tag{4}$$

Given the one-to-one mapping between the bundle representation (q, p, m) and the ray representation (r, p, m), we can also define the *(optimal) ray correspondence* $\mu(p, m)$, which determines the ray(s) that are chosen by a utility maximizing consumer:

$$\mu(p, m) = \left\{ r \in \Delta^{L-1} : r = \frac{q}{\sum_{\ell=1}^{L} q_\ell}, \text{ for } q \in x(p, m) \right\}. \tag{5}$$

We assume two additional properties for an approximate utility maximizer characterized by the pair (U, e). First, when faced with the budget (p^t, m^t), she always picks bundles that give at least as much utility as what could be secured by a perfect utility maximizer selecting from the budget set $B(p^t, e\, m^t)$.

Assumption 2. *For all $t \leq T$:*

$$U(\mathbf{q}^t) \geq \max_{\mathbf{q}} U(\mathbf{q}) \textit{ subject to } \mathbf{q} \in B(\mathbf{p}^t, e\mathbf{m}^t) \textit{ with probability 1.} \tag{6}$$

[9] From the continuity of the CCEI function $\tau(D^T)$, it follows that ϕ_α is a measurable function.

[10] By strictly increasing we mean $U(q) > U(q')$ when $q > q'$ (i.e. $q_\ell \geq q'_\ell$ for all goods ℓ and $q \neq q'$). As alluded to earlier, this rules out Leontief preferences, which are indistinguishable from random behavior as defined in our null hypothesis.

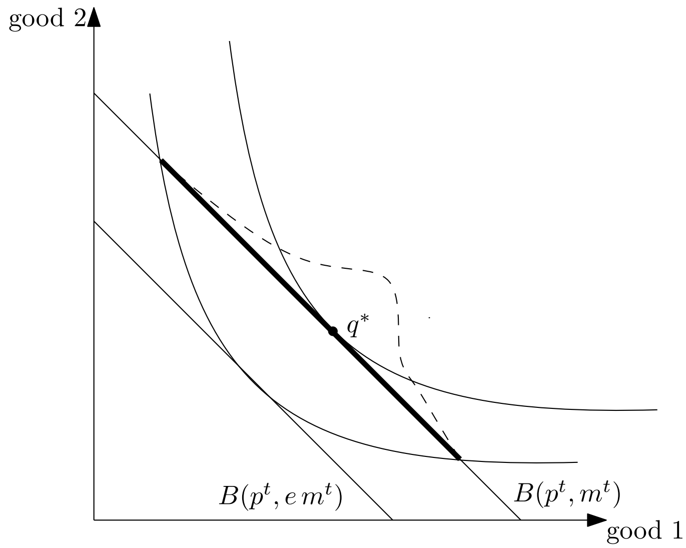

Figure 3: Illustration of Assumptions 2 and 3

Figure 3 provides intuition for this assumption.[11] The figure shows the budget hyperplanes $B(p^t, m^t)$ and $B(p^t, e\, m^t)$ for some $e < 1$, and the two indifference curves corresponding to maximum utility for these budgets. Assumption 2 imposes that the random bundle $\mathbf{q}^t$, of which the distribution is restricted to the budget hyperplane $B(p^t, m^t)$, must provide at least as much utility as the optimal bundle at $B(p^t, e\, m^t)$. In other words, the support of $\mathbf{q}^t$ is restricted to the thick part of the budget hyperplane $B(p^t, m^t)$. An example of such a distribution is illustrated by the dashed curve. It is easy to see that, for given U, the support region grows as e decreases. Conversely, if e gets close to 1, the support of $\mathbf{q}^t$ is increasingly restricted to a close neighborhood of the optimal bundle q^*.

The second property that we impose is that a random data set for an approximate utility maximizer must be *U-possibly perfect*. This means that, for any open neighborhood R of the utility maximizing rays $\mu(p^t, m^t)$, the probability with which the consumer selects a ray in R is strictly positive. Intuitively, a data set is U-possibly perfect if it is possible that the consumer's choices are arbitrarily close to perfect optimization.

Assumption 3 (U-possibly perfect)**.** *For all observations $t \leq T$, all open sets $B \subseteq \mathbb{R}^L_{++} \times \mathbb{R}_{++}$ and all (relative) open sets $R \subseteq \Delta^{L-1}$ satisfying $\mu(p, m) \subseteq R$ for all $(p, m) \in B$, if $P((\mathbf{p}^t, \mathbf{m}^t) \in B) > 0$, then:*

$$P\left(\mathbf{r}^t \in R \middle| (\mathbf{p}^t, \mathbf{m}^t) \in B\right) > 0. \tag{7}$$

[11]Let $V(p, m) = U(x(p, m))$ where x is the demand correspondence defined in (4). It is well-known that V is continuous. Equation (6) is equivalent to $U(\mathbf{q}^t) \geq V(\mathbf{p}^t, e\mathbf{m}^t)$, which clearly corresponds to a measurable event.

Going back to Figure 3, we know that Assumption 2 restricts the support of $\mathbf{q}^t$ to the thick part of the budget hyperplane $B(p^t, m^t)$. Assumption 3 then requires that the probability of having $\mathbf{q}^t$ in any open neighborhood of the optimal bundle q^* is strictly positive. For example, this holds when the distribution of $\mathbf{q}^t$ on the budget hyperplane $B(p^t, m^t)$ is continuous and has strictly positive density at q^*. This is the case for the dashed curve in Figure 3.

Definition 2. A random data set $\mathbf{D}^T = (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)_{t \leq T}$ is generated by an *approximate utility maximizer* if there exists a pair (U, e) with $e \in (0, 1]$ such that Assumptions 1, 2 and 3 are satisfied.

The following theorem connects Assumption 2 to the notion of e-GARP and thus to the CCEI function τ that we use as our test statistic.

Theorem 2. *Suppose $e \in (0, 1]$ and let $D^T = (q^t, p^t, m^t)_{t \leq T}$ be a (deterministic) data set. There exists a utility function U that satisfies Assumption 1 and:*

$$U(q^t) \geq \max_q \; U(q) \text{ subject to } q \in B(p^t, em^t), \text{ for all } t \leq T, \tag{8}$$

if and only if D^T satisfies e-GARP.

A straightforward corollary of Theorem 2 is the following. If $\mathbf{D}^T$ is generated by an approximate utility maximizer characterized by the pair (U, e), then $\mathbf{D}^T$ satisfies e-GARP with probability 1.

The next identification result shows that a data set cannot be generated by both a random consumer and an approximate utility maximizer given a full support assumption on the distribution of budgets.[12] In other words, the two types of consumers are empirically distinguishable.

Theorem 3. *Let $\mathbf{D}^T = (\mathbf{q}^t, \mathbf{p}^t, \mathbf{m}^t)_{t \leq T}$ be a random data set. Suppose that $(\mathbf{p}^t, \mathbf{m}^t)$ has support $\mathbb{R}^L_{++} \times \mathbb{R}_{++}$ for all $t \leq T$. Then $\mathbf{D}^T$ cannot be generated by both a random consumer and an approximate utility maximizer.*

Summarizing, Theorem 1 motivates our permutation test by characterizing its size, and Theorem 3 gives conditions such that our null and alternative hypothesis are distinguishable. Our last result concerns the power of our procedure, i.e. the probability

[12] We say that a random vector $\mathbf{X}$ in $\mathbb{R}^N$ has support $A \subseteq \mathbb{R}^N$ if (i) $P(\mathbf{X} \in A) = 1$ and (ii) for any non-empty open set $E \subseteq A$ we have $P(\mathbf{X} \in E) > 0$. This definition of support is not standard but it allows us to say that the support of a price vector $\mathbf{p}$ is $\mathbb{R}^L_{++}$ whereas the support would usually be a closed set. Of course we could talk about relative topologies (in which case $\mathbb{R}^L_{++}$ could be a closed set) or simply say that the support of a price vector is $\mathbb{R}^L_{+}$, but we feel this may cause more confusion than clarity.

of rejecting the null hypothesis when the alternative is true. To address this issue, we present conditions under which the asymptotic power of the test is 1.

Theorem 4. *Suppose the random data set $\mathbf{D}^T = (\mathbf{q}^t, \mathbf{p}^t, \mathbf{m}^t)_{t \leq T}$ is generated by an approximate utility maximizer characterized by the pair (U, e) (with $e \in (0, 1]$). Suppose that the individual observations are independent and identically distributed and that $(\mathbf{p}^t, \mathbf{m}^t)$ has support $\mathbb{R}^L_{++} \times \mathbb{R}_{++}$ for all $t \leq T$. Then for any significance level $\alpha \in (0, 1)$, the probability with which Procedure 1 rejects H_0 goes to 1 as $T \to \infty$. That is, $\lim_{T \to \infty} P(\phi_\alpha(\mathbf{D}^T) = 1) = 1$.*

Two final qualifications apply to our results. First, Theorem 4 assumes that observations are independent and identically distributed (IID), which may not always be a sensible assumption. The IID assumption can be weakened at the cost of some additional complexity. This is addressed by Theorem 6 in the Appendix. Next, it is perhaps not obvious how to interpret the full support assumption in Theorems 3 and 4. We view the assumption as a requirement that there is sufficient price variation in the data. The full support assumption is further relaxed by Theorem 5 in the Appendix, again at the cost of introducing somewhat more elaborate notation. We also note that full support conditions are not uncommon in the related literature (see, for example, Matzkin (1994) and references therein).

