NONPARAMETRIC ANALYSIS OF RANDOM UTILITY MODELS

YUICHI KITAMURA
Department of Economics, Cowles Foundation, Yale University

JÖRG STOYE
Department of Economics, Cornell University, Department of Economics, University of Bonn, and
Hausdorff Center for Mathematics

This paper develops and implements a nonparametric test of random utility models. The motivating application is to test the null hypothesis that a sample of cross-sectional demand distributions was generated by a population of rational consumers. We test a necessary and sufficient condition for this that does not restrict unobserved heterogeneity or the number of goods. We also propose and implement a control function approach to account for endogenous expenditure. An econometric result of independent interest is a test for linear inequality constraints when these are represented as the vertices of a polyhedral cone rather than its faces. An empirical application to the U.K. Household Expenditure Survey illustrates computational feasibility of the method in demand problems with five goods.

KEYWORDS: Stochastic rationality.

1. INTRODUCTION

THIS PAPER DEVELOPS new tools for the nonparametric analysis of random utility models (RUM). We test the null hypothesis that a repeated cross section of demand data might have been generated by a population of rational consumers, without restricting either unobserved heterogeneity or the number of goods. Equivalently, we empirically test McFadden and Richter's (1991) *axiom of revealed stochastic preference*. To do so, we develop a new statistical test that promises to be useful well beyond the motivating application.

We start from first principles and end with an empirical application. Core contributions made along the way are as follows.

First, the testing problem appears formidable: A structural parameterization of the null hypothesis would involve an essentially unrestricted distribution over all nonsatiated utility functions. However, the problem can, without loss of information, be rewritten as one in which the universal choice set is finite. Intuitively, this is because a RUM only restricts the population proportions with which preferences between different budgets are directly revealed. The corresponding sample information can be preserved in an appropriate discretization of consumption space.

Yuichi Kitamura: yuichi.kitamura@yale.edu
Jörg Stoye: stoye@cornell.edu

More specifically, observable choice proportions must be in the convex hull of a finite (but long) list of vectors. Intuitively, these vectors characterize rationalizable nonstochastic choice types, and observable choice proportions are a mixture over them that corresponds to the population distribution of types. This builds on McFadden (2005), but with an innovation that is crucial for testing: While the set just described is a finite polytope, the null hypothesis can be written as a cone. Furthermore, computing the list of vectors is hard, but we provide algorithms to do so efficiently.

Next, the statistical problem is to test whether an estimated vector of choice proportions is inside a nonstochastic, finite polyhedral cone. This is reminiscent of multiple linear inequality testing and shares with it the difficulty that inference must take account of many nuisance parameters. However, in our setting, inequalities are characterized only implicitly through the vertices of their intersection cone. It is not computationally possible to make this characterization explicit. We provide a novel test and prove that it controls size uniformly over a reasonable class of data generating processes (d.g.p.s) without either computing facets of the cone or resorting to globally conservative approximation. This is a contribution of independent interest that has already seen other applications (Deb, Kitamura, Quah, and Stoye (2018), Hubner (2018), Lazzati, Quah, and Shirai (2018a,b)). Also, while our approach can become computationally costly in high dimensions, it avoids a statistical curse of dimensionality (i.e., rates of approximation do not deteriorate), and our empirical exercise shows that it is practically applicable to at least five-dimensional commodity spaces.

Finally, we leverage recent results on control functions (Imbens and Newey (2009); see also Blundell and Powell (2003)) to deal with endogeneity for unobserved heterogeneity of unrestricted dimension. These contributions are illustrated on the U.K. Family Expenditure Survey, one of the workhorse data sets of the literature. In that data, estimated demand distributions are not stochastically rationalizable, but the rejection is not statistically significant.

The remainder of this paper is organized as follows. Section 2 discusses the related literature. Section 3 lays out the model, develops a geometric characterization of its empirical content, and presents algorithms that allow one to compute this characterization in practice. All of this happens at population level, that is, all identifiable quantities are known. Section 4 explains our test and its implementation under the assumption that one has an estimator of demand distributions and an approximation of its sampling distribution. Section 5 explains how to get the estimator and a bootstrap approximation to its distribution by both smoothing over expenditure and adjusting for endogenous expenditure. Section 6 contains a Monte Carlo investigation of the test's finite sample performance, and Section 7 contains our empirical application. Section 8 concludes. The Supplemental Material (Kitamura and Stoye (2018)) collects all proofs (Appendix A), pseudocode for some algorithms (Appendix B), and some algebraic elaborations (Appendix C).

2. RELATED LITERATURE

Our framework for testing random utility models is built from scratch in the sense that it only presupposes classic results on nonstochastic revealed preference, notably the characterization of individual level rationalizability through the weak (Samuelson (1938)), strong (Houthakker (1950)), or generalized (Afriat (1967)) axioms of revealed preference (WARP, SARP, and GARP henceforth). At the population level, stochastic rationalizability was analyzed in classic work by McFadden and Richter (1991), and updated by McFadden (2005). This work was an important inspiration for us, and we will further clarify the relationship later, but they did not consider statistical testing or attempt to make

the test operational, and could not have done so with computational constraints even of 2005.

An influential related research project is embodied in a sequence of papers by Blundell, Browning, and Crawford (2003, 2007, 2008; BBC henceforth), where the 2003 paper focuses on testing rationality and bounding welfare, and later papers focus on bounding counterfactual demand. BBC assume the same observables that we do and apply their method to the same data, but they analyze a nonstochastic demand system generated by nonparametric estimation of Engel curves. This could be loosely characterized as revealed preference analysis of a representative consumer and, in practice, of average demand. Lewbel (2001) gives conditions on a RUM that ensure integrability of average demand, so BBC effectively add those assumptions to ours. Also, the nonparametric estimation step in practice constrains the dimension of commodity space, which equals three in their empirical applications.[1]

Manski (2007) analyzes stochastic choice from subsets of an abstract, finite choice universe. He states the testing and extrapolation problems in the abstract, solves them explicitly in simple examples, and outlines an approach to nonasymptotic inference. (He also considers models with more structure.) While we start from a continuous problem, the settings become similar after our initial discretization step. However, methods in Manski (2007) will only be practical for choice universes with a handful of elements, an order of magnitude less than in Section 7 herein. In a related paper, Manski (2014) uses our computational toolkit for choice extrapolation.

Our setting much simplifies if there are only two goods, an interesting but obviously very specific case. Blundell, Kristensen, and Matzkin (2014) bound counterfactual demand in this setting through bounding quantile demands. They justify this through an invertibility assumption. Hoderlein and Stoye (2015) show that with two goods, this assumption has no observational implications.[2] Hence, Blundell, Kristensen, and Matzkin (2014) use the same assumptions as we do; however, the restriction to two goods is fundamental. Blundell, Kristensen, and Matzkin (2017) conceptually extend this approach to many goods, in which case invertibility is a restriction. A nonparametric estimation step again limits the dimensionality of commodity space. They apply the method to similar data and the same goods as BBC, meaning that their nonparametric estimation problem is two-dimensional.

Hausman and Newey (2016) nonparametrically bound average welfare under assumptions resembling ours, though their approach additionally imposes smoothness restrictions to facilitate nonparametric estimation and interpolation. Their main identification results apply to an arbitrary number of goods, but the approach is based on nonparametric smoothing; hence, the curse of dimensionality needs to be addressed. The empirical application is to two goods.

With more than two goods, pairwise testing of a stochastic analog of WARP amounts to testing a necessary but not sufficient condition for stochastic rationalizability. This is explored by Hoderlein and Stoye (2014) in a setting that is otherwise ours and also on the same data. Kawaguchi (2017) tests a logically intermediate condition, again on the same data. A different test of necessary conditions was proposed by Hoderlein (2011), who shows that certain features of rationalizable individual demand, like adding up and standard properties of the Slutsky matrix, are inherited by average demand under weak

---

[1]BBC's implementation exploits only WARP and, therefore, a necessary but not sufficient condition for rationalizability. This is remedied in Blundell, Browning, Cherchye, Crawford, De Rock, and Vermeulen (2015).

[2]A similar point is made and exploited, by Hausman and Newey (2016).

conditions. The resulting test is passed by the same data that we use. Dette, Hoderlein, and Neumeyer (2016) propose a similar test using quantiles.

Section 4 of this paper is (implicitly) about testing multiple inequalities, the subject of a large literature in economics and statistics. See, in particular, Gourieroux, Holly, and Monfort (1982) and Wolak (1991), and also Chernoff (1954), Kudo (1963), Perlman (1969), Shapiro (1988), and Takemura and Kuriki (1997) as well as Andrews (1991), Bugni, Canay, and Shi (2015), and Guggenberger, Hahn, and Kim (2008). Furthermore, for the also related setting of inference on parameters defined by moment inequalities, see Andrews and Soares (2010), Bugni (2010), Canay (2010), Chernozhukov, Hong, and Tamer (2007), Imbens and Manski (2004), Romano and Shaikh (2010), Rosen (2008), and Stoye (2009). The major difference from these works is that moment inequalities, if linear (which most of the papers do not assume), define a polyhedron through its faces, while the restrictions generated by our model correspond to its vertices. One cannot in practice switch between these representations in high dimensions, so we had to develop a new approach. This problem also occurs in a related problem in psychology, namely testing whether binary choice probabilities are in the so-called linear order polytope. Here, the problem of computing explicit moment inequalities is researched but unresolved (e.g., see Doignon and Rexhep (2016) and references therein), and we believe that our test is of interest for that literature. Finally, Hoderlein and Stoye (2014) only compare two budgets at a time, and Kawaguchi (2017) tests necessary conditions that are directly expressed as moment inequalities. Therefore, inference in both papers is much closer to the aforementioned literature.

3. ANALYSIS OF POPULATION LEVEL PROBLEM

We now show how to verify rationalizability of a known set of cross-sectional demand distributions on J budgets. The main results are a tractable geometric characterization of stochastic rationalizability and algorithms for its practical implementation.

3.1. *Setting up the Model*

Throughout this paper, we assume the existence of $J < \infty$ fixed budgets $\mathcal{B}_j$ characterized by price vectors $p_j \in \mathbf{R}_+^K$ and expenditure levels $W_j > 0$. Normalizing $W_j = 1$ for now, we can write these budgets as

$$\mathcal{B}_j = \left\{ y \in \mathbf{R}_+^K : p_j' y = 1 \right\}, \quad j = 1, \dots, J.$$

We also start by assuming that the corresponding cross-sectional distributions of demand are known. Thus, assume that demand in budget $\mathcal{B}_j$ is described by the random variable $y(p_j)$. Then we know that

$$P_j(x) := \Pr\big(y(p_j) \in x\big), \quad x \subset \mathbf{R}_+^K, \tag{3.1}$$

for $j = 1, \dots, J$.[3] We will henceforth call $(P_1, \dots, P_J)$ a *stochastic demand system*.

The question is whether this system is rationalizable by a RUM. To define the latter, let

$$u : \mathbf{R}_+^K \mapsto \mathbf{R}$$

---

[3]To keep the presentation simple, here and henceforth we are informal about probability spaces and measurability. See McFadden (2005) for a formally rigorous setup.

denote a utility function over consumption vectors $y \in \mathbf{R}_+^K$. Consider for a moment an individual consumer endowed with some fixed u. Then her choice from a budget characterized by normalized price vector p would be

$$y \in \arg \max_{y \in \mathbf{R}_+^K : p'y \leq 1} u(y), \tag{3.2}$$

with arbitrary tie-breaking if the solution is not unique. For simplicity, we restrict utility functions by monotonicity ("more is better") so that choice is on budget planes, but this is not conceptually necessary.

The RUM postulates that

$$u \sim P_u,$$

that is, u is not constant but is distributed according to a constant (in j) probability law P_u. In our motivating application, P_u describes the distribution of preferences in a population of consumers, but other interpretations are conceivable. For each j, the random variable $y(p_j)$ then is the distribution of y defined in (3.2) that is induced by p_j and P_u. The formal statement follows.

DEFINITION 3.1: The stochastic demand system $(P_1, \ldots, P_J)$ is (*stochastically*) *rationalizable* if there exists a distribution P_u over utility functions u so that

$$P_j(x) = \int 1\left\{\arg \max_{y \in \mathbf{R}_+^K : p'_j y = 1} u(y) \in x\right\} dP_u, \quad x \subset \mathcal{B}_j, j = 1, \ldots, J. \tag{3.3}$$

This model is completely parameterized by P_u, but it only partially identifies P_u because many distinct P_u will induce the same stochastic demand system. We do not place substantive restrictions on P_u; thus we allow for minimally constrained, infinite-dimensional unobserved heterogeneity across consumers.

Definition 3.1 reflects some simplifications that we will drop later. First, W_j and p_j are nonrandom, which is the framework of McFadden and Richter (1991) and others but may not be realistic in applications. In the econometric analysis in Section 5 as well as in our empirical analysis in Section 7, we treat W_j as a random variable that may furthermore covary with u. Also, we initially assume that P_u is the same across price regimes. Once W_j (hence p_j, after income normalization) is a random variable, this is essentially the same as imposing $W_j \perp\!\!\!\perp u$, an assumption we maintain in Section 4 but drop in Section 5 and in our empirical application. However, for all of these extensions, our strategy will be to effectively reduce them to (3.3), so testing this model is at the heart of our contribution.

3.2. *A Geometric Characterization*

The model embodied in (3.3) is extremely general; again, a parameterization would involve an essentially unrestricted distribution over utility functions. However, we next develop a simple geometric characterization of the model's empirical content and, hence, of stochastic rationalizability.

To get an intuition, consider the simplest example in which (3.3) can be tested.

EXAMPLE 3.1: There are two intersecting budgets; thus, $J = 2$ and there exists $y \in \mathbf{R}_{++}^K$ with $p'_1 y = p'_2 y$.
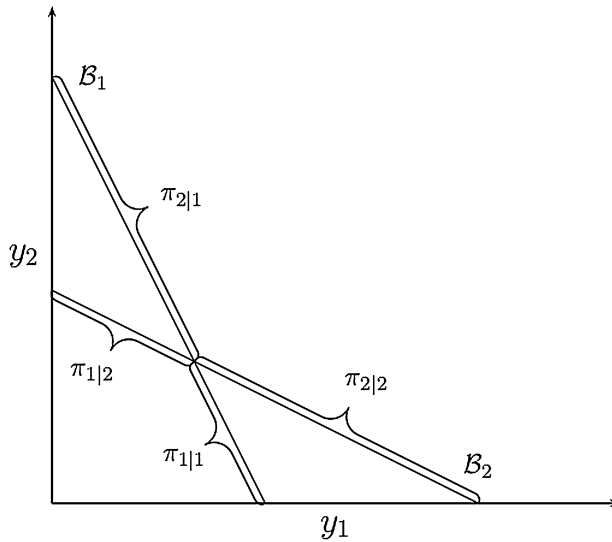
FIGURE 1.—Visualization of Example 3.1.

Consider Figure 1, whose labels will become clear. (The restriction to $\mathbf{R}^2$ is only for the figure.) It is well known that in this example, individual choice behavior is rationalizable unless choice from each budget is below the other budget, in which case a consumer would "revealed prefer" each budget to the other one. Does this restrict repeated cross-sectional choice probabilities? Yes! Supposing for simplicity that there is no probability mass on the intersection of budget planes, it is easy to see (e.g., by applying Fréchet–Hoeffding bounds) that the cross-sectional probabilities of the two line segments labeled $(\pi_{1|1}, \pi_{1|2})$ must not sum to more than 1. This condition is also sufficient (Matzkin (2006)).

Things rapidly get complicated as budgets are added, but the basic insight scales. The only relevant information for testing (3.3) is what fractions of consumers revealed prefer budget j to k for different (j, k). This information is contained in the cross-sectional choice probabilities of the line segments highlighted in Figure 1 (plus, for noncontinuous demand, the intersection). The picture will be much more involved in interesting applications (see Figure 2 for an example), but the idea remains the same. This insight allows one to replace the universal choice set $\mathbf{R}_+^K$ with a finite set and stochastic demand systems with lists of corresponding choice probabilities. But then there are only finitely many rationalizable nonstochastic cross-budget choice patterns. A rationalizable stochastic demand system must be a mixture over them and, therefore, must lie inside a certain finite polytope.

Formalizing this insight requires some notation.

DEFINITION 3.2: Let $\mathcal{X} := \{x_1, \ldots, x_I\}$ be the coarsest partition of $\bigcup_{j=1}^{J} \mathcal{B}_j$ such that for any $i \in \{1, \ldots, I\}$ and $j \in \{1, \ldots, J\}$, x_i is either completely on, completely strictly above, or completely strictly below budget plane $\mathcal{B}_j$. Equivalently, any $y_1, y_2 \in \bigcup_{j=1}^{J} \mathcal{B}_j$ are in the same element of the partition if and only if (iff) $\mathrm{sg}(p_j' y_1 - 1) = \mathrm{sg}(p_j' y_2 - 1)$ for all $j = 1, \ldots, J$.

Elements of $\mathcal{X}$ will be called *patches*. Patches that are part of more than one budget plane will be called *intersection patches*. Each budget can be uniquely expressed as a union

of patches; the number of patches that jointly comprise budget $\mathcal{B}_j$ will be called I_j. Note that $\sum_{j=1}^{J} I_j \geq I$, strictly so (because of multiple counting of intersection patches) if any two budget planes intersect.

REMARK 3.1: Note that $I_j \leq I \leq 3^J$; hence, I_j and I are finite.

The partition $\mathcal{X}$ is the finite universal choice set alluded to earlier. The basic idea is that all choices from a given budget that are on the same patch induce the same directly revealed preferences, so are equivalent for the purpose of our test. Conversely, stochastic rationalizability does not at all constrain the distribution of demand on any patch. Therefore, rationalizability of $(P_1, \ldots, P_J)$ can be decided by only considering the cross-sectional probabilities of patches on the respective budgets. We formalize this as follows.

DEFINITION 3.3: The *vector representation* of $(\mathcal{B}_1, \ldots, \mathcal{B}_J)$ is a $(\sum_{j=1}^{J} I_j)$ vector

$$(x_{1|1}, \ldots, x_{I_1|1}, x_{1|2}, \ldots, x_{I_J|J}),$$

where $(x_{1|j}, \ldots, x_{I_j|j})$ lists all patches comprising $\mathcal{B}_j$. The ordering of patches on budgets is arbitrary but henceforth fixed. Note that intersection patches appear once for each budget that contains them.

DEFINITION 3.4: The *vector representation* of $(P_1, \ldots, P_J)$ is the $(\sum_{j=1}^{J} I_j)$ vector

$$\pi := (\pi_{1|1}, \ldots, \pi_{I_1|1}, \pi_{1|2}, \ldots, \pi_{I_J|J}),$$

where $\pi_{i|j} := P_j(x_{i|j})$.

Thus, the vector representation of a stochastic demand system lists the probability masses that it assigns to patches.

EXAMPLE 3.1—Continued: This example has a total of five patches, namely the four line segments identified in the figure and the intersection. The vector representations of $(\mathcal{B}_1, \mathcal{B}_2)$ and (P_1, P_2) have six components because the intersection patch is counted twice. If one disregards intersection patches (as we do later), the vector representation of (P_1, P_2) is $(\pi_{1|1}, \pi_{2|1}, \pi_{1|2}, \pi_{2|2})$; see Figure 1.

Next, a stochastic demand system is rationalizable iff it is a mixture of rationalizable nonstochastic demand systems. To intuit this, one may literally think of the latter as characterizing rational individuals. It follows that the vector representation of a rationalizable stochastic demand system must be the corresponding mixture of vector representations of rationalizable nonstochastic demand systems. Thus, we state the following definition.

DEFINITION 3.5: The *rational demand matrix* A is the (unique, up to ordering of columns) matrix such that the vector representation of each rationalizable nonstochastic demand system is exactly one column of A. The number of columns of A is denoted H.

REMARK 3.2: We have $H \leq \prod_{j=1}^{J} I_j$; hence, H is finite.

We then can state the following theorem.

THEOREM 3.1: *The following statements are equivalent*:
  (i) *The stochastic demand system* $(P_1, \ldots, P_J)$ *is rationalizable*.
 (ii) *Its vector representation* π *fulfills* $\pi = A\nu$ *for some* $\nu \in \Delta^{H-1}$, *the unit simplex in* $\mathbf{R}^H$.
(iii) *Its vector representation* π *fulfills* $\pi = A\nu$ *for some* $\nu \geq 0$.

Theorem 3.1 reduces the problem of testing (3.3) to testing a null hypothesis about finite (though possibly rather long) vector of probabilities. Furthermore, this hypothesis can be expressed as a finite cone, a simple but novel observation that will be crucial for testing.[4]

We conclude this subsection with a few remarks.

Simplification if Demand Is Continuous

Intersection patches are of lower dimension than budget planes. Thus, if the distribution of demand is continuous, their probabilities must be zero, and they can be eliminated from $\mathcal{X}$. This may considerably simplify A. Also, each remaining patch belongs to exactly one budget plane, so that $\sum_{j=1}^{J} I_j = I$. We impose this simplification henceforth and in our empirical application, but none of our results depends on it.

GARP versus SARP

Rationalizability of nonstochastic demand systems can be defined using either GARP or SARP. SARP will define a smaller matrix A, but nothing else changes. However, columns that are consistent with GARP but not SARP must select at least three intersection patches, so that GARP and SARP define the same A if $\mathcal{X}$ was simplified to reflect continuous demand.

Generality

At its heart, Theorem 3.1 only uses that choice from finitely many budgets reveals finitely many distinct revealed preference relations. Thus, it applies to any setting with finitely many budgets, irrespective of budgets' shapes. For example, the result was applied to kinked budget sets in Manski (2014) and could be used to characterize rationalizable choice proportions over binary menus, that is, the linear order polytope. The result furthermore applies to the "random utility" extension of any other revealed preference characterization that allows for discretization of choice space; see Deb et al. (2018) for an example.

3.3. *Examples*

We next illustrate with a few examples. For simplicity, we presume continuous demand and, therefore, disregard intersection patches.