5 Empirical exercises

Theorem 4 is a large sample result. In what follows, we first conduct a simulation exercise that investigates the power of our permutation test in the finite sample case. Subsequently, we demonstrate the empirical usefulness of our testing procedure by applying it to the experimental data set of Fisman, Kariv, and Markovits (2007) and the real-life Stanford Basket data set that was also used by Echenique, Lee, and Shum (2011). To illustrate the versatility of our approach, we end by exploring the added value of considering the more restricted class of quasi-linear preferences.

The permutation test that we outlined in the previous section starts by calculating the CCEI, $\tau(D^T)$, for a data set D^T of observed prices and quantities associated with a single consumer. Subsequently, it permutes the budget rays across observations, and computes the CCEI, $\tau(D^T_\sigma)$, for each permuted data set D^T_σ. This means that we must compute $\tau(D^T_\sigma)$ for each of the $(T!)$ possible permutations σ. However, for large enough data sets, this quickly becomes computationally intractable. For example, for a data set of $T = 50$ observations, this requires $(50!) \geq 3 \times 10^{64}$ permutations. To ensure computational feasibility, it is standard practice in the literature on permutation tests to take a large enough sample of random permutations when the number of observations becomes too large. In our following exercises, our test uses all possible permutations when the number of observations equals at most $T = 7$. In the other cases, we randomly sample 10 000 permutations with replacement. In order to speed up our computations, we further employ the following 'heuristic': if after running the test using 1 000 permutations we find a p-value strictly greater than 0.2, we abort the test and report the results using only these 1 000 permutations. This saves us the trouble of refining our p-value when there is little chance of ever approaching the 10% significance level.

The CCEI-value of a data set is computed using a standard binary search algorithm. We choose the number of iterations that guarantees that the CCEI is calculated to within an error of 2^{-17} of the true value. We make sure to always test the data set for GARP so that if the data set is perfectly rationalizable, we return a CCEI value of 1.

Simulated data. To compute the power of our test, we need to generate data that are consistent with e-GARP for chosen values of e. To generate a budget set when there are L goods we start by drawing L numbers uniformly from the interval $[1, 10]$. Denote these numbers by $\alpha_1, \ldots, \alpha_L$. Next, for each good ℓ we set the price equal to $10/\alpha_\ell$, and we fix the expenditure level at 10. This ensures that if the consumer allocates her income exclusively to good ℓ, she could purchase exactly α_ℓ units. Once the budgets are selected, we generate 100 random data sets satisfying e-GARP by using the Markov Chain Monte Carlo algorithm of Demuynck (2020).

Table 1 contains the simulation results for various numbers of goods ($L = 2, 4, 8, 16$) and different number of observations ($T = 6, 8, 10, 14, 20$). The different cells reveal the power of our statistical test for alternative combinations of L, T, e and α. For example, the cell ($T = 8, L = 2$ with $e = 0.99$ and $\alpha = 0.10$) has a value 0.55. This says that for 55 percent of our simulations, we reject the null hypothesis of random behavior at the 10% significance level. Generally, we find that the power of our permutation test increases in the number of goods (L) and the number of observations (T). It naturally decreases when e decreases. The power is close to 1 as soon as we have 20 observations. We conclude that our test has sufficient power whenever T is moderately large. This especially holds true when the number of goods L is sufficiently large.

Experimental data. An advantage of experimental data is that they allow for gathering a high number of consumption observations for one and the same individual at low cost. In addition, the experimental designer has full control over the various budgets faced by the experimental subjects. This type of data is exactly in line with the setting in our simulation exercise above, which motivates that our procedure has sufficient

		$e = .99$			$e = .95$			$e = .90$		
		α			α			α		
	T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
$L = 2$	6	0.12	0.02	0.00	0.07	0.02	0.00	0.03	0.00	0.00
	8	0.55	0.22	0.00	0.34	0.08	0.00	0.25	0.07	0.00
	10	0.88	0.63	0.14	0.77	0.44	0.03	0.63	0.39	0.03
	14	1.00	1.00	0.81	0.98	0.90	0.53	0.92	0.75	0.24
	20	1.00	1.00	1.00	1.00	1.00	0.96	1.00	0.99	0.80
$L = 4$	6	0.25	0.06	0.00	0.25	0.07	0.00	0.22	0.04	0.00
	8	0.93	0.61	0.06	0.81	0.49	0.01	0.68	0.34	0.02
	10	1.00	0.96	0.43	0.96	0.89	0.34	0.86	0.68	0.16
	14	1.00	1.00	1.00	1.00	1.00	0.96	1.00	1.00	0.77
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
$L = 8$	6	0.58	0.25	0.01	0.53	0.22	0.01	0.40	0.15	0.01
	8	0.99	0.87	0.18	0.97	0.82	0.18	0.89	0.75	0.09
	10	1.00	1.00	0.88	0.99	0.99	0.76	0.98	0.90	0.60
	14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$L = 16$	6	0.79	0.53	0.08	0.82	0.45	0.12	0.82	0.48	0.05
	8	1.00	0.98	0.50	1.00	0.95	0.45	1.00	0.95	0.53
	10	1.00	1.00	0.97	1.00	1.00	0.95	1.00	1.00	0.95
	14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 1: Power results for simulated data

power.

We illustrate this for the data set on giving versus keeping of Fisman, Kariv, and Markovits (2007). This experiment was designed to investigate individual preferences for giving by exposing subjects to a series of dictator games under varying incomes and conversion rates between giving and keeping.[13] The data set has two components. The first component contains information for 76 subjects (i.e. 76 consumers) on 50 choices between keeping and giving to one other individual (i.e. 2 goods), and the second component contains information for 65 subjects on choices between keeping and giving to either individual A or individual B (i.e. 3 goods). We refer to Fisman, Kariv, and Markovits (2007) for more details on the data.

Table 2 summarizes our results. Attractively, we find that the experimental data allow us to statistically discriminate between utility maximizing and random behavior. All rejection rates are well above the nominal significance levels, even without imposing specific additional structure on the consumers' utility functions (see also below). For instance, we reject the null hypothesis of random choice behavior at the 1% significance level for 72% of the subjects (for the choices with 2 goods) and 83% of the subjects (for the choices with 3 goods). In our opinion, this convincingly demonstrates that our permutation test can have substantial empirical bite in practice.

Real-life data. We next study the Stanford Basket data set that was also used by Echenique, Lee, and Shum (2011). This data set captures consumer expenditures on 14

[13]In particular, subjects made several choices by filling in questions of the form: *"Divide X tokens: Hold ______ at a points, and Pass ______ at b points (the Hold and Pass amounts must sum to X)"*. The parameters X, a and b were varied across the decision problems.

types of goods that fall in the "food" category, covering the period from June 1991 to June 1993 (i.e. 104 weeks). There are 494 consumers and, after aggregating up to brand level and dropping goods which have no price data for some weeks, we retain a total of 430 goods. We aggregate the data so that one period represents 4 weeks, resulting in a maximum of 26 periods per participant. All our aggregation steps closely follow the same procedures as in Echenique, Lee, and Shum (2011).

If we compute CCEI values for the 494 consumers in the Stanford Basket Data set, we find that 416 (84.2%) have a CCEI value below unity, i.e. they violate the sharp GARP condition. Still, we find that the CCEI values are generally high. The average CCEI equals 0.9504, with a standard deviation of 0.0578. Although the minimum CCEI value equals no more than 0.4278, we observe that the first quartile, median and third quartile amount to 0.93, 0.97 and 0.99, respectively. This may suggest that the observed behavior is generally close to approximate utility maximization.

Our test procedure allows us to investigate the statistical support for this claim. In particular, we can use our procedure to assess for which subjects we reject the null of random behavior. The results of this exercise are also given in Table 2. Generally, we find that the statistical support for utility maximizing behavior is rather weak when using a significance level of 1%, with the rejection rate of the null hypothesis amounting to only 10%. The picture is somewhat more nuanced for the 10% significance level, with a rejection rate of 40%.