EXAMPLE 3.1—Continued: Dropping the intersection patch, we have $I = 4$ patches. Index vector representations as in Figure 1. Then the only excluded behavior is $(1, 0,$

---

[4]The idea of patches, as well as the equivalence of (i) and (ii) in Theorem 3.1, were anticipated by McFadden (2005). While the explanation of patches is arguably unclear and (i) ⇔ (ii) is not explicitly pointed out, the idea is unquestionably there. The observation that (ii) ⇔ (iii) (more importantly, the idea of using this for testing) is new.

$1, 0)'$; thus

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{matrix} x_{1|1} \\ x_{2|1} \\ x_{1|2} \\ x_{2|2} \end{matrix}.$$

The column cone of A can be explicitly written as $\{(\nu_1, \nu_2 + \nu_3, \nu_2, \nu_1 + \nu_3)' : \nu_1, \nu_2, \nu_3 \geq 0\}$. As expected, the only restriction on π beyond adding up constraints is that $\pi_{1|1} + \pi_{1|2} \leq 1$.

EXAMPLE 3.2: The following example is the simplest example in which WARP does not imply SARP, so that applying Example 3.1 to all pairs of budgets will only test a necessary condition. More subtly, it can be shown that the conditions in Kawaguchi (2017) are only necessary as well. Let $K = J = 3$ and assume a maximal pattern of intersection of budgets; for example, prices could be $(p_1, p_2, p_3) = ((1/2, 1/4, 1/4), (1/4, 1/2, 1/4), (1/4, 1/4, 1/2))$. Each budget has four patches for a total of $I = 12$ patches, and one can compute

$$A = \begin{pmatrix} 0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&1&0&0&1&0&0&1&0&0 \\ 0&1&0&0&0&0&1&0&0&0&0&1&0&0&0&0&1&0&0&0&0&0&1&0 \\ 1&0&1&1&0&0&0&1&1&1&1&0&1&0&1&0&0&0&1&0&1&0&0&0&1 \\ 0&0&0&0&0&0&0&0&1&0&0&0&0&1&1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&1&1&1&0&0&0&0&0&0 \\ 1&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&1&0&0&0&0 \\ 0&1&1&0&1&1&1&1&0&0&0&1&1&0&0&0&0&0&0&0&0&1&1&1&1 \\ 1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1&1&1&1&1 \end{pmatrix} \begin{matrix} x_{1|1} \\ x_{2|1} \\ x_{3|1} \\ x_{4|1} \\ x_{1|2} \\ x_{2|2} \\ x_{3|2} \\ x_{4|2} \\ x_{1|3} \\ x_{2|3} \\ x_{3|3} \\ x_{4|3} \end{matrix}.$$

(3.4)

Interpreting A requires knowing the geometry of patches. Given the ordering of patches used in the above, the choice of $x_{3|1}$, $x_{2|2}$, and $x_{3|3}$ from their respective budgets would reveal a preference cycle; thus, A does not contain the column $(0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0)'$. We revisit this example in our Monte Carlo study.

EXAMPLE 3.3: Our empirical application has sequences of $J = 7$ budgets in $\mathbf{R}^K$ for $K = 3, 4, 5$ and sequences of $J = 8$ budgets in $\mathbf{R}^3$. Figure 2 visualizes one of the budgets in $\mathbf{R}^3$ and its intersection with six other budgets. There are total of 10 patches (plus 15 intersection patches). The largest A-matrices in the application are of sizes $78 \times 336{,}467$ and $79 \times 313{,}440$. In exploratory work using more complex examples, we also computed a matrix with over 2 million columns.

3.4. *Computing A*

It should be clear now (and we formally show below) that the size of A, hence the cost of computing it, may escalate rapidly as examples get more complicated. We next elaborate how to compute A from a vector of prices $(p_1, \ldots, p_J)$. For ease of exposition, we drop intersection patches (thus SARP = GARP) and add remarks on generalization along the way. We split the problem into two subproblems, namely checking whether a binary "candidate" vector a is in fact a column of A and finding all such vectors.
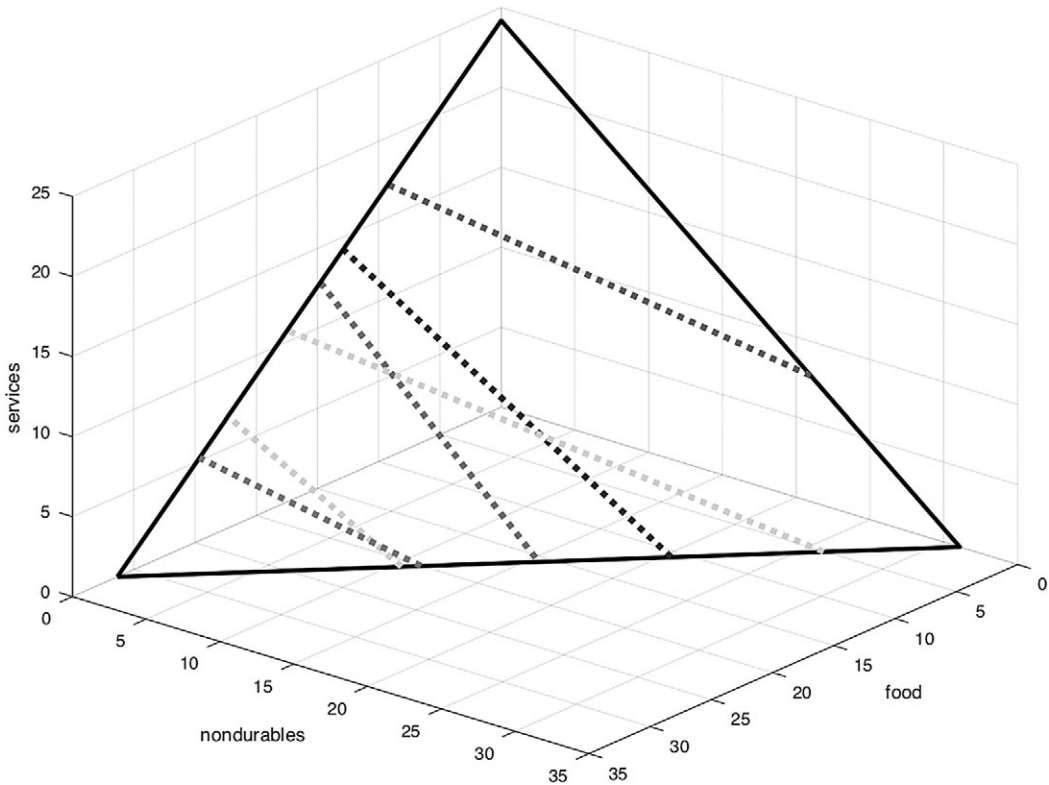
FIGURE 2.—Visualization of one budget set in the empirical application.

3.4.1. *Checking Rationalizability of a*

Consider any binary I-vector a with one entry of 1 on each subvector corresponding to one budget. This vector corresponds to a nonstochastic demand system. It is a column of A if this demand system respects SARP, in which case we call *a rationalizable*.

To check such rationalizability, we initially extract a direct revealed preference relation over budgets. Specifically, if (an element of) $x_{i|j}$ is chosen from budget $\mathcal{B}_j$, then all budgets that are above $x_{i|j}$ are direct revealed preferred to $\mathcal{B}_j$. This information can be extracted extremely quickly.[5]

We next exploit a well known representation: Preference relations over J budgets can be identified with directed graphs on J labeled nodes by equating a directed link from node i to node j with revealed preference for $\mathcal{B}_i$ over $\mathcal{B}_j$. A preference relation fulfills SARP iff this graph is acyclic. This can be tested in quadratic time (in J) through a depth-first or breadth-first search. Alternatively, the Floyd–Warshall algorithm (Floyd (1962)) is theoretically slower but also computes rapidly in our application. Importantly, increasing K does not directly increase the size of graphs checked in this step, though it allows for

---

[5]In practice, we compute a $(I \times J)$ matrix X, where, for example, the ith row of X is $(0, -1, 1, 1, 1)$ if $x_{i|1}$ is on budget $\mathcal{B}_1$, below budget $\mathcal{B}_2$, and above the remaining budgets. This allows us to vectorize construction of direct revealed preference relations, including strict versus weak revealed preference, though we do not use the distinction.

more intricate patterns of overlap between budgets and, therefore, potentially for richer revealed preference relations.

If intersection patches are retained, then one must distinguish between weak and strict revealed preference, and the above procedure tests SARP as opposed to GARP. To test GARP, one could use the Floyd–Warshall or a recent algorithm that achieves quadatic time (Talla Nobibon, Smeulders, and Spieksma (2015)).

3.4.2. *Collecting Rationalizable Vectors*

A total of $\prod_{j=1}^{J} I_J$ vectors a could in principle be checked for rationalizability. Doing this by brute force rapidly becomes infeasible, including in our empirical application. However, these vectors can be usefully identified with the leaves (i.e., the terminal nodes) of a tree constructed as follows: (i) The root of the tree has no label and has I_1 children labeled $(x_{1|1}, \ldots, x_{I_1|1})$. (ii) Each child in turn has I_2 children labeled $(x_{1|2}, \ldots, x_{I_2|2})$ and so on for a total of J generations beyond the root. Then there is a one-to-one mapping from leaves of the tree to conceivable vectors a, namely by identifying every leaf with the nonstochastic demand system that selects its ancestors. Furthermore, each nonterminal node of the tree can be identified with an incomplete a vector that specifies choice only on the first $j < J$ budgets. The methods from Section 3.4.1 can be used to check rationalizability of such incomplete vectors as well.

Our suggested algorithm for computing A is a depth-first search of this tree. Importantly, rationalizability of the implied (possibly incomplete) vector a is checked at each node that is visited. If this check fails, the node *and all its descendants* are abandoned. A column of A is discovered whenever a terminal node has been reached without detecting a choice cycle. Pseudocode for the tree search algorithm is displayed in Appendix B.

3.4.3. *Remarks on Computational Complexity*

The cost of computing A will escalate rapidly under any approach, but some meaningful comparison is possible. To do so, we consider three sequences, all indexed by J, whose first terms are displayed in Table I. First, any two distinct nonstochastic demand systems induce distinct direct revealed preference relations; hence, H is bounded above by $\bar{H}_J$, the number of distinct directed acyclic graphs on J labeled nodes. This sequence—and hence the worst-case cost of enumerating the columns of A, not to mention computing them—is well understood (Robinson (1973)), increases exponentially in J, and is displayed in the first row of Table I.

Next, a worst-case bound on the number of terminal nodes of the aforementioned tree, hence on vectors that a brute force algorithm would check, is $2^{J(J-1)}$. This is simply because the number of conceivable candidate vectors a equals $\prod_{j=1}^{J} I_j$, and every I_j is bounded above by 2^{J-1}. The corresponding sequence is displayed in the last row of Table I.

TABLE I

WORST-CASE BOUNDS[a]

J	1	2	3	4	5	6	7	8
$\bar{H}_J$	1	3	25	543	29,281	3.78×10^6	1.14×10^9	7.84×10^{11}
Tree search	1	4	64	2048	167,936	3.49×10^7	1.76×10^{10}	2.08×10^{13}
Brute force	1	4	64	4096	1,048,576	1.08×10^8	4.40×10^{12}	7.21×10^{16}

[a]The bounds are on the number of rationalizable vectors a and on the number of candidate vectors visited by different algorithms.