Quasi-linear preferences. One possible conclusion from the results in Table 2 is that the restrictions imposed by nearly utility maximizing behavior are often not sufficiently restrictive to significantly distinguish such behavior from purely random behavior. So many types of behavior can count as approximate utility maximization that it is often hard to differentiate it from randomness. To explore this in more detail, we applied our testing procedure when using the (stronger) alternative hypothesis of approximate utility maximization with quasi-linear preferences. Particularly, we say that a utility function U is quasi-linear if there exists an outside good y such that we can write:

$$U(q, y) = V(q) + y.$$

The model of quasi-linear utility maximization is substantially more restrictive than the standard utility maximization model. As such, if people effectively behave like approximate quasi-linear utility maximizers, we should more easily detect this when using an appropriate statistical test. For this exercise, we make use of the revealed preference characterization of quasi-linear utility maximization that was developed by

	Experimental data		Real-life data	
Sign. level	Rejection rates Gen. pref., 2 goods	Rejection rates Gen. pref., 3 goods	Rejection rates Gen. pref.	Rejection rates Quasi-linear pref.
$\alpha = 0.10$	0.88	0.94	0.40	0.49
$\alpha = 0.05$	0.82	0.92	0.30	0.39
$\alpha = 0.01$	0.72	0.83	0.10	0.24

Table 2: Rejection rates for experimental and real-life data

Brown and Calsimiglia (2007, Theorem 2.2), which we adapt to our particular setting.[14]

As expected, the goodness-of-fit of this more restricted model, measured once more by the CCEI, is significantly lower. In this case, the first quartile, median and third quartile amount to respectively 0.7875, 0.8390 and 0.8810. Next, the results of our statistical test are again summarized in Table 2. It is interesting to note that there are many people for which we reject the null in favor of the alternative hypothesis of nearly quasi-linear utility maximization, but not in favor of the standard nearly utility maximization model, particularly when using a significance level of 1%. This shows that it might often be useful to focus on a more restricted class of utility functions to verify the utility maximization hypothesis. If the observed behavior is consistent with a more restrictive utility maximization model, it will generally be easier to distinguish such optimizing behavior from purely random behavior.

6 Conclusion

We present a novel statistical testing procedure for the hypothesis of (approximate) utility maximization on the basis of nonparametric revealed preference conditions. It allows us to compute critical values for the CCEI in order to statistically distinguish utility maximization from random behavior. A specific feature of our test procedure is that it shifts the burden of proof: we only reject random consumption behavior if there is substantially strong evidence favoring utility maximizing behavior. We take as null hypothesis that consumers behave randomly, and as alternative hypothesis that consumers are approximate utility maximizers. Our statistical test makes use of a permutation method to operationalize the principle of randomization. This permutation procedure is also valid for small samples and allows us to characterize the asymptotic power of the test.

[14]Specifically, the CCEI for quasi-linear utility maximization can be calculated by testing the data for cyclical monotonicity (as defined in Brown and Calsimiglia (2007)), which is the approach we take here.

We illustrate the practical usefulness of our test for both experimental and observational scanner data. Our application to experimental data shows the use of experiments to statistically discriminate between utility maximizing and random behavior. A main advantage of experimental data is that it allows for gathering a high number of consumption observations for one and the same individual at low cost. This can yield a strong statistical test even when focusing on the standard utility maximization model. Finally, our application to real-life data illustrates the possibility of adding additional structure on the preferences of the consumer (in our case, quasi-linearity) to strengthen the test. If the observed behavior is (approximately) utility maximizing for such more structured preferences, it will generally be easier to statistically distinguish optimizing behavior from random behavior.

References

Afriat, S. N., 1967. The construction of utility functions from expenditure data. International Economic Review 8 (1), 67–77.

Andreoni, J., Miller, J., 2002. Giving according to garp: An experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.

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

Beatty, T. K. M., Crawford, I. A., 2011. How demanding is the revealed preference approach to demand. American Economic Review 101, 2782–2795.

Becker, G. S., 1962. Irrational behavior and economic theory. Journal of Political Economy 70, 1–13.

Billingsley, P., 1986. Probability and Measure, 2nd Edition. John Wiley and Sons.

Bronars, S. G., 1987. The power of nonparametric tests of preference maximization. Econometrica, 693–698.

Brown, D. J., Calsimiglia, C., 2007. The nonparametric approach to applied welfare analysis. Economic theory 31, 183–188.

Chambers, C., Echenique, F., 2016. Revealed Preference Theory. Cambridge University Press.

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., Müller, W., Silverman, D., 2014. Who is (more) rational? The American Economic Review 104, 1518–1550.

Dean, M., Martin, D., 2016. Measuring rationality with the minimum cost of revealed preference violations. Review of Economics and Statistics 98, 524–534.

Demuynck, T., 2020. A markov chain monte carlo procedure to generate revealed preference consistent datasets. Tech. rep., ECARES Working Paper 220-33.

Diewert, W. E., 1973. Afriat and revealed preference theory. The Review of Economic Studies 40, 419–425.

Dziewulski, P., 2020. Just-noticeable difference as a behavioural foundation of the critical cost-efficiency index. Journal of Economic Theory 188, 105071.

Echenique, F., Lee, S., Shum, M., 2011. The money pump as a measure of revealed preference violations. Journal of Political Economy 119, 1201–1223.

Epstein, L. G., Yatchew, A. J., 1985. Non-parametric hypothesis testing procedures and applications to demand analysis. Journal of Econometrics 30 (1), 149 – 169.

Fisman, R., Kariv, S., Markovits, D., 2007. Individual preferences for giving. American Economic Review 97, 1858–1876.

Fostel, A., Scarf, H. E., Todd, M. J., 2004. Two new proofs of Afriat's theorem. Economic Theory 24, 211–219.

Houthakker, H. S., 1950. Revealed preference and the utility function. Economica 17, 159–174.

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

Kitamira, Y., Stoye, J., 2018. Nonparametric analysis of random utility models. Econometrica 86, 1883–1909.

Matzkin, R. L., 1994. Chapter 42 restrictions of economic theory in nonparametric methods. Vol. 4 of Handbook of Econometrics. Elsevier, pp. 2523 – 2558.

Pesarin, F., Salmaso, L., 2010. The permutation testing approach: A review. Statistica 70, 481–509.

Polisson, M., Quah, J. K.-H., Renou, L., 2020. Revealed preferences over risk and uncertainty. American Economic Review 110, 1782–1820.

Samuelson, P. A., 1938. A note on the pure theory of consumer's behaviour. Economica 5, 61–71.

Selten, R., 1991. Properties of a measure of predictive success. Mathematical Social Sciences 21, 153–167.

Varian, H. R., 1982. The nonparametric approach to demand analysis. Econometrica, 945–973.

Varian, H. R., 1985. Non-parametric analysis of optimizing behavior with measurement error. Journal of Econometrics 30 (1), 445 – 458.

Varian, H. R., 1990. Goodness-of-fit in optimizing models. Journal of Econometrics 46, 125–140.

Varian, H. R., 1991. Goodness-of-fit for revealed preference tests. Tech. rep.

A Proofs

A.1 Proof of Theorem 1

Fix $\alpha \in (0,1)$ and suppose the data set $\mathbf{D}^T = (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m^t})_{t \leq T}$ is generated by a random consumer. Let Π contain all permutations $\sigma : \{1, \ldots, T\} \to \{1, \ldots, T\}$. It is clear that $\mathbf{D}^T$ and $\mathbf{D}^T_\sigma$ have the same distribution for all permutations $\sigma \in \Pi$. As such, we have that, for all $\sigma \in \Pi$:

$$\mathbb{E}[\phi_\alpha(\mathbf{D}^T)] = \mathbb{E}[\phi_\alpha(\mathbf{D}^T_\sigma)].$$

Averaging the right hand side over all permutations gives:

$$\begin{aligned}\mathbb{E}[\phi_\alpha(\mathbf{D}^T)] &= \frac{1}{T!}\sum_{\sigma \in \Pi} \mathbb{E}[\phi_\alpha(\mathbf{D}^T_\sigma)] \\ &= \frac{1}{T!}\mathbb{E}\left[\sum_{\sigma \in \Pi} \phi_\alpha(\mathbf{D}^T_\sigma)\right],\end{aligned}$$

where the last equality follows from exchanging integration and summation. For the sum within the expectation sign, we have:

$$\sum_{\sigma\in\Pi}\phi_\alpha(\mathbf{D}^T_\sigma) = \sum_{\sigma\in\Pi}\mathbb{1}\left[\frac{\left|\left\{\rho\in\Pi : \tau(\mathbf{D}^T_\rho)\geq\tau(\mathbf{D}^T_\sigma)\right\}\right|}{T!}\leq\alpha\right]. \quad (9)$$

Consider a realization D^T of $\mathbf{D}^T$ and a ranking of all permuted data sets D^T_σ ($\sigma \in \Pi$) according to their CCEI, $\tau(D^T_\sigma)$, from smallest to largest. Then the term in the summation on the right hand side of (9) will be zero for the lowest values of the ranking and will be equal to 1 from the $(1-\alpha)$th quantile onward. As such, for all realizations D^T, we have:

$$\sum_{\sigma\in\Pi}\phi_\alpha(D^T_\sigma)\leq\alpha\,(T!).$$

From this, we obtain:

$$\begin{aligned}\mathbb{E}(\phi_\alpha(\mathbf{D}^T)) &\leq \frac{1}{T!}\mathbb{E}\left[(T!)\alpha\right]\\ &= \alpha.\end{aligned}$$

A.2 Proof of Theorem 2

Suppose that $D^T = (q^t, p^t, m^t)_{t\leq T}$ satisfies (8) for some well-behaved utility function U and $e\in(0,1]$. Let $q^t R^e q^v$, i.e. $e\,m^t \geq p^t\cdot q^v$. Then $U(q^t)\geq \max_{p^t\cdot\tilde{q}\leq em^t} U(\tilde{q})\geq U(q^v)$, so $U(q^t)\geq U(q^v)$. Similarly, we can show that $q^t P^e q^v$ implies $U(q^t) > U(q^v)$. Then if e-GARP is violated, we have that there is a sequence $t_1,\ldots,t_M\leq T$ such that:

$$q^{t_1}R^e\ldots R^e q^{t_M} \text{ and } q^{t_M}P^e q^{t_1}.$$

However, this implies:

$$U(q^{t_1})\geq\ldots\geq U(q^{t_M}) \text{ and } U(q^{t_M}) > U(q^{t_1}),$$

a contradiction.