Finally, some tedious combinatorial bookkeeping (see Appendix C) reveals that the depth-first search algorithm visits at most $\sum_{j=2}^{J} \bar{H}_{j-1} 2^{j(J+2-j)-2}$ nodes. This sequence is displayed in the middle row of the table, and the gain over a brute force approach is clear.[6]

It is easy to show that the ratio of any two sequences in the table grows exponentially. Also, all of the bounds are in principle attainable (though restricting K may improve them) and are indeed attained in Examples 3.1 and 3.2. In our empirical application, the bounds are far from binding (see Tables IV and V for relevant values of H), but brute force was not always feasible, and the tree search improved on it by orders of magnitude in some cases where it was.

3.4.4. *Further Refinement*

A modest amount of problem-specific adjustment may lead to further improvement. The key to this is contained in the following result.

THEOREM 3.2: *Suppose that not all budget planes mutually intersect; in particular, there exists $M < J$ such that $\mathcal{B}_J$ is either above all or below all of $(\mathcal{B}_1, \ldots, \mathcal{B}_M)$. Suppose also that choices from $(\mathcal{B}_1, \ldots, \mathcal{B}_{J-1})$ are rationalizable. Then choices from $(\mathcal{B}_1, \ldots, \mathcal{B}_J)$ are rationalizable iff choices from $(\mathcal{B}_{M+1}, \ldots, \mathcal{B}_J)$ are.*

If the geometry of budgets allows it (this is particularly likely if budgets move outward over time and even guaranteed if some budget planes are parallel), Theorem 3.2 can be used to construct columns of A recursively from columns of A-matrices that correspond to a smaller J. The gain can be tremendous because, at least with regard to worst-case cost, one effectively moves one or more columns to the left in Table I. A caveat is that application of Theorem 3.2 may require us to manually reorder budgets so that it applies. Also, while the internal ordering of $(\mathcal{B}_1, \ldots, \mathcal{B}_M)$ and $(\mathcal{B}_{M+1}, \ldots, \mathcal{B}_{J-1})$ does not matter, the theorem may apply to distinct partitions of the same set of budgets. In that case, any choice of partition will accelerate computations, but we have no general advice on which is best. We tried the refinement in our empirical application, and it considerably improved computation time for some of the largest matrices. However, the tree search proved so fast that to keep it transparent, our replication code omits this step.

4. STATISTICAL TESTING

This section lays out our statistical testing procedure in the idealized situation where, for finite J, repeated cross-sectional observations of demand over J periods are available to the econometrician. Formally, for each $1 \le j \le J$, suppose one observes N_j random draws of y distributed according to P_j defined in (3.1). Define $N = \sum_{j=1}^{J} N_j$ for later use. Clearly, P_j can be estimated consistently as $N_j \uparrow \infty$ for each j, $1 \le j \le J$. The question is whether the estimated distributions may, up to sampling uncertainty, have arisen from a RUM. We define a test statistic and critical value and show that the resulting test is uniformly asymptotically valid over an interesting range of d.g.p.s.

---

[6]The comparison favors brute force because some nodes visited by a tree search are nonterminal, in which case rationalizability is easier to check. For example, this holds for 16 of the 64 nodes a tree search visits in Example 3.2.

4.1. *Null Hypothesis and Test Statistic*

By Theorem 3.1, we wish to test the following hypothesis.

$(\mathbf{H}_A)$ There exist $\nu \geq 0$ such that $A\nu = \pi$.

This hypothesis is equivalent to

$$(\mathbf{H}_B) \quad \min_{\eta \in \mathcal{C}}[\pi - \eta]'\Omega[\pi - \eta] = 0,$$

where Ω is a positive definite matrix (restricted to be diagonal in our inference procedure) and $\mathcal{C} := \{A\nu | \nu \geq 0\}$ is a convex cone in $\mathbf{R}^I$. The solution η_0 of $(\mathbf{H}_B)$ is the projection of $\pi \in \mathbf{R}_+^I$ onto $\mathcal{C}$ under the weighted norm $\|x\|_\Omega = \sqrt{x'\Omega x}$. The corresponding value of the objective function is the squared length of the projection residual vector. The projection η_0 is unique, but the corresponding ν is not. Stochastic rationality holds if and only if the length of the residual vector is zero.

A natural sample counterpart of the objective function in $(\mathbf{H}_B)$ would be $\min_{\eta \in \mathcal{C}}[\hat{\pi} - \eta]'\Omega[\hat{\pi} - \eta]$, where $\hat{\pi}$ estimates π, for example, by sample choice frequencies. The usual scaling yields

$$\begin{aligned}
\mathcal{J}_N &:= N \min_{\eta \in \mathcal{C}}[\hat{\pi} - \eta]'\Omega[\hat{\pi} - \eta] \\
&= N \min_{\nu \in \mathbf{R}_+^H}[\hat{\pi} - A\nu]'\Omega[\hat{\pi} - A\nu].
\end{aligned} \tag{4.1}$$

Once again, ν is not unique at the optimum, but $\eta = A\nu$ is. Call its optimal value $\hat{\eta}$. Then $\hat{\eta} = \hat{\pi}$ and $\mathcal{J}_N = 0$ if the estimated choice probabilities $\hat{\pi}$ are stochastically rationalizable; obviously, our null hypothesis will be accepted in this case.

4.2. *Simulating a Critical Value*

We next explain how to get a valid critical value for $\mathcal{J}_N$ under the assumption that $\hat{\pi}$ estimates the probabilities of patches by corresponding sample frequencies and that one has R bootstrap replications $\hat{\pi}^{*(r)}$, $r = 1, \ldots, R$. Thus, $\hat{\pi}^{*(r)} - \hat{\pi}$ is a natural bootstrap analog of $\hat{\pi} - \pi$. We will make enough assumptions to ensure that its distribution consistently estimates the distribution of $\hat{\pi} - \pi_0$, where π_0 is the true value of π. The main difficulty is that one cannot use $\hat{\pi}$ as a bootstrap analog of π_0.

Our bootstrap procedure relies on a tuning parameter τ_N chosen such that $\tau_N \downarrow 0$ and $\sqrt{N}\tau_N \uparrow \infty$.[7] Also, we restrict Ω to be diagonal and positive definite, and let $\mathbf{1}_H$ be an H-vector of ones.[8] The restriction on Ω is important: Together with a geometric feature

---

[7]In this section's simplified setting and if $\hat{\pi}$ collects sample frequencies, a reasonable choice would be

$$\tau_N = \sqrt{\frac{\log \underline{N}}{\underline{N}}},$$

where $\underline{N} = \min_j N_j$ and N_j is the number of observations on budget $\mathcal{B}_j$: see (4.8). This choice corresponds to the "BIC choice" (Bayesian information criterion) in Andrews and Soares (2010). We will later propose a different τ_N based on how π is in fact estimated.

[8]In principle, $\mathbf{1}_H$ could be any strictly positive H-vector, though a data-based choice of such a vector is beyond the scope of the paper.

of the column vectors of the matrix A, it ensures that constraints that are fulfilled but with small slack become binding through the *cone tightening* algorithm we are about to describe. A nondiagonal weighting matrix can disrupt this property. For further details on this point and its proof, the reader is referred to Appendix A. Our procedure is as follows:

(i) Obtain the τ_N-*tightened restricted* estimator $\hat{\eta}_{\tau_n}$, which solves

$$\min_{\eta \in \mathcal{C}_{\tau_N}} N[\hat{\pi} - \eta]'\Omega[\hat{\pi} - \eta] = \min_{[\nu - \tau_N \mathbf{1}_H / H] \in \mathbf{R}_+^H} N[\hat{\pi} - A\nu]'\Omega[\hat{\pi} - A\nu].$$

(ii) Define the τ_N-*tightened recentered* bootstrap estimators

$$\hat{\pi}_{\tau_N}^{*(r)} := \hat{\pi}^{*(r)} - \hat{\pi} + \hat{\eta}_{\tau_N}, \quad r = 1, \dots, R.$$

(iii) The bootstrap test statistic is

$$\mathcal{J}_N^{*(r)} = \min_{[\nu - \tau_N \mathbf{1}_H / H] \in \mathbf{R}_+^H} N[\hat{\pi}_{\tau_N}^{*(r)} - A\nu]'\Omega[\hat{\pi}_{\tau_N}^{*(r)} - A\nu]$$

for $r = 1, \dots, R$.

(iv) Use the empirical distribution of $\mathcal{J}_N^{*(r)}$, $r = 1, \dots, R$, to obtain the critical value for $\mathcal{J}_N$.

The object $\hat{\eta}_{\tau_N}$ is the true value of π in the bootstrap population, that is, it is the bootstrap analog of π_0. It differs from $\hat{\pi}$ through a "double recentering." To disentangle the two recenterings, suppose first that $\tau_N = 0$. Then inspection of step (i) of the algorithm shows that $\hat{\pi}$ would be projected onto the cone $\mathcal{C}$. This is a relatively standard recentering "onto the null" that resembles recentering of the J statistic in overidentified generalized method of moments (GMM). However, with $\tau_N > 0$, there is a second recentering because the cone $\mathcal{C}$ itself has been tightened. We next discuss why this recentering is needed.

4.3. *Discussion*

Our testing problem is related to the large literature on inequality testing but adds an important twist. Writing $\{a_1, a_2, \dots, a_H\}$ for the column vectors of A, one has

$$\mathcal{C} = \text{cone}(A) := \{\nu_1 a_1 + \cdots + \nu_H a_H : \nu_h \geq 0\},$$

that is, the set $\mathcal{C}$ is a finitely generated cone. The following result, known as the *Weyl–Minkowski theorem*, provides an alternative representation that is useful for theoretical developments of our statistical testing procedure.[9]

THEOREM 4.1—Weyl–Minkowski Theorem for Cones: *A subset $\mathcal{C}$ of $\mathbf{R}^I$ is a finitely generated cone*

$$\mathcal{C} = \{\nu_1 a_1 + \cdots + \nu_H a_H : \nu_h \geq 0\} \quad \text{for some } A = [a_1, \dots, a_H] \in \mathbf{R}^{I \times H} \qquad (4.2)$$

if and only if it is a finite intersection of closed half-spaces

$$\mathcal{C} = \{t \in \mathbf{R}^I | Bt \leq 0\} \quad \text{for some } B \in \mathbf{R}^{m \times I}. \qquad (4.3)$$

---

[9]See Gruber (2007), Grünbaum, Kaibel, Klee, and Ziegler (2003), and Ziegler (1995), especially Theorem 1.3, for these results and other materials concerning convex polytopes used in this paper.

The expressions in (4.2) *and* (4.3) *are called a* $\mathcal{V}$-*representation (as in vertices) and an* $\mathcal{H}$-*representation (as in half-spaces) of* $\mathcal{C}$, *respectively.*

The "only if" part of the theorem (which is Weyl's theorem) shows that our rationality hypothesis $\pi \in \mathcal{C}$, $\mathcal{C} = \{A\nu | \nu \geq 0\}$ in terms of a $\mathcal{V}$-representation can be reformulated in an $\mathcal{H}$-representation using an appropriate matrix B, at least in theory. If such B were available, our testing problem would resemble tests of

$$H_0: B\theta \geq 0, \quad B \in \mathbf{R}^{p \times q} \text{ is known}$$

based on a quadratic form of the empirical discrepancy between $B\theta$ and η minimized over $\eta \in \mathbf{R}_+^q$. This type of problem has been studied extensively; see references in Section 2. Its analysis is intricate because the limiting distribution of such a statistic depends discontinuously on the true value of $B\theta$. One common way to get a critical value is to consider the globally least favorable case, which is $\theta = 0$. A less conservative strategy widely followed in the econometric literature on moment inequalities is *generalized moment selection* (GMS; see Andrews and Soares (2010), Bugni (2010), and Canay (2010)). If we had the $\mathcal{H}$-representation of $\mathcal{C}$, we might conceivably use the same technique. However, the duality between the two representations is purely theoretical: In practice, B cannot be computed from A in high-dimensional cases like our empirical application.

We therefore propose a *tightening* of the cone $\mathcal{C}$ that is computationally feasible and will have a similar effect as GMS. The idea is to tighten the constraint on ν in (4.1). In particular, define $\mathcal{C}_{\tau_N} := \{A\nu | \nu \geq \tau_N \mathbf{1}_H / H\}$ and define $\hat{\eta}_{\tau_N}$ as

$$
\begin{aligned}
\hat{\eta}_{\tau_N} &:= \operatorname*{argmin}_{\eta \in \mathcal{C}_{\tau_N}} N[\hat{\pi} - \eta]' \Omega [\hat{\pi} - \eta] \\
&= \operatorname*{argmin}_{A\nu:[\nu - \tau_N \mathbf{1}_H / H] \in \mathbf{R}_+^H} N[\hat{\pi} - A\nu]' \Omega [\hat{\pi} - A\nu].
\end{aligned}
\tag{4.4}
$$

Our proof establishes that constraints in the $\mathcal{H}$-representation that are almost binding at the original problem's solution (i.e., their slack is difficult to distinguish from zero at the sample size) will be binding with zero slack after tightening. Suppose that $\sqrt{N}(\hat{\pi} - \pi) \to_d N(0, S)$ and let $\hat{S}$ consistently estimate S. Let $\tilde{\eta}_{\tau_N} := \hat{\eta}_{\tau_N} + \frac{1}{\sqrt{N}} N(0, \hat{S})$ or a bootstrap random variable, and use the distribution of

$$
\begin{aligned}
\tilde{\mathcal{J}}_N &:= \min_{\eta \in \mathcal{C}_{\tau_N}} N[\tilde{\eta}_{\tau_N} - \eta]' \Omega [\tilde{\eta}_{\tau_N} - \eta] \\
&= \min_{[\nu - \tau_N \mathbf{1}_H / H] \in \mathbf{R}_+^H} N[\tilde{\eta}_{\tau_N} - A\nu]' \Omega [\tilde{\eta}_{\tau_N} - A\nu]
\end{aligned}
\tag{4.5}
$$

to approximate the distribution of $\mathcal{J}_N$. This has the same theoretical justification as the inequality selection procedure. Unlike the latter, however, it avoids the use of an $\mathcal{H}$-representation, thus offering a computationally feasible testing procedure.

To further illustrate the duality between $\mathcal{H}$- and $\mathcal{V}$-representations, we revisit the first two examples. It is not possible to compute B-matrices in our empirical application.

EXAMPLE 3.1—Continued: With two intersecting budget planes, the cone $\mathcal{C}$ is represented by

$$B = \begin{array}{c} \begin{array}{cccc} x_{1|1} & x_{2|1} & x_{1|2} & x_{2|2} \end{array} \\ \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 0 & 0 & -1 \end{pmatrix} \end{array}. \qquad (4.6)$$

The first two rows of B are nonnegativity constraints (the other two such constraints are redundant), the next two rows are an equality constraint forcing the sum of probabilities to be constant across budgets, and only the last constraint is a substantive economic constraint. If the estimator $\hat{\pi}$ fulfills the first four constraints by construction, then the testing problem simplifies to a test of $(1, 0, 0, -1)\pi \leq 0$, the same condition identified earlier.

EXAMPLE 3.2—Continued: Eliminating nonnegativity and adding-up constraints for brevity, numerical evaluation reveals

$$B = \begin{array}{c} \begin{array}{cccccccccccc} x_{1|1} & x_{2|1} & x_{3|1} & x_{4|1} & x_{1|1} & x_{1|2} & x_{3|1} & x_{4|1} & x_{1|1} & x_{2|1} & x_{3|1} & x_{4|1} \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & -1 & -1 & -1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \end{pmatrix} \end{array}. \qquad (4.7)$$

The first three rows are constraints on pairs of budgets that mirror the last row of (4.6). The next two constraints are not implied by these or by additional constraints in Kawaguchi (2017), but they imply the latter.

4.4. *Theoretical Justification*

We now provide a detailed justification. First, we formalize the notion that choice probabilities are estimated by sample frequencies. Thus, for each budget set $\mathcal{B}_j$, denote the choices of N_j individuals, indexed by $n = 1, \ldots, N_j$, by

$$d_{i|j,n} = \begin{cases} 1 & \text{if individual } n \text{ chooses } x_{i|j}, \\ 0 & \text{otherwise,} \end{cases} \quad n = 1, \ldots, N_J.$$

Assume that one observes J random samples $\{\{d_{i|j,n}\}_{i=1}^{I_j}\}_{n=1}^{N_j}$, $j = 1, 2, \ldots, J$. For later use, define

$$d_{j,n} := \begin{bmatrix} d_{1|j,n} \\ \vdots \\ d_{I_j|j,n} \end{bmatrix}, \quad N = \sum_{j=1}^{J} N_J.$$

An obvious way to estimate the vector π is to use choice frequencies

$$\hat{\pi}_{i|j} = \sum_{n=1}^{N_j} d_{i|j,n}/N_j, \quad i = 1, \ldots, I_j, j = 1, \ldots, J. \tag{4.8}$$

The next lemma, among other things, shows that our tightening of the $\mathcal{V}$-representation of $\mathcal{C}$ is equivalent to tightening its $\mathcal{H}$-representation but leaving B unchanged. For a matrix B, let col(B) denote its column space.

LEMMA 4.1: *For $A \in \mathbf{R}^{I \times H}$, let*

$$\mathcal{C} = \{A\nu | \nu \geq 0\}.$$

Also let

$$\mathcal{C} = \{t : Bt \leq 0\}$$

be its $\mathcal{H}$-representation for some $B \in \mathbf{R}^{m \times I}$ such that $B = \begin{bmatrix} B^{\leq} \\ B^{=} \end{bmatrix}$, where the submatrices $B^{\leq} \in \mathbf{R}^{\tilde{m} \times I}$ and $B^{=} \in \mathbf{R}^{(m-\tilde{m}) \times I}$ correspond to inequality and equality constraints, respectively. For $\tau > 0$, define

$$\mathcal{C}_\tau = \left\{A\nu | \nu \geq (\tau/H)\mathbf{1}_H\right\}.$$

Then one also has

$$\mathcal{C}_\tau = \{t : Bt \leq -\tau\phi\}$$

for some $\phi = (\phi_1, \ldots, \phi_m)' \in \text{col}(B)$ with the properties that (i) $\bar{\phi} := [\phi_1, \ldots, \phi_{\tilde{m}}]' \in \mathbf{R}^{\tilde{m}}_{++}$ and (ii) $\phi_k = 0$ for $k > \tilde{m}$.

Lemma 4.1 is *not* just a restatement of the Minkowski–Weyl theorem for polyhedra, which would simply say $\mathcal{C}_\tau = \{A\nu | \nu \geq (\tau/H)\mathbf{1}_H\}$ is alternatively represented as an intersection of closed half-spaces. The lemma instead shows that the inequalities in the $\mathcal{H}$-representation becomes tighter by $\tau\phi$ after tightening the $\mathcal{V}$-representation by $\tau_N \mathbf{1}_H/H$, with the same matrix of coefficients B appearing for both $\mathcal{C}$ and $\mathcal{C}_\tau$. Note that for notational convenience, we rearrange rows of B so that the genuine inequalities come first and pairs of inequalities that represent equality constraints come last.[10] This is without loss of generality; in particular, the researcher does not need to know which rows of B these are. Then as we show in the proof, the elements in ϕ corresponding to the equality constraints are automatically zero when we tighten the space for *all* the elements of ν in the $\mathcal{V}$-representation. This is a useful feature that makes our methodology work in the presence of equality constraints.

The following assumptions are used for our asymptotic theory.

ASSUMPTION 4.1: *For all $j = 1, \ldots, J$, $\frac{N_j}{N} \to \rho_j$ as $N \to \infty$, where $\rho_j > 0$.*

ASSUMPTION 4.2: *We observe J repeated cross sections of random samples $\{\{d_{i|j,n(j)}\}_{i=1}^{I_j}\}_{n(j)=1}^{N_j}, j = 1, \ldots, J$.*

---

[10]In the matrix displayed in (4.6), the third and fourth rows would then come last.

The econometrician also observes the normalized price vector p_j, which is fixed in this section, for each $1 \le j \le J$. Let $\mathcal{P}$ denote the set of all πs that satisfy Condition S.1 in Appendix A for some (common) value of (c_1, c_2).

THEOREM 4.2: *Choose τ_N so that $\tau_N \downarrow 0$ and $\sqrt{N}\tau_N \uparrow \infty$. Also, let Ω be diagonal, where all the diagonal elements are positive. Then under Assumptions* 4.1 *and* 4.2,

$$\liminf_{N \to \infty} \inf_{\pi \in \mathcal{P} \cap \mathcal{C}} \Pr\{\mathcal{J}_N \le \hat{c}_{1-\alpha}\} = 1 - \alpha,$$

where $\hat{c}_{1-\alpha}$ is the $1 - \alpha$ quantile of $\tilde{\mathcal{J}}_N$, $0 \le \alpha \le \frac{1}{2}$.

While it is obvious that our tightening contracts the cone, the result depends on a more delicate feature, namely that we (potentially) turn nonbinding inequalities from the $\mathcal{H}$-representation into binding ones but not vice versa. This feature is not universal to cones as they get contracted. Our proof establishes that it generally obtains if Ω is the identity matrix and all corners of the cone are acute. In this paper's application, we can further exploit the cone's geometry to extend the result to any diagonal Ω.[11] Our method immediately applies to other testing problems featuring $\mathcal{V}$-representations if analogous features can be verified.

5. METHODS FOR TYPICAL SURVEY DATA

The methodology outlined in Section 4 requires that (i) the observations available to the econometrician are drawn on a finite number of budgets and (ii) the budgets are given exogenously, that is, unobserved heterogeneity and budgets are assumed to be independent. These conditions are naturally satisfied in some applications. The empirical setting in Section 7, however, calls for modifications because condition (i) is certainly violated in it and imposing condition (ii) would be very restrictive. These are typical issues for a survey data set. This section addresses them.