For the reverse, let $D^T = (q^t, p^t, m^t)_{t\leq T}$ satisfy e-GARP. From Fostel, Scarf, and Todd (2004), we know that there exist numbers U^t and $\lambda^t > 0$ such that, for all observations $t, v \leq T$:

$$U^t - U^v \leq \lambda^v p^v \cdot (q^t - eq^v). \tag{10}$$

Consider the utility function:

$$V(q) = \min_{t \leq T} \left\{ U^t + \lambda^t p^t \cdot (q - eq^t) \right\}.$$

The function V is increasing, continuous and concave. Let us first show that $V(q^t) \geq U^t$. If not, there must exist an observation v such that:

$$V(q^t) = U^v + \lambda^v p^v \cdot (q^t - eq^v) < U^t.$$

This contradicts (10). Now, towards a contradiction, assume that for all U the data set D^T does not satisfy (8). Then there is an observation t such that:

$$V(q^t) < \max_q V(q) \text{ s.t. } p^t \cdot q \leq e\, m^t.$$

Let q^* solve the maximization problem on the right hand side. Then:

$$\begin{aligned} U^t &\leq V(q^t) \\ &< V(q^*) \\ &\leq U^t + \lambda^t p^t \cdot (q^* - eq^t) \\ &= U^t + \lambda^t (p^t \cdot q^* - e(p^t \cdot q^t)) \\ &\leq U^t, \end{aligned}$$

where the last inequality comes from the fact that $p^t \cdot q^* \leq e\, m^t = e\, p^t \cdot q^t$. This gives the desired contradiction.

A.3 Proof of Theorem 4

We will prove Theorem 4 in 3 steps by first proving two stronger results (Theorems 5 and 6). Theorem 5 relaxes both the IID and full support assumption in terms of two higher level conditions. Theorem 6 relaxes the IID condition.

Figure 4 gives the dependencies between the various Theorems and Lemmata leading to Theorem 4. The Lemmata and their proofs can be found in Appendix A.5.

Theorem 6 and its proof use conditional probabilities, i.e. probabilities of events after conditioning on some σ-algebra (or the σ-algebra generated by a random vector). As

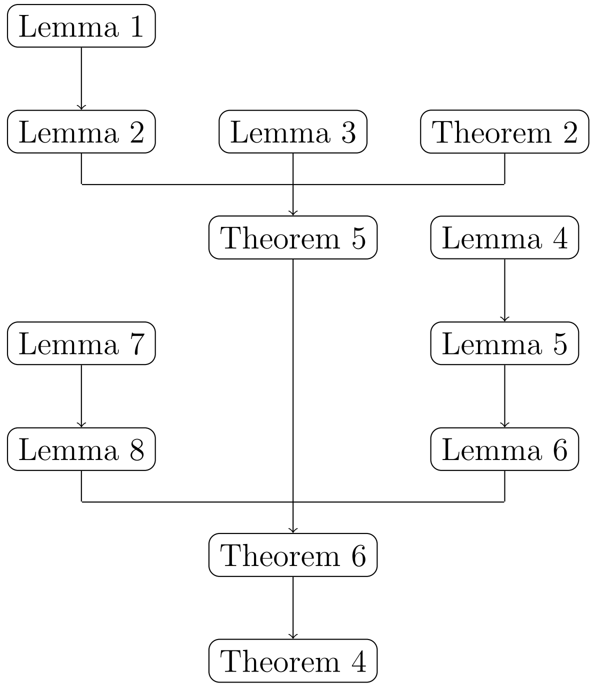

Figure 4: Structure for the proof of Theorem 4

the concept may not be familiar to many readers we feel a brief primer might be in order.

A.3.1 Conditional Probabilities

Let $(\Omega, \mathcal{F}, P)$ be a probability space. For $A, B \in \mathcal{F}$ with $P(B) > 0$ let $P(A|B)$ be the familiar probability of A conditional on B. That is, $P(A|B) = P(A \cap B)/P(B)$. Sometimes we would like to condition on probability 0 events (i.e. an event $B \in \mathcal{F}$ satisfying $P(B) = 0$). Clearly the formula $P(A|B) = P(A \cap B)/P(B)$ can no longer used in this case. The solution is to condition on an entire σ-algebra of interest instead of one event.

Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let $\mathcal{R}^N$ denote the Borel σ-algebra on $\mathbb{R}^N$. A random vector $X : \Omega \to \mathbb{R}^N$ is said to be $\mathcal{G}$-measurable if $X^{-1}(A) \in \mathcal{G}$ for all $A \in \mathcal{R}^N$. For $A \in \mathcal{F}$ let $P(A|\mathcal{G}) : \Omega \to [0, 1]$ be a $\mathcal{G}$-measurable random variable which satisfies:

$$P(A \cap G) = \int_G P(A|\mathcal{G}) dP, \qquad \text{for all } G \in \mathcal{G}. \tag{11}$$

The random variable $P(A|\mathcal{G})$ is called the probability of A conditional on $\mathcal{G}$. Instead of a σ-algebra we may condition on some random vector $X : \Omega \to \mathbb{R}^N$. Let $\sigma(X)$ be the smallest σ-algebra for which X is measurable. Define $P(A|X)$ by $P(A|X) \equiv P(A|\sigma(X))$. We call $P(A|X)$ the probability of A conditional on X. See Billingsley (1986) for more on the intuition behind conditional probabilities as well as a proof that such a random variable always exists. We can now proceed to our proof of Theorem 4.

A.3.2 Step 1

Let $\mathbf{D}^T = (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)_{t\leq T}$ be a random data set. For a set of observations $t, v, \ldots, s \leq T$, let $\mathbf{D}^T_{-t,-v,\ldots,-s}$ be the data set formed by removing observations $t, v, \ldots, s$. That is, $\mathbf{D}^T_{-t,-v,\ldots,-s} = (\mathbf{r}^\ell, \mathbf{p}^\ell, \mathbf{m}^\ell)_{\substack{\ell\leq T \\ \ell\notin\{t,v,\ldots,s\}}}$.

The following theorem provides a set of sufficient conditions such that the asymptotic power result holds.

Theorem 5. *Suppose the data set* $\mathbf{D}^T = (\mathbf{q}^t, \mathbf{p}^t, \mathbf{m}^t)_{t\leq T}$ *is generated by an approximate utility maximizer characterized by the pair* (U, e) *with* $e \in (0, 1]$ *and suppose there is a number* $0 < \pi < 1$ *so that the following two conditions hold:*

1. *For any* $T \geq 2$ *and any two distinct pairs of observations* $t, v \leq T$*:*

$$P\Big(\tau\left((\mathbf{r}^v, \mathbf{p}^t, \mathbf{m}^t), (\mathbf{r}^t, \mathbf{p}^v, \mathbf{m}^v)\right) \geq e \Big| \mathbf{D}^T_{-t,-v}\Big) \leq \pi \text{ with probability 1.}$$

2. *For any* $T \geq 3$ *and any distinct triples of observations* $t, v, s \leq T$*:*

$$P\Big(\tau\left((\mathbf{r}^v, \mathbf{p}^t, \mathbf{m}^t), (\mathbf{r}^s, \mathbf{p}^v, \mathbf{m}^v)\right) \geq e \Big| \mathbf{D}^T_{-t,-v,-s}\Big) \leq \pi \text{ with probability 1.}$$

Then for any significance level $\alpha \in (0, 1)$*, the probability with which Procedure 1 rejects* H_0 *goes to* 1 *as* $T \to \infty$*. That is,* $\lim_{T\to\infty} P(\phi_\alpha(\mathbf{D}^T) = 1) = 1$*.*

Conditions 1 and 2 in Theorem 5 require that, with some positive probability, certain permuted sub-data-sets violate e-GARP (regardless of the values taken by the remaining observations). Provided this violation is possible (i.e. occurs with positive probability), we can show that the asymptotic power result of the proposition holds.

Proof. Suppose, for all T, the data set $\mathbf{D}^T$ is generated by an approximate utility maximizer characterized by the pair (U, e) with $e \in (0, 1]$. Let $(\Pi, 2^\Pi, \mathbb{Q})$ denote the uniform probability space on Π, i.e. Q is the probability measure on $(\Pi, 2^\Pi)$ such that, for all $S \subseteq \Pi$:[15]

$$Q(\boldsymbol{\sigma} \in S) = \frac{|S|}{T!}.$$

[15] Here we use the bold notation $\boldsymbol{\sigma}$ to denote the random variable $\boldsymbol{\sigma} : \Pi \to \Pi$ where $\boldsymbol{\sigma}(\sigma) = \sigma$ for all $\sigma \in \Pi$.