Let P_u denote the marginal probability law of u, which we assume does not depend on j. We do not, however, assume that the laws of other random elements, such as income, are time homogeneous. Let $w = \log(W)$ denote log total expenditure, and suppose the researcher chooses a value $\underline{w}_j$ for w for each period j. Note that our algorithm and asymptotic theory remain valid if multiple values of w are chosen for each period. Let $w_{n(j)}$ be the log total expenditure of consumer $n(j)$, $1 \le n(j) \le N_j$, observed in period j.

ASSUMPTION 5.1: *We observe J repeated cross sections of random samples $\{(\{d_{i|j,n(j)}\}_{i=1}^{I_j},$ $w_{n(j)})\}_{n(j)=1}^{N_j}$, $j = 1, \dots, J$.*

The econometrician also observes the unnormalized price vector $\tilde{p}_j$, which is fixed, for each $1 \le j \le J$.

---

[11]It is possible to replace Ω with its consistent estimator and retain uniform asymptotic validity if we further impose a restriction on the class of distributions over which we define the size of our test. Note, however, that $\mathcal{P}$ in our Theorem 4.2 (and its variants in Theorems 5.1 and 5.2) allows for some elements of the vector π being zeros. This makes the use of the reciprocals of estimated variances for the diagonals of the weighting matrix potentially problematic, as it invalidates the asymptotic uniform validity since the required triangular central limit theorem does not hold under parameter sequences where the elements of π converge to zeros. The use of fixed Ω, which we recommend in implementing our procedure, makes contributions from these terms asymptotically negligible, thereby circumventing this problem.

We first assume that the total expenditure is exogenous, in the sense that $w \perp\!\!\!\perp u$ holds under every $P^{(j)}$, $1 \le j \le J$. This exogeneity assumption will be relaxed shortly. Letting $\pi_{i|j}(w) := \Pr\{d_{i|j,n(j)} = 1 | w_{n(j)} = \underline{w}_j\}$ and writing $\pi_j := (\pi_{1|j}, \ldots, \pi_{I_j|j})'$ and $\pi := (\pi_1', \ldots, \pi_J')' = (\pi_{1|1}, \pi_{2|1}, \ldots, \pi_{I_J|J})'$, the stochastic rationality condition is given by $\pi \in \mathcal{C}$ as before. Note that this π can be estimated by standard nonparametric procedures. For concreteness, we use a series estimator, as defined and analyzed in Appendix A. The smoothed version of $\mathcal{J}_N$ (also denoted $\mathcal{J}_N$ for simplicity) is obtained using the series estimator for $\hat{\pi}$ in (4.1). In Appendix A, we also present an algorithm for obtaining the bootstrapped version $\tilde{\mathcal{J}}_N$ of the smoothed statistic.

In what follows, F_j signifies the joint distribution of $(d_{i|j,n(j)}, w_{n(j)})$. Let $\mathcal{F}$ be the set of all $(F_1, \ldots, F_J)$ that satisfy Condition S.2 in Appendix A for some $(c_1, c_2, \delta, \zeta(\cdot))$.

THEOREM 5.1: *Let Condition S.3 hold. Also let Ω be diagonal, where all the diagonal elements are positive. Then under Assumptions 4.1 and 5.1,*

$$\liminf_{N \to \infty} \inf_{(F_1, \ldots, F_J) \in \mathcal{F}} \Pr\{\mathcal{J}_N \le \hat{c}_{1-\alpha}\} = 1 - \alpha,$$

where $\hat{c}_{1-\alpha}$ is the $1 - \alpha$ quantile of $\tilde{\mathcal{J}}_N$, $0 \le \alpha \le \frac{1}{2}$.

Next, we relax the assumption that consumer's utility functions are realized independently from W. For each fixed value $\underline{w}_j$ and the unnormalized price vector $\tilde{p}_j$ in period j, $1 \le j \le J$, define the endogeneity corrected conditional probability[12]

$$\pi\big(\tilde{p}_j/e^{\underline{w}_j}, x_{i|j}\big) := \int \mathbf{1}\big\{D_j(\underline{w}_j, u) \in x_{i|j}\big\} \, dP_u,$$

where $D_j(w, u) := D(\tilde{p}_j/e^w, u)$. Then Theorem 3.1 still applies to

$$\pi_{\text{EC}} := \big[\pi(p_1, x_{1|1}), \ldots, \pi(p_1, x_{I_1|1}), \pi(p_2, x_{1|2}), \ldots, \pi(p_2, x_{I_2|2}), \ldots,$$
$$\pi(p_J, x_{1|J}), \ldots, \pi(p_J, x_{I_J|J})\big]'.$$

Suppose there exists a control variable ε such that $w \perp\!\!\!\perp u | \varepsilon$ holds under every $P^{(j)}$, $1 \le j \le J$. See (S.8) in Appendix A for an example. We propose to use a fully nonparametric, control-function-based two-step estimator, denoted by $\widehat{\pi_{\text{EC}}}$, to define our endogeneity-corrected (EC) test statistic $\mathcal{J}_{\text{EC}_N}$; see Appendix A for details. For this, the bootstrap procedure needs to be adjusted appropriately to obtain the bootstrapped statistic $\tilde{\mathcal{J}}_{\text{EC}_N}$: once again, the reader is referred to Appendix A. This is the method we use for the empirical results reported in Section 7. Let $z_{n(j)}$ be the $n(j)$th observation of the instrumental variable z in period j.

ASSUMPTION 5.2: *We observe J repeated cross sections of random samples $\{(\{d_{i|j,n(j)}\}_{i=1}^{I_j}, x_{n(j)}, z_{n(j)})\}_{n=1}^{N_j}$, $j = 1, \ldots, J$.*

---

[12]This is the conditional choice probability if p is (counterfactually) assumed to be exogenous. We call it *endogeneity corrected* instead of counterfactual to avoid confusion with rationality constrained, counterfactual prediction.

The econometrician also observes the unnormalized price vector $\tilde{p}_j$, which is fixed, for each $1 \le j \le J$.

In what follows, F_j signifies the joint distribution of $(d_{i|j,n(j)}, w_{n(j)}, z_{n(j)})$. Let $\mathcal{F}_{EC}$ be the set of all $(F_1, \ldots, F_J)$ that satisfy Condition S.4 in Appendix A for some $(c_1, c_2, \delta_1, \delta, \zeta_r(\cdot), \zeta_s(\cdot), \zeta_1(\cdot))$. Then we can state the following theorem.

THEOREM 5.2: *Let Condition S.5 hold. Also let Ω be diagonal where all the diagonal elements are positive. Then under Assumptions 4.1 and 5.2,*

$$\liminf_{N \to \infty} \inf_{(F_1, \ldots, F_J) \in \mathcal{F}_{EC}} \Pr\{\mathcal{J}_{EC_N} \le \hat{c}_{1-\alpha}\} = 1 - \alpha,$$

where $\hat{c}_{1-\alpha}$ is the $1 - \alpha$ quantile of $\tilde{\mathcal{J}}_{EC_N}$, $0 \le \alpha \le \frac{1}{2}$.

6. MONTE CARLO SIMULATIONS

We next analyze the performance of cone tightening in a small Monte Carlo study. To keep examples transparent and to focus on the core novelty, we model the idealized setting of Section 4, that is, sampling distributions are multinomial over patches. In addition, we focus on Example 3.2, for which an $\mathcal{H}$-representation in the sense of Weyl–Minkowski duality is available; see displays (3.4) and (4.7) for the relevant matrices.[13] This allows us to alternatively test rationalizability through a moment inequalities test that ensures uniform validity through GMS.[14]

Data were generated from a total of 31 d.g.p.s described below and for sample sizes of $N_j \in \{100, 200, 500, 1000\}$; recall that these are per budget, that is, each simulated data set is based on three such samples. The d.g.p.s are parameterized by the π vectors reported in Table II. They are related as follows: π_0 is in the interior of $\mathcal{C}$, π_2, π_4, and π_6 are outside the interior, and π_2, π_3, and π_5 are on its boundary. Furthermore, $\pi_1 = (\pi_0 + \pi_2)/2$, $\pi_3 = (\pi_0 + \pi_4)/2$, and $\pi_5 = (\pi_0 + \pi_6)/2$. Thus, the line segment connecting π_0 and π_2 intersects the boundary of $\mathcal{C}$ precisely at π_1 and similarly for the next two pairs of vectors. We compute "power curves" along those three line segments at 11 equally spaced points, that is, changing mixture weights in increments of 0.1. This is replicated 500 times at a bootstrap size of $R = 499$. The nominal size of the test is $\alpha = 0.05$ throughout. Ideally, it should be exactly attained at the vectors $\{\pi_1, \pi_3, \pi_5\}$.

Results are displayed in Table III. Noting that the vectors are not too different, we would argue that the simulations indicate reasonable power. Adjustments that ensure uniform validity of tests do tend to cause conservatism for both GMS and cone tightening, but size control markedly improves with sample size.[15] While cone tightening appears less conservative than GMS in these simulations, we caution that the tuning parameters and the distance metrics underlying the test statistics are not directly comparable.

The differential performance across the three families of d.g.p.s is expected because the d.g.p.s were designed to pose different challenges. For both π_1 and π_3, one constraint

---

[13]This is also true of Example 3.1, but that example is too simple because the test reduces to a one-sided test about the sum of two probabilities, and the issues that motivate cone tightening or GMS go away. We verified that all testing methods successfully recover this and achieve excellent size control, including when tuning parameters are set to 0.

[14]The implementation uses a "modified method of moments" criterion function, that is, S_1 in the terminology of Andrews and Soares (2010), and the hard thresholding GMS function, that is, studentized intercepts above $-\kappa_N$, were set to 0 and all others to $-\infty$. The tuning parameter is set to $\kappa_N = \sqrt{\ln(N_j)}$.

[15]We attribute some very slight nonmonotonicities in the "power curves" to simulation noise.

TABLE II

THE π VECTORS USED FOR THE MONTE CARLO SIMULATIONS IN TABLE III

	π_0	π_1	π_2	π_3	π_4	π_5	π_6
$x_{1\mid1}$	0.181	0.190	0.2	0.240	0.3	0.167	0.152
$x_{2\mid1}$	0.226	0.238	0.25	0.213	0.2	0.167	0.107
$x_{3\mid1}$	0.226	0.238	0.25	0.213	0.2	0.333	0.441
$x_{4\mid1}$	0.367	0.333	0.3	0.333	0.3	0.333	0.3
$x_{1\mid2}$	0.181	0.190	0.2	0.240	0.3	0.333	0.486
$x_{2\mid2}$	0.226	0.238	0.25	0.213	0.2	0.167	0.107
$x_{3\mid2}$	0.226	0.238	0.25	0.213	0.2	0.167	0.107
$x_{4\mid2}$	0.367	0.333	0.3	0.333	0.3	0.333	0.3
$x_{1\mid3}$	0.181	0.190	0.2	0.240	0.3	0.333	0.486
$x_{2\mid3}$	0.226	0.238	0.25	0.213	0.2	0.167	0.107
$x_{3\mid3}$	0.226	0.238	0.25	0.213	0.2	0.167	0.107
$x_{4\mid3}$	0.367	0.333	0.3	0.333	0.3	0.333	0.3

TABLE III

MONTE CARLO RESULTS[a]

Method	N_j	π_0	.	.	.	.	π_1	.	.	.	.	π_2
Cone	100	0.002	0.004	0.008	0.008	0.018	0.024	0.060	0.110	0.178	0.238	0.334
tightening	200	0	0	0.004	0.008	0.012	0.040	0.088	0.164	0.286	0.410	0.544
	500	0	0	0.004	0.004	0.026	0.066	0.166	0.310	0.500	0.690	0.856
	1000	0	0	0	0	0.010	0.058	0.206	0.466	0.764	0.924	0.984
GMS	100	0	0	0	0.002	0.004	0.004	0.010	0.024	0.032	0.066	0.100
	200	0	0	0	0	0.004	0.010	0.018	0.048	0.082	0.174	0.262
	500	0	0	0.002	0.002	0.026	0.050	0.150	0.266	0.442	0.640	0.814
	1000	0	0	0	0	0.008	0.050	0.194	0.460	0.758	0.924	0.976

Method	N_j	π_0	.	.	.	.	π_3	.	.	.	.	π_4
Cone	100	0.002	0.002	0.008	0.006	0.012	0.016	0.036	0.070	0.098	0.148	0.200
tightening	200	0	0	0.004	0.008	0.010	0.036	0.068	0.112	0.194	0.296	0.404
	500	0	0	0.004	0.004	0.026	0.064	0.156	0.296	0.456	0.664	0.786
	1000	0	0	0	0	0.010	0.058	0.200	0.460	0.756	0.916	0.974
GMS	100	0	0	0	0.002	0.004	0.004	0.006	0.018	0.028	0.066	0.120
	200	0	0	0	0	0.004	0.004	0.012	0.024	0.056	0.110	0.216
	500	0	0	0.002	0	0.004	0.012	0.034	0.084	0.196	0.348	0.544
	1000	0	0	0	0	0.002	0.028	0.080	0.180	0.406	0.694	0.892

Method	N_j	π_0	.	.	.	.	π_5	.	.	.	.	π_6
Cone	100	0.002	0.002	0.006	0.004	0.020	0.052	0.126	0.326	0.548	0.766	0.934
tightening	200	0	0	0.002	0.004	0.008	0.044	0.202	0.490	0.836	0.962	0.996
	500	0	0	0.002	0.004	0.012	0.072	0.374	0.880	0.992	1	1
	1000	0	0	0	0	0.006	0.052	0.606	0.992	1	1	1
GMS	100	0	0	0	0.004	0.008	0.022	0.072	0.170	0.388	0.604	0.808
	200	0	0	0	0	0.006	0.020	0.112	0.300	0.640	0.880	0.974
	500	0	0	0	0.002	0.004	0.072	0.292	0.716	0.952	1	1
	1000	0	0	0	0	0	0.042	0.498	0.956	1	1	1

[a]See Table II for definition of π vectors. Recall that $\{\pi_1, \pi_3, \pi_5\}$ are on the boundary of $\mathcal{C}$ and π_0 is interior to it. All entries computed from 500 simulations and 499 replications per bootstrap.

is binding and three more are close enough to binding that, at the relevant sample sizes, they cannot be ignored. This is more the case for π_3 compared to π_1. It means that GMS or cone tightening will be necessary, but also that both are expected to be conservative. At vector π_5, three constraints bind and two more are somewhat close. This is a worst case scenario for naive (not using cone tightening or GMS) inference, which will rarely pick up all binding constraints. Indeed, we verified that inference with $\tau_N = 0$ or $\kappa_N = 0$ leads to overrejection. Finally, π_2 and π_4 fulfill the necessary conditions identified by Kawaguchi (2017), so that his test will have no asymptotic power at a parameter value in the first two panels of Table III.

7. EMPIRICAL APPLICATION

We apply our methods to data from the U.K. Family Expenditure Survey, the same data used by BBC. Our testing of a RUM can, therefore, be compared with their revealed preference analysis of a representative consumer. To facilitate this comparison, we use the same selection from these data, namely the time periods from 1975 through 1999, and households with a car and at least one child. The number of data points used varies from 715 (in 1997) to 1509 (in 1975), for a total of 26,341. For each year, we extract the budget corresponding to that year's median expenditure and, following Section 5, estimate the distribution of demand on that budget with polynomials of order 3. Like BBC, we assume that all consumers in one year face the same prices, and we use the same price data. While budgets have a tendency to move outward over time, there is substantial overlap of budgets at median expenditure. To account for endogenous expenditure, we again follow Section 5, using total household income as the instrument. This is also the same instrument used in BBC (2008).

We present results for blocks of eight consecutive periods and the same three composite goods (food, nondurable consumption goods, and services) considered in BBC.[16] For all blocks of seven consecutive years, we analyze the same basket, but also increase the dimensionality of commodity space to four or even five. This is done by first splitting nondurables into clothing and other nondurables and then further into clothing, alcoholic beverages, and other nondurables. Thus, the separability assumptions that we (and others) implicitly invoke are successively relaxed. We are able to go further than much of the existing literature in this regard because, while computational expense increases with K, our approach is not subject to a statistical curse of dimensionality.[17]

Regarding the test's statistical power, increasing the dimensionality of commodity space can, in principle, cut both ways. The number of rationality constraints increases, and this helps if some of the new constraints are violated, but adds noise otherwise. Also, the maintained assumptions become weaker: In principle, a rejection of stochastic rationalizability at three but not four goods might just indicate a failure of separability.

Tables IV and V summarize our empirical findings. They display test statistics, p-values, and the numbers I of patches and H of rationalizable demand vectors; thus, matrices A are of size $(I \times H)$. All entries that show $\mathcal{J}_N = 0$ and a corresponding p-value of 1 were verified to be true zeros, that is, $\hat{\pi}_{EC}$ is rationalizable. All in all, it turns out that

---

[16]As a reminder, Figure 2 illustrates the application. The budget is the 1993 budget embedded in the 1986–1993 block of periods, that is, the figure corresponds to a row of Table V.

[17]Tables IV and V were computed in a few days on Cornell's Economics Computer Cluster Organization (ECCO) cluster (32 nodes). An individual cell of a table can be computed in reasonable time on any desktop computer. Computation of a matrix A took up to 1 hour and computation of one $\mathcal{J}_N$ took about 5 seconds on a laptop.

TABLE IV

EMPIRICAL RESULTS WITH SEVEN PERIODS[a]

	Three Goods				Four Goods				Five Goods		
	I	H	$\mathcal{J}_N$	p	I	H	$\mathcal{J}_N$	p	I	H	$\mathcal{J}_N$	p
75–81	36	6409	3.67	0.38	52	39,957	5.43	0.29	55	53,816	4.75	0.24
76–82	39	4209	11.6	0.14	65	82,507	5.75	0.39	65	82,507	5.34	0.31
77–83	41	7137	9.81	0.17	65	100,728	6.07	0.39	68	133,746	4.66	0.38
78–84	32	3358	7.38	0.24	62	85,888	2.14	0.70	67	116,348	1.45	0.71
79–85	35	5628	0.114	0.96	71	202,686	0.326	0.92	79	313,440	0.219	0.94
80–86	38	7104	0.0055	0.998	58	68,738	1.70	0.81	66	123,462	7.91	0.21
81–87	26	713	0.0007	0.998	42	9621	0.640	0.89	52	28,089	6.33	0.27
82–88	15	42	0	1	21	177	0.298	0.60	31	1283	9.38	0.14
83–89	13	14	0	1	15	31	0.263	0.49	15	31	9.72	0.13
84–90	15	42	0	1	15	42	0.251	0.74	15	42	10.25	0.24
85–91	15	63	0.062	0.77	19	195	3.59	0.45	21	331	3.59	0.44
86–92	24	413	1.92	0.71	33	1859	7.27	0.35	35	3739	9.46	0.28
87–93	45	17,880	1.33	0.74	57	52,316	6.60	0.44	70	153,388	6.32	0.38
88–94	39	4153	1.44	0.70	67	136,823	6.95	0.38	77	313,289	6.91	0.38
89–95	26	840	0.042	0.97	69	134,323	4.89	0.35	78	336,467	5.84	0.31
90–96	19	120	0.040	0.95	56	52,036	4.42	0.19	76	272,233	3.55	0.25
91–97	17	84	0.039	0.93	40	7379	3.32	0.26	50	19,000	3.27	0.24
92–98	13	21	0.041	0.97	26	897	0.060	0.93	26	897	0.011	0.99
93–99	9	3	0.037	0.66	15	63	0	1	15	63	0	1

[a]Notations: I = number of patches; H = number of rationalizable discrete demand vectors; $\mathcal{J}_N$ = test statistic; p = p-value.

TABLE V

EMPIRICAL RESULTS WITH 8 PERIODS[a]

	Three Goods			
	I	H	$\mathcal{J}_N$	p
75–82	51	71,853	11.4	0.17
76–83	64	114,550	9.66	0.24
77–84	52	57,666	9.85	0.20
78–85	49	76,746	7.52	0.26
79–86	55	112,449	0.114	0.998
80–87	41	13,206	3.58	0.58
81–88	27	713	0	1
82–89	16	42	0	1
83–90	16	42	0	1
84–91	20	294	0.072	0.89
85–92	27	1239	2.24	0.68
86–93	46	17,880	1.54	0.75
87–94	48	39,913	1.55	0.75
88–95	42	12,459	1.68	0.70
89–96	27	840	0.047	0.97
90–96	24	441	0.389	0.83
91–98	22	258	1.27	0.52
92–99	14	21	0.047	0.96

[a]Notations: I = number of patches; H = number of rationalizable discrete demand vectors; $\mathcal{J}_N$ = test statistic; p = p-value.

estimated choice probabilities are typically not stochastically rationalizable, but also that this rejection is not statistically significant.[18]

We identified a mechanism that may explain this phenomenon. Consider the 84–91 entry in Table V, where $\mathcal{J}_N$ is especially low. It turns out that one patch on budget $\mathcal{B}_5$ is below $\mathcal{B}_8$ and two patches on $\mathcal{B}_8$ are below $\mathcal{B}_5$. By the reasoning of Example 3.1, the probabilities of these patches must add to less than 1. The estimated sum equals 1.006, leading to a tiny and statistically insignificant violation. This phenomenon occurs frequently and seems to cause the many positive but insignificant values of $\mathcal{J}_N$. The frequency of its occurrence, in turn, has a simple cause that may also appear in other data: If two budgets are slight rotations of each other and demand distributions change continuously in response, then population probabilities of patches like the above will sum to *just* less than 1. If these probabilities are estimated independently across budgets, the estimates will frequently add to slightly more than 1. With seven or eight mutually intersecting budgets, there are many opportunities for such reversals, and positive but insignificant test statistics may become ubiquitous.

The phenomenon of estimated choice frequencies typically not being rationalizable means that there is a need for a statistical testing theory and also a theory of rationality constrained estimation. The former is this paper's main contribution. We leave the latter for future research.

8. CONCLUSION

This paper presented asymptotic theory and computational tools for nonparametric testing of random utility models. Again, the null to be tested was that data were generated by a RUM, interpreted as describing a heterogeneous population, where the only restrictions imposed on individuals' behavior were "more is better" and SARP. In particular, we allowed for unrestricted, unobserved heterogeneity and stopped far short of assumptions that would recover invertibility of demand. We showed that testing the model is nonetheless possible. The method is easily adapted to choice problems that are discrete to begin with, and one can easily impose more, or fewer, restrictions at the individual level.