Our aim is to construct an upper bound on $P(\phi_\alpha(\mathbf{D}^T) = 0)$ that converges to 0 as $T \to \infty$. Notice that, by Theorem 2:

$$1 = P(\mathbf{D}^T \text{ satisfies } e\text{-GARP}) = P(\tau(\mathbf{D}^T) \geq e).$$

As such:

$$\begin{aligned}
P(\phi_\alpha(\mathbf{D}^T) = 0) &= P\left(\phi_\alpha(\mathbf{D}^T) = 0 \middle| \tau(\mathbf{D}^T) \geq e\right) \\
&\leq P\left(\frac{\left|\sigma \in \Pi : \tau(\mathbf{D}^T_\sigma) \geq \tau(\mathbf{D}^T)\right|}{|\Pi|} > \alpha \middle| \tau(\mathbf{D}^T) \geq e\right) \\
&\leq P\left(\frac{\left|\sigma \in \Pi : \tau(\mathbf{D}^T_\sigma) \geq e\right|}{|\Pi|} > \alpha\right) \\
&\leq \mathbb{E}_P\left[\mathbb{E}_Q\left[\mathbb{1}\left(\tau(\mathbf{D}^T_{\boldsymbol{\sigma}}) \geq e\right)\right] \geq \alpha\right],
\end{aligned}$$

where we use $\mathbb{E}_P$ to denote the expectation with respect to P and $\mathbb{E}_Q$ to denote the expectation with respect to Q. Next, applying Markov's inequality gives:

$$\begin{aligned}
P(\phi_\alpha(\mathbf{D}^T) = 0) &\leq \frac{1}{\alpha}\mathbb{E}_P\big[\mathbb{E}_Q\left[\mathbb{1}\left(\tau(\mathbf{D}^T_{\boldsymbol{\sigma}}) \geq e\right)\right]\big] \\
&= \frac{1}{\alpha}\mathbb{E}_Q\big[\mathbb{E}_P\left[\mathbb{1}\left(\tau(\mathbf{D}^T_{\boldsymbol{\sigma}}) \geq e\right)\right]\big],
\end{aligned}$$

where the last equality follows from exchanging integration and summation (as Π is finite).

We say that a permutation $\sigma \in \Pi$ has n fixed points if there are exactly n observations $t \in \{1, \dots, T\}$ such that $\sigma(t) = t$. Let Π_n be the set of permutations with n fixed points.

Then:

$$
\begin{aligned}
P(\phi_\alpha(\mathbf{D}^T)=0) &\leq \frac{1}{\alpha}\mathbb{E}_Q\big[\mathbb{E}_P\left[\mathbb{1}\left(\tau(\mathbf{D}^T_{\boldsymbol{\sigma}})\geq e\right)\right]\big] \\
&= \frac{1}{\alpha}\sum_{n=0}^{T} Q(\Pi_n)\sum_{\sigma\in\Pi_n}\frac{1}{|\Pi_n|}\mathbb{E}_P\left[\mathbb{1}\left(\tau(\mathbf{D}^T_\sigma)\geq e\right)\right] \\
&\leq \frac{1}{\alpha}\sum_{n=0}^{T} Q(\Pi_n)\sum_{\sigma\in\Pi_n}\frac{1}{|\Pi_n|}\pi^{\frac{T-n}{4}} && \text{(by Lemma 2)} \\
&\leq \frac{1}{\alpha}\sum_{n=0}^{T}\frac{1}{n!}\pi^{\frac{T-n}{4}} && \text{(by Lemma 3)} \\
&\leq \frac{1}{\alpha}\pi^{\frac{T}{4}}\sum_{n=0}^{\infty}\frac{\pi^{-\frac{n}{4}}}{n!} \\
&= \frac{1}{\alpha}\pi^{\frac{T}{4}}\exp\left(\pi^{-\frac{1}{4}}\right)\overset{T\to\infty}{\to}0,
\end{aligned}
$$

which completes the proof. □

A.3.3 Step 2

In the second step, we demonstrate the following result:

Theorem 6. *Assume that* $\mathbf{D}^T=(\mathbf{r}^t,\mathbf{p}^t,\mathbf{m}^t)_{t\leq T}$ *is generated by an approximate utility maximizer. Further assume that for all open sets* $B\subseteq\mathbb{R}^L_{++}\times\mathbb{R}_{++}$ *and all (relative) open subsets* $R\subseteq\Delta^{L-1}$ *satisfying* $\mu(p,m)\subseteq R$ *for all* $(p,m)\in B$ *there exists an* $a>0$ *such that for all* $T\in\mathbb{N}$ *and all* $t\leq T$*:*

$$
P\left((\mathbf{r}^t,\mathbf{p}^t,\mathbf{m}^t)\in R\times B|\mathbf{D}^T_{-t}\right)>a \text{ with probability 1.}
$$

Then for any significance level $\alpha\in(0,1)$*, the probability with which Procedure 1 rejects* H_0 *goes to* 1 *as* $T\to\infty$*. That is,* $\lim_{T\to\infty}P(\phi_\alpha(\mathbf{D}^T)=1)=1$*.*

Theorem 6 relaxes the IID assumption of Theorem 4 and instead only requires that any triple $(\mathbf{r}^t,\mathbf{p}^t,\mathbf{m}^t)$ can be observed with a positive probability provided that the corresponding $\mathbf{q}^t$ is arbitrarily close to maximizing utility. In contrast, the full support assumption is maintained.

Proof. The proof proceeds by showing that the conditions of Theorem 6 imply that conditions 1 and 2 of Theorem 5 hold.

For condition 1, consider the open sets $B^1,B^2\subseteq\mathbb{R}^L_{++}\times\mathbb{R}_{++}$ and the (relatively) open sets $R^1,R^2\subseteq\Delta^{L-1}$ for which the existence is guaranteed by Lemma 6 for the particular efficiency level $e \in (0,1]$. As $\mu(p,m) \subseteq R^1$ for all $(p,m) \in B^1$ and $\mu(p,m) \in R^2$ for all $(p,m) \in B^2$, we have from Lemma 8 that there is a $b > 0$ such that for all $T \geq 2$ and all $t, v \leq T$:

$$P\left(\begin{array}{l} (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t) \in R^1 \times B^1 \\ (\mathbf{r}^v, \mathbf{p}^v, \mathbf{m}^v) \in R^2 \times B^2 \end{array} \quad \text{and} \quad \middle| \mathbf{D}^T_{-t,-v} \right) > b \text{ with probability 1.}$$

If we choose π such that:

$$0 < 1 - \pi < b,$$

we have that condition 1 of Theorem 5 holds if we can show that:

$$\tau\big((r^t, p^v, m^v), (r^v, p^t, m^t)\big) < e, \text{ for all } (r^t, p^t, m^t) \in R^1 \times B^1, (r^v, p^v, m^v) \in R^2 \times B^2. \tag{12}$$

Using (17) in Lemma 6 we can show that (r^t, p^v, m^v) is "e-revealed preferred" to (r^v, p^t, m^t) in the sense that:

$$\begin{aligned} & e > \frac{m^t}{m^v} \frac{p^v \cdot r^v}{p^t \cdot r^v}, \\ \iff \quad & e > \frac{p^v \cdot \gamma(r^v, p^t, m^t)}{m^v}, \\ \iff \quad & e\, m^v > p^v \cdot \gamma(r^v, p^t, m^t), \end{aligned}$$

where

$$\gamma(r, p, m) = m \frac{r}{p \cdot r}$$

gives the demanded quantities given the ray r, price vector p and income level m. A similar argument shows that (r^v, p^t, m^t) is e-revealed preferred to (r^t, p^v, m^v) and thus condition (12) holds.

To show condition 2 of Theorem 5, as before let R^1, R^2 and B^1, B^2 be the sets of which existence is guaranteed by Lemma 6 for the particular efficiency level $e \in (0,1]$. From part 2 of Lemma 8, we have that there is a $b > 0$ such that, for all $T \geq 3$ and

$t, v, s \leq T$:

$$P\left((\mathbf{p}^t, \mathbf{m}^t) \in B^1 \text{ and } (\mathbf{r}^v, \mathbf{p}^v, \mathbf{m}^v) \in R^2 \times B^2 \text{ and } \mathbf{r}^s \in R^1 \middle| \mathbf{D}^T_{-t,-v,-s}\right)$$
$$\geq P\left(\begin{array}{ll} (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t) \in R^1 \times B^1 & \text{and} \\ (\mathbf{r}^v, \mathbf{p}^v, \mathbf{m}^v) \in R^2 \times B^2 & \text{and} \\ (\mathbf{r}^s, \mathbf{p}^s, \mathbf{m}^s) \in R^1 \times B^1 & \end{array} \middle| \mathbf{D}^T_{-t,-v,-s} \right) > b \text{ with probability 1.}$$

Given $0 < 1 - \pi < b$, condition 2 of Theorem 5 holds if we can show that:

$$\tau\left((r^v, p^t, m^t), (r^s, p^v, m^v)\right) < e, \text{ for all,}$$
$$(p^t, m^t) \in B^1, (r^v, p^v, m^v) \in R^2 \times B^2, r^s \in R^1. \quad (13)$$

Using (17) in Lemma 6 we can show that (r^v, p^t, m^t) is e-revealed preferred to (r^s, p^v, m^v). Indeed:

$$\begin{aligned} & e > \frac{m^v}{m^t} \frac{p^t \cdot r^s}{p^v \cdot r^s}, \\ \iff \quad & e > \frac{p^t \cdot \gamma(r^s, p^v, m^v)}{m^v}, \\ \iff \quad & e\, m^t > p^t \cdot \gamma(r^s, p^v, m^v). \end{aligned}$$

A similar argument can be used to show that (r^s, p^v, m^v) is e-revealed preferred to (r^v, p^t, m^t), which shows (13). □

A.3.4 Step 3

We finalize the proof of Theorem 4 by showing that the conditions of Theorem 6 hold. Let B be an open subset of $\mathbb{R}^L_{++} \times \mathbb{R}_{++}$ and let R be a (relatively) open subsets of Δ^{L-1} such that:

$$\forall (p, m) \in B : \mu(p, m) \subseteq R.$$

Then for $T \in \mathbb{N}$, $t \leq T$, with probability 1:

$$\begin{aligned} P\left((\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t) \in R \times B \middle| \mathcal{D}^T_{-t}\right) &= P\left((\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t) \in R \times B\right), \\ &= P\left(\mathbf{r}^t \in R | (\mathbf{p}^t, \mathbf{m}^t) \in B\right) P\left((\mathbf{p}^t, \mathbf{m}^t) \in B\right) \equiv \gamma > 0 \end{aligned}$$

The equivalences follow from the IID assumption and the definition of conditional probabilities. The last inequality follows from Assumption 3 and the full support assumption. From the IID assumption, we notice that γ only depends on R and B, but not on the particular values of T or t.

A.4 Proof of Theorem 3

Let $\mathbf{D}^T = (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)_{t \leq T}$ be a random data set. Let F be the distribution of the first observation $(\mathbf{r}^1, \mathbf{p}^1, \mathbf{m}^1)$. Let $\tilde{\mathbf{D}}^{\tilde{T}} = (\tilde{\mathbf{q}}^t, \tilde{\mathbf{p}}^t, \tilde{\mathbf{m}}^t)_{t \leq \tilde{T}}$ be the random data set composed of $\tilde{T}$ IID draws from F. Clearly, if $\mathbf{D}^T$ is generated by a random consumer then so is $\tilde{\mathbf{D}}^{\tilde{T}}$, and if $\mathbf{D}^T$ is generated by an approximate utility maximizer characterized by the pair (U, e) then so is $\tilde{\mathbf{D}}^{\tilde{T}}$.

From Theorem 1 it follows that if the random data set $\mathbf{D}^T$ is generated by a random consumer, then:

$$P(\phi_\alpha(\tilde{\mathbf{D}}^{\tilde{T}}) = 1) \leq \alpha,$$

for all $\tilde{T} \in \mathbb{N}$ and $\alpha \in (0, 1)$.

On the other hand, from Theorem 4 it follows that if $\mathbf{D}^T$ is generated by an approximate utility maximizer, then:

$$\lim_{\tilde{T} \to \infty} P(\phi_\alpha(\tilde{\mathbf{D}}^{\tilde{T}}) = 1) = 1,$$

for all $\alpha \in (0, 1)$. As these two conditions are mutually exclusive we see that $\mathbf{D}^T$ cannot be generated by both a random consumer and an approximate utility maximizer.

A.5 Lemmata

In what follows $(\Omega, \mathcal{F}, P)$ denotes the underlying probability space.

Lemma 1. *Let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$ let $A \in \mathcal{G}$ $B \in \mathcal{F}$, $P(A) > 0$ and $P(B|\mathcal{G}) \leq \pi$ with probability 1. Then:*

$$P(B|A) \leq \pi.$$

Proof. From the definition of a conditional probability (equation (11)) we have:

$$P(B \cap A) = \int_A P(B|\mathcal{G}) \, dP \leq \pi P(A).$$

On the other hand, as $P(A) > 0$:

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \leq \pi \frac{P(A)}{P(A)} = \pi.$$

□

Lemma 2. *Suppose $\mathbf{D}^T$ is satisfies the conditions of Theorem 5. Then there exists a* *number* $\pi \in (0,1)$ *so that for any* $\sigma \in \Pi$*:*

$$P(\tau(\mathbf{D}_\sigma^T) \geq e) \leq \pi^{\frac{T-n}{4}},$$

where $n = |\{i \leq T : \sigma(i) = i\}|$ *is the number of fixed points of* σ*.*

Proof. Any permutation can be decomposed into an exhaustive set of disjoint cycles. As there are T elements in total in the set $\{1, \dots, T\}$, the maximum length of a cycle in σ is T. Also, the statement of the lemma assumes that there are n cycles of length 1. Let us denote by C_m the number of cycles in the permutation σ of length m. Then calculating the elements by cycle gives:

$$T = \sum_{k=1}^{T} kC_k = n + \sum_{k=2}^{T} kC_k. \tag{14}$$

Consider a cycle of length 2. As this cycle has no fixed points, it must take the form:

$$\widetilde{\mathbf{D}} = \Big((\mathbf{r}^{\sigma(i)}, \mathbf{p}^i, \mathbf{m}^i), (\mathbf{r}^i, \mathbf{p}^{\sigma(i)}, \mathbf{m}^{\sigma(i)})\Big),$$

for some $i \leq T$.

Next, any cycle of length $k \geq 3$ allows for constructing $\lfloor \frac{k}{3} \rfloor$ non-overlapping data sets of size 3 that take the form:

$$\bigg(\Big(\mathbf{r}^{\sigma(i)}, \mathbf{p}^i, \mathbf{m}^i\Big), \Big(\mathbf{r}^{\sigma(\sigma(i))}, \mathbf{p}^{\sigma(i)}, \mathbf{m}^{\sigma(i)}\Big), \Big(\mathbf{r}^{\sigma(\sigma(\sigma(i)))}, \mathbf{p}^{\sigma(\sigma(i))}, \mathbf{m}^{\sigma(\sigma(i))}\Big)\bigg),$$

for some $i \leq T$ and $\lfloor a \rfloor$ denoting the greatest integer below a. Consider the data subsets generated from these three element data sets by dropping the last observation:

$$\overline{\mathbf{D}} = \bigg(\Big(\mathbf{r}^{\sigma(i)}, \mathbf{p}^i, \mathbf{m}^i\Big), \Big(\mathbf{r}^{\sigma(\sigma(i))}, \mathbf{p}^{\sigma(i)}, \mathbf{m}^{\sigma(i)}\Big)\bigg).$$

Clearly, all these data sets $\overline{\mathbf{D}}$ have no indices in common with the other data sets constructed from the same cycle. Let us enumerate the constructed sub data sets $\widetilde{\mathbf{D}}$ and $\overline{\mathbf{D}}$ of size 2 by $\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_K$. For each $k \leq K$ let E_k be the event that $\tau(\mathbf{D}_k) \geq e$. Then:

$$P(\tau(\mathbf{D}_\sigma) \geq e) \leq P\Big(E_1 \cap \ldots \cap E_K\Big).$$

Let us show that the right hand side is less than or equal to π^K. First, if $P(E_1 \cap E_2 \cap$

$\ldots \cap E_k) = 0$ then immediately:

$$P(\tau(\mathbf{D}_\sigma) \geq e) \leq \pi^K.$$

On the other hand, if $P(E_1 \cap E_2 \cap \ldots \cap E_k) > 0$ then we have:

$$\begin{aligned} P(\tau(\mathbf{D}_\sigma) \geq e) &\leq P\Big(E_1 \cap \ldots \cap E_K\Big) \\ &= \prod_{k=1}^{K} P\Big(E_k | E_{k-1} \cap \ldots \cap E_1\Big) \leq \pi^K, \end{aligned}$$

where the last line follows from Lemma 1 and conditions 1 and 2 in Theorem 5.

Next, the total number of both types of data subsets of size 2 is bounded from below as follows:

$$\begin{aligned} K = C_2 + \sum_{k=3}^{T} C_k \left\lfloor \frac{k}{3} \right\rfloor &\geq C_2 + \sum_{k=3}^{T} C_k \frac{k}{4} \\ &\geq \frac{1}{4} \sum_{k=2}^{T} k\, C_k \\ &= \frac{1}{4}(T - n). && \text{(by equation (14))} \end{aligned}$$

The first inequality follows from the fact that, for $k \geq 3$, $\lfloor k/3 \rfloor \geq k/4$. The result follows. $\square$

Lemma 3. *We have*

$$Q(\Pi_n) \leq \frac{1}{n!}. \tag{15}$$

Proof. A derangement is defined as a permutation with no fixed points. Let $!m$ be the number of derangements of $\{1, \ldots, m\}$. It directly follows that $!m/m! \leq 1$ for all $m \in \mathbb{N}$ and, thus:

$$Q(\Pi_n) = \frac{\binom{N}{n}\big(!(N-n)\big)}{N!} = \frac{!(N-n)}{n!((N-n)!)} \leq \frac{1}{n!}.$$

Here, we counted the number of elements in Π_n by first counting the possible ways to pick n fixed points and then, for each set of fixed points, counting the possible ways in which the remaining elements can be deranged. $\square$

Lemma 4. *Let $U : \mathbb{R}^L_+ \to \mathbb{R}$ be a well-behaved utility function and $m > 0$ an income level. For all $\delta > 0$ and $j \in \{1, 2\}$, there exists a price vector $\overline{p} \in \mathbb{R}^L_{++}$ with $\overline{p}_j < \delta$ and* *for all $\ell \neq j$: $\overline{p}_\ell = Lm$, such that that for all optimal demands $q \in x(\overline{p}, m)$ we have $q_j > \frac{1}{\delta}$.*

Proof. Fix an income level $m > 0$. Towards a contradiction, assume that there exists a $\delta > 0$ and a $j \in \{1, 2\}$ such that for all $\overline{p} \in \mathbb{R}^n_{++}$ with $\overline{p}_j < \delta$ and $\overline{p}_\ell = Lm$ (for $\ell \neq j$) there is an optimal demand $q \in x(\overline{p}, m)$ such that $q_j \leq \frac{1}{\delta}$.

Let $(p^n)_{n\in\mathbb{N}}$ be the sequence of prices in $\mathbb{R}^L_{++}$ with:

$$p^n_\ell = \begin{cases} 1/n & \text{if } \ell = j, \\ Lm & \text{if } \ell \neq j. \end{cases}$$

Then by assumption, for all $n \geq 1/\delta$, we can find an optimal bundle $q^n \in x(p^n, m)$, such that $q^n_j \leq \frac{1}{\delta}$. Given that for all $\ell \neq j$, $q^n_\ell \leq \frac{1}{L}$, we have that the sequence of bundles $(q^n)_{n\in\mathbb{N}}$ is bounded. By proceeding along a sub-sequences, if necessary, we may therefore assume that $\lim_{n\to\infty} q^n = \tilde{q} \in \mathbb{R}^L_+$. In particular, $\tilde{q}_j \leq 1/\delta$.

Next, note that for all $n \in \mathbb{N}$:

$$\{q \in \mathbb{R}^L_+ : p^n \cdot q \leq m\} \subseteq \{q \in \mathbb{R}^L_+ : p^{n+1} \cdot q \leq m\},$$

and so, by optimality of the bundles q^n, we must also have that:

$$U(q^n) \leq U(q^{n+1}) \leq \ldots \leq U(\tilde{q}), \tag{16}$$

where the last inequality uses continuity of the utility function U. Then from the continuity and strict increasing property of U there must be a $\varepsilon > 0$ and a $k \neq j$, such that the bundle $\tilde{q}'$ where:

$$\tilde{q}'_\ell = \begin{cases} \tilde{q}_j + 1 & \text{if } \ell = j, \\ \tilde{q}_k - \varepsilon & \text{if } \ell = k, \\ \tilde{q}_\ell & \text{else,} \end{cases}$$

satisfies $U(\tilde{q}') > U(\tilde{q})$.

From the definition of p^n, we can find a number D such that $\sup_n \|p^n\| < D$. Take $N \in \mathbb{N}$ large enough such that for all $n \geq N$ (such an N exists as $q^N \to \tilde{q}$):

$$\frac{1}{n} + D\|\tilde{q} - q^n\| < \varepsilon Lm.$$

Then for all $n \geq N$, we have that:

$$\begin{aligned}
p^n \cdot \tilde{q}' &= \sum_{\ell=1}^{L} p^n_\ell \tilde{q}'_\ell, \\
&= -\varepsilon p^n_k + p^n_j + \sum_{\ell=1}^{L} p^n_\ell \tilde{q}_\ell, \\
&= -\varepsilon Lm + 1/n + p^n \cdot q^n + p^n \cdot (\tilde{q} - q^n), \\
&\leq -\varepsilon Lm + 1/n + m + \|p^n\| \|\tilde{q} - q^n\|, \\
&\leq -\varepsilon Lm + 1/n + D\|\tilde{q} - q^n\| + m < m.
\end{aligned}$$

Here the first inequality uses the Cauchy-Schwartz inequality together with the fact that $p^n \cdot q^n = m$. The final line shows that $p^n \cdot \tilde{q}' < m$, which contradicts the fact that q^n is optimal for (p^n, m) (i.e. $q^n \in x(p^n, m)$) and $U(q^n) \leq U(\tilde{q}) < U(\tilde{q}')$. This gives the desired contradiction. □

Lemma 5. *Let $U : \mathbb{R}^L_+ \to \mathbb{R}$ be well-behaved utility function and let $e > 0$. There exists $(p^1, m^1), (p^2, m^2) \in \mathbb{R}^L_{++} \times \mathbb{R}_{++}$ which satisfy:*

$$\begin{aligned}
&\frac{m^2}{m_1} \frac{p^1 \cdot r^1}{p^2 \cdot r^1} < e, \qquad \textit{for all } r^1 \in \mu(p^1, m^1) \textit{ and },\\
&\frac{m^1}{m_2} \frac{p^2 \cdot r^2}{p^1 \cdot r^2} < e, \qquad \textit{for all } r^2 \in \mu(p^2, m^2)
\end{aligned}$$

Proof. Fix $m_1, m_2 > 0$. From Lemma 4, we know that for all $\delta > 0$ there exists prices p^1, p^2 with for $j, k \in \{1, 2\}$: $p^j_k = Lm^j$ if $j \neq k$ and such that for all $q^j \in x(p^j, m^j)$:

$$q^j_j > \tfrac{1}{\delta},$$

Then for $j, k \in \{1, 2\}$ with $j \neq k$:

$$\begin{aligned}
p^j \cdot q^k &= \sum_{\ell=1}^{L} p^j_\ell q^k_\ell, \\
&= p^j_k q^k_k + \sum_{\ell \neq k} p^j_\ell q^k_\ell, \\
&\geq p^j_k q^k_k > \frac{Lm^j}{\delta}.
\end{aligned}$$

As such:

$$\frac{m^j}{p^j \cdot q^k} < \frac{m^j \delta}{L m^j} = \frac{\delta}{L} = e,$$

where we have chosen the value $\delta = Le$. Using the transformation from bundles to rays:

$$q^k = m^k \frac{r^k}{p^k \cdot r^k},$$

we obtain that for $j, k \in \{1, 2\}$ with $j \neq k$:

$$\frac{m^j}{m^k} \frac{p^k \cdot r^k}{p^j \cdot r^k} < e,$$

as we wanted to show. □

Lemma 6. *Let $U : \mathbb{R}^L_+ \to \mathbb{R}$ be a well-behaved utility function. For all $e > 0$, there exists non-empty open sets of price income combinations $B^1, B^2 \subseteq \mathbb{R}^L_{++} \times \mathbb{R}_{++}$ and (relatively) open sets of rays $R^1, R^2 \subseteq \Delta^{L-1}$ such that:*

1. *For all $k = 1, 2$ and $(p^k, m^k) \in B^k$:*

$$\mu(p^k, m^k) \subseteq R^k.$$

2. *For $k, j \in \{1, 2\}$ with $k \neq j$:*

$$\frac{m^j}{m^k} \frac{p^k \cdot r^k}{p^j \cdot r^k} < e \text{ for all } (p^k, m^k) \in B^k, (p^j, m^j) \in B^j, r^k \in R^k. \tag{17}$$

Proof. Given $e > 0$, we can find, by Lemma 5, budgets $(p^1, m^1), (p^2, m^2)$ which satisfy for $j, k \in \{1, 2\}$ and $j \neq k$:

$$\frac{m^j}{m^k} \frac{p^k \cdot r^k}{p^j \cdot r^k} < \frac{e}{2}, \qquad \text{for all } r^k \in \mu(p^k, m^k). \tag{18}$$

Let B^k_δ be the open ball of radius $\delta > 0$ around (p^k, m^k) and let:[16]

$$R^k_\delta = \bigcup_{(\tilde{p}^k, \tilde{m}^k) \in B^k_\delta} \mu(\tilde{p}^k, \tilde{m}^k)$$

be the union of all optimal rays over the budgets in B^k_δ.

[16]We choose δ small enough such that $B^k_\delta \subseteq \mathbb{R}^L_{++} \times \mathbb{R}_{++}$.

Let us first show that there exists such number $\delta > 0$ such that for $j, k \in \{1, 2\}$ and $j \neq k$:

$$\frac{\tilde{m}^j}{\tilde{m}^k}\frac{\tilde{p}^k \cdot \tilde{r}^k}{\tilde{p}^j \cdot \tilde{r}^k} < \frac{e}{2}, \qquad \text{for all } (\tilde{p}^j, \tilde{m}^j) \in B^j_\delta, (\tilde{p}^k, \tilde{m}^k) \in B^k_\delta, \tilde{r}^k \in R^k_\delta.$$

Towards a contradiction, if the above is not true, we can find a sequence of budgets $(\tilde{p}^j_n, \tilde{m}^j_n) \in B^j_{1/n}$, $(\tilde{p}^k_n, m^k_n) \in B^k_{1/n}$ and a sequence of rays $\tilde{r}^k_n \in R^k_{1/n}$ such that for all n large enough:

$$\frac{\tilde{m}^j_n}{\tilde{m}^k_n}\frac{\tilde{p}^k_n \cdot \tilde{r}^k_n}{\tilde{p}^j_n \cdot \tilde{r}^k_n} \geq \frac{e}{2}.$$

As for all n large enough, $\tilde{r}^k_n \in R^k_{1/n}$ we can find (using the definition of R^k_δ) a sequence of budgets $(\overline{p}^k_n, \overline{m}^k_n) \in B^k_{1/n}$ such that $\tilde{r}^k_n \in \mu(\overline{p}^k_n, \overline{m}^k_n)$. As $(\overline{p}^k_n, \overline{m}^k_n) \overset{n}{\to} (p^k, m^k)$, we have, by the upper hemi-continuity of the optimal ray correspondence (which follows from Berge's maximization theorem), that, possibly along a subsequence, $\tilde{r}^k_n \overset{n}{\to} r^k \in \mu(p^k, m^k)$. As also $(\tilde{p}^j_n, \tilde{m}^j_n) \overset{n}{\to} (p^j, m^j)$ and $(\tilde{p}^k_n, \tilde{m}^k_n) \overset{n}{\to} (p^k, m^k)$, we obtain:

$$\frac{\tilde{m}^j_n}{\tilde{m}^k_n}\frac{\tilde{p}^k_n \cdot \tilde{r}^k_n}{\tilde{p}^j_n \cdot \tilde{r}^k_n} \overset{n}{\to} \frac{m^j}{m^k}\frac{p^k \cdot r^k}{p^j \cdot r^k} \geq \frac{e}{2},$$

which contradicts (18).

Given this $\delta > 0$, we set $B^1 = B^1_\delta$ and $B^2 = B^2_\delta$ which are both open sets. Although R^1_δ and R^2_δ satisfy both conditions of the lemma, the sets might not be open. However as for all $j, k \in \{1, 2\}$ with $j \neq k$:

$$\sup_{(p^j, m^j) \in B^j, (p^k, m^k) \in B^k, r^k \in R^k_\delta} \frac{m^j}{m^k}\frac{p^k \cdot r^k}{p^j \cdot r^k} \leq \frac{e}{2},$$

we can expand the sets R^k_δ slightly to obtain open sets $R^k \supseteq R^k_\delta$ such that the conditions of the lemma still hold. □

Lemma 7. *Let $\mathcal{G}_1, \mathcal{G}_2, \mathcal{H}$ be sub σ-algebra's of $\mathcal{F}$ such that $\mathcal{H} \subseteq \mathcal{G}_1$ and $\mathcal{H} \subseteq \mathcal{G}_2$. Let $a, b > 0$, $B \in \mathcal{F}$ and $A \in \mathcal{G}_2$ be such that:*

$$P(A|\mathcal{G}_1) \geq a \text{ with probability 1,}$$
$$P(B|\mathcal{G}_2) \geq b \text{ with probability 1.}$$

Then:

$$P(A \cap B|\mathcal{H}) \geq \frac{ab}{2} \text{ with probability 1.}$$

Proof. Towards a contradiction, assume that there is a $C \in \mathcal{H}$ with $P(C) > 0$ such that

$P(A\cap B|\mathcal{H}) < \frac{ab}{2}$ on C. Using (11):

$$P(A\cap B\cap C) = \int_C P(A\cap B|\mathcal{H})\,d\,P \le \frac{ab}{2}P(C).$$

From $\mathcal{H}\subseteq \mathcal{G}_1$, we have that $C\in\mathcal{G}_1$. As such:

$$P(A\cap C) = \int_C P(A|\mathcal{G}_1)\,d\,P \ge aP(C).$$

Next, notice that $A, C\in\mathcal{G}_2$ gives $A\cap C\in\mathcal{G}_2$. Then:

$$P(A\cap B\cap C) = \int_{A\cap C} P(B|\mathcal{G}_2)\,d\,P \ge bP(A\cap C) \ge abP(C),$$

This gives $abP(C)\le \frac{ab}{2}P(C)$ which can only hold if $P(C)=0$, a contradiction. □

Lemma 8. *Assume that the conditions of Theorem 6 hold. Then for all open sets of budgets $B^1, B^2\subseteq \mathbb{R}^L_{++}\times\mathbb{R}_{++}$ and (relatively) open sets of rays $R^1, R^2\subseteq\Delta^{L-1}$, with for $k\in\{1,2\}$:*

$$\mu(p,m)\subseteq R^k \textit{ for all } (p,m)\in B^k:$$

there exists a number $b>0$ such that for all $T\in\mathbb{N}$:

1. *If $T\ge 2$, then for all $t, v\le T$ and for all $j,k\in\{1,2\}$:*

$$P\left(\begin{array}{ll}(\mathbf{r}^t,\mathbf{p}^t,\mathbf{m}^t)\in R^j\times B^j & \textit{and}\\ (\mathbf{r}^v,\mathbf{p}^v,\mathbf{m}^v)\in R^k\times B^k & \end{array}\middle|\,\mathbf{D}^T_{-t,-v}\right) > b \textit{ with probability 1.}$$

2. *If $T\ge 3$, then for all $t,v,s\le T$ and for all $j,k,\ell\in\{1,2\}$:*

$$P\left(\begin{array}{ll}(\mathbf{r}^t,\mathbf{p}^t,\mathbf{m}^t)\in R^j\times B^j & \textit{and}\\ (\mathbf{r}^v,\mathbf{p}^v,\mathbf{m}^v)\in R^k\times B^k & \textit{and}\\ (\mathbf{r}^s,\mathbf{p}^s,\mathbf{m}^s)\in R^\ell\times B^\ell & \end{array}\middle|\,\mathbf{D}^T_{-t,-v,-s}\right) > b \textit{ with probability 1.}$$

Proof. For an observation $\mathbf{r}^t,\mathbf{p}^t,\mathbf{q}^t$ and an element of the sample space $\omega\in\Omega$ let $(\mathbf{r}^t,\mathbf{p}^t,\mathbf{q}^t)(\omega)$ be defined as $(\mathbf{r}^t(\omega),\mathbf{p}^t(\omega),\mathbf{q}^t(\omega))$. The proof of the present lemma is a consequence of Lemma 7. For condition 1, for $j,k\in\{1,2\}$ and $t,v\le T$ set:

$$\begin{aligned}&\mathcal{G}_1=\mathbf{D}^T_{-v},\ \mathcal{G}_2=\mathbf{D}^T_{-t},\ \mathcal{H}=\mathbf{D}^T_{-t,-v},\\ &A=\{\omega\in\Omega:(\mathbf{r}^v,\mathbf{p}^v,\mathbf{m}^v)(\omega)\in R^k\times B^k\},\\ &B=\{\omega\in\Omega:(\mathbf{r}^t,\mathbf{p}^t,\mathbf{m}^t)(\omega)\in R^j\times B^j\}.\end{aligned}$$

Then by assumption:

$$\begin{aligned}
&P(A|\mathcal{G}_1) > a, && \text{with probability 1,}\\
&P(B|\mathcal{G}_2) > a, && \text{with probability 1 and}\\
&A \in \mathcal{G}_2, \mathcal{H} \subseteq \mathcal{G}_1, \mathcal{H} \subseteq \mathcal{G}_2.
\end{aligned}$$

Lemma 7 then gives:

$$P(A \cap B|\mathcal{H}) > \frac{a^2}{2} \text{ with probability 1.}$$

For condition 2, for $j, k, \ell \in \{1, 2\}$ and $t, v, s \leq T$ set:

$$\begin{aligned}
&\tilde{\mathcal{G}}_1 = \mathbf{D}^T_{-s},\ \tilde{\mathcal{G}}_2 = \mathbf{D}^T_{-t,-v},\ \tilde{\mathcal{H}} = \mathbf{D}^T_{-t,-v,-s},\\
&\tilde{A} = \{\omega \in \Omega : (\mathbf{r}^s, \mathbf{p}^s, \mathbf{m}^s)(\omega) \in R^\ell \times B^\ell\},\\
&\tilde{B} = \left\{\omega \in \Omega : \begin{array}{l} (\mathbf{r}^t, \mathbf{p}^t, \mathbf{m}^t)(\omega) \in R^j \times B^j \text{ and} \\ (\mathbf{r}^v, \mathbf{p}^v, \mathbf{m}^v)(\omega) \in R^k \times B^k \end{array}\right\}.
\end{aligned}$$

Notice that $\tilde{B} = A \cap B$. Then:

$$\begin{aligned}
&P(\tilde{A}|\tilde{\mathcal{G}}_1) > a \text{ with probability 1,}\\
&P(\tilde{B}|\tilde{\mathcal{G}}_2) = P(A \cap B|\mathcal{H}) \geq \frac{a^2}{2} \text{ with probability 1, and}\\
&\tilde{A} \in \tilde{\mathcal{G}}_2, \tilde{\mathcal{H}} \subseteq \tilde{\mathcal{G}}_1, \tilde{\mathcal{H}} \subseteq \tilde{\mathcal{G}}_2,
\end{aligned}$$

we can again apply Lemma 7 to obtain:

$$P(\tilde{A} \cap \tilde{B}|\tilde{\mathcal{H}}) \geq \frac{a^3}{4} > 0 \text{ with probability 1.}$$

Setting $b = \frac{a^3}{4}$ completes the proof. □