Possibilities for extensions and refinements abound, and some of these have already been explored. We close by mentioning further salient issues.

(i) We provide algorithms (and code) that work for reasonably sized problem, but it would be extremely useful to make further improvements in this dimension.

(ii) The extension to infinitely many budgets is of obvious interest. Theoretically, it can be handled by considering an appropriate discretization argument (McFadden (2005)). For the proposed projection-based econometric methodology, such an extension requires evaluating choice probabilities locally over points in the space of p via nonparametric smoothing, and then use the choice probability estimators in the calculation of the $\mathcal{J}_N$-statistic. The asymptotic theory then needs to be modified. Another approach that can mitigate the computational constraint is to consider a partition of the space of p such that

---

[18]In additional analyses not presented here, we replicated these tables using polynomials of degree 2, as well as setting $\tau_N = 0$. The qualitative finding of many positive but insignificant test statistics remains. In isolation, this finding may raise questions about the test's power. However, the test exhibits reasonable power in our Monte Carlo exercise and also rejects rationalizability in an empirical application elsewhere (Hubner (2018)). We also checked whether small but positive test statistics are caused by adding up constraints, that is, by the fact that all components of π that correspond to one budget must jointly be on some unit simplex. The estimator $\hat{\pi}$ can slightly violate this. Adding-up failures occur but are at least 1 order of magnitude smaller than the distance from a typical $\hat{\pi}$ to the corresponding projection $\hat{\eta}$.

$\mathbf{R}_+^K = \mathcal{P}_1 \cup \mathcal{P}_2 \cup \cdots \cup \mathcal{P}_M$. Suppose we calculate the $\mathcal{J}_N$-statistic for each of these partitions. Given the resulting M statistics, say $\mathcal{J}_N^1, \ldots, \mathcal{J}_N^M$, we can consider $\mathcal{J}_N^{\max} := \max_{1 \le m \le M} \mathcal{J}_N^m$ or a weighted average of them. These extensions and their formal statistical analysis are of practical interest.

(iii) While we allow for endogenous expenditure and, therefore, for the distribution of u to vary with observed expenditure, we do assume that samples for all budgets are drawn from the same underlying population. This assumption can obviously not be dropped completely. However, it will frequently be of interest to impose it only conditionally on observable covariates, which must then be controlled for. This may be especially relevant for cases where different budgets correspond to independent markets, but also to adjust for slow demographic change as does, strictly speaking, occur in our data. It requires incorporating nonparametric smoothing in estimating choice probabilities as in Section 5, and then averaging the corresponding $\mathcal{J}_N$-statistics over the covariates. This extension will be pursued.

(iv) Natural next steps after rationality testing are extrapolation to (bounds on) counterfactual demand distributions and welfare analysis, that is, along the lines of BBC (2008) or, closer to our own setting, Adams (2016) and Deb et al. (2018). This extension is being pursued. Indeed, the tools from Section 3 have already been used for choice extrapolation (using algorithms from an earlier version of this paper) in Manski (2014).

(v) The econometric techniques proposed here can be potentially useful in much broader contexts. Indeed, they have already been used to nonparametrically test game theoretic models with strategic complementarities (Lazzati, Quah, and Shirai (2018a,b)), a novel model of "price preference" (Deb et al. (2018)), and the collective household model Hubner (2018). Even more generally, existing proposals for testing in moment inequality models (Andrews and Guggenberger (2009), Andrews and Soares (2010), Bugni, Canay, and Shi (2015), Romano and Shaikh (2010)) work with explicit inequality constraints, that is, (in the linear case) $\mathcal{H}$-representations. In settings in which theoretical restrictions inform a $\mathcal{V}$-representation of a cone or, more generally, a polyhedron, the $\mathcal{H}$-representation will typically not be available in practice. We expect that our method can be used in many such cases.

REFERENCES

Adams, A. (2016): "Mutually Consistent Revealed Preference Bounds," Manuscript, University of Oxford. [1907]

Afriat, S. N. (1967): "The Construction of Utility Functions From Expenditure Data," *International Economic Review*, 8 (1), 67–77. [1884]

Andrews, D. W. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858. [1886]

Andrews, D. W., and P. Guggenberger (2009): "Validity of Subsampling and "Plug-in Asymptotic" Inference for Parameters Defined by Moment Inequalities," *Econometric Theory*, 25, 669–709. [1907]

Andrews, D. W., and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78, 119–157. [1886,1895,1897,1902,1907]

Blundell, R., and J. L. Powell (2003): "Endogeneity in Nonparametric and Semiparametric Regression Models," *Econometric Society Monographs*, 36, 312–357. [1884]

Blundell, R., M. Browning, L. Cherchye, I. Crawford, B. De Rock, and F. Vermeulen (2015): "Sharp for SARP: Nonparametric Bounds on Counterfactual Demands," *American Economic Journal: Microeconomics*, 7 (1), 43–60. [1885]

Blundell, R., M. Browning, and I. Crawford (2003): "Nonparametric Engel Curves and Revealed Preference," *Econometrica*, 71 (1), 205–240. [1885]

——— (2007): "Improving Revealed Preference Bounds on Demand Responses," *International Economic Review*, 48 (4), 1227–1244. [1885]

———— (2008): "Best Nonparametric Bounds on Demand Responses," *Econometrica*, 76 (6), 1227–1262.[1885]

BLUNDELL, R., D. KRISTENSEN, AND R. MATZKIN (2014): "Bounding Quantile Demand Functions Using Revealed Preference Inequalities," *Journal of Econometrics*, 179 (2), 112–127.[1885]

———— (2017): "Individual Counterfactuals With Multidimensional Unobserved Heterogeneity," cemmap working paper CWP60/17. [1885]

BUGNI, F. A. (2010): "Bootstrap Inference in Partially Identified Models Defined by Moment Inequalities: Coverage of the Identified Set," *Econometrica*, 78 (2), 735–753.[1886,1897]

BUGNI, F. A., I. A. CANAY, AND X. SHI (2015): "Specification Tests for Partially Identified Models Defined by Moment Inequalities," *Journal of Econometrics*, 185 (1), 259–282.[1886,1907]

CANAY, I. A. (2010): "{EL} Inference for Partially Identified Models: Large Deviations Optimality and Bootstrap Validity," *Journal of Econometrics*, 156 (2), 408–425.[1886,1897]

CHERNOFF, H. (1954): "On the Distribution of the Likelihood Ratio," *Annals of Mathematical Statistics*, 25, 573–578.[1886]

CHERNOZHUKOV, V., H. HONG, AND E. TAMER (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Model," *Econometrica*, 75 (5), 1243–1284.[1886]

DEB, R., Y. KITAMURA, J. K.-H. QUAH, AND J. STOYE (2018): "Revealed Price Preference: Theory and Empirical Analysis," ArXiv e-prints, https://arxiv.org/abs/1801.02702. [1884,1890,1907]

DETTE, H., S. HODERLEIN, AND N. NEUMEYER (2016): "Testing Multivariate Economic Restrictions Using Quantiles: The Example of Slutsky Negative Semidefiniteness," *Journal of Econometrics*, 191 (1), 129–144.[1886]

DOIGNON, J.-P., AND S. REXHEP (2016): "Primary Facets of Order Polytopes," *Journal of Mathematical Psychology*, 75, 231–245, Special Issue in Honor of R. Duncan Luce.[1886]

FLOYD, R. W. (1962): "Algorithm 97: Shortest Path," *Communications of the ACM*, 5 (6), 345.[1892]

GOURIEROUX, C., A. HOLLY, AND A. MONFORT (1982): "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models With Inequality Constraints on the Regression Parameters," *Econometrica*, 50 (1), 63–80.[1886]

GRUBER, P. (2007): *Convex and Discrete Geometry*. Grundlehren der Mathematischen Wissenschaften. New York: Springer.[1896]

GRÜNBAUM, B., V. KAIBEL, V. KLEE, AND G. M. ZIEGLER (2003): *Convex Polytopes*. New York: Springer.[1896]

GUGGENBERGER, P., J. HAHN, AND K. KIM (2008): "Specification Testing Under Moment Inequalities," *Economics Letters*, 99, 375–378.[1886]

HAUSMAN, J. A., AND W. K. NEWEY (2016): "Individual Heterogeneity and Average Welfare," *Econometrica*, 84 (3), 1225–1248.[1885]

HODERLEIN, S. (2011): "How Many Consumers Are Rational?" *Journal of Econometrics*, 164 (2), 294–309.[1885]

HODERLEIN, S., AND J. STOYE (2014): "Revealed Preferences in a Heterogeneous Population," *The Review of Economics and Statistics*, 96 (2), 197–213.[1885,1886]

———— (2015): "Testing Stochastic Rationality and Predicting Stochastic Demand: The Case of Two Goods," *Economic Theory Bulletin*, 3 (2), 313–328.[1885]

HOUTHAKKER, H. S. (1950): "Revealed Preference and the Utility Function," *Economica*, 17 (66), 159–174.[1884]

HUBNER, S. (2018): "It's Complicated: A Nonparametric Test of Preference Stability Between Singles and Couples," Preprint, University of Oxford. [1884,1906,1907]

IMBENS, G. W., AND C. MANSKI (2004): "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 72, 1845–1857.[1886]

IMBENS, G. W., AND W. K. NEWEY (2009): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 77 (5), 1481–1512.[1884]

KAWAGUCHI, K. (2017): "Testing Rationality Without Restricting Heterogeneity," *Journal of Econometrics*, 197 (1), 153–171.[1885,1886,1891,1898,1904]

KITAMURA, Y., AND J. STOYE (2018): "Supplement to 'Nonparametric Analysis of Random Utility Models'," *Econometrica Supplemental Material*, 86, https://doi.org/10.3982/ECTA14478. [1884]

KUDO, A. (1963): "A Multivariate Analogue of the One-Sided Test," *Biometrika*, 50 (3/4), 403–418.[1886]

LAZZATI, N., J. QUAH, AND K. SHIRAI (2018a): "Nonparametric Analysis of Monotone Choice," Manuscript, University of California Santa Cruz. [1884,1907]

———— (2018b): "Strategic Complementarity in Spousal Smoking Behavior," Manuscript, University of California Santa Cruz. [1884,1907]

LEWBEL, A. (2001): "Demand Systems With and Without Errors," *American Economic Review*, 91 (3), 611–618.[1885]

MANSKI, C. F. (2007): "Partial Identification of Counterfactual Choice Probabilities," *International Economic Review*, 48 (4), 1393–1410.[1885]

——— (2014): "Identification of Income-Leisure Preferences and Evaluation of Income Tax Policy," *Quantitative Economics*, 5 (1), 145–174.[1885,1890,1907]

MATZKIN, R. (2006): "Heterogeneous Choice," in *Advances in Econometrics: Proceedings of the 9th World Congress*. Budapest: Cambridge University Press. [1888]

MCFADDEN, D., AND K. RICHTER (1991): "Stochastic Rationality and Revealed Stochastic Preference," in *Preferences, Uncertainty and Rationality*, ed. by J. Chipman, D. McFadden, and K. Richter. Boulder: Westview Press, 161–186. [1883,1884,1887]

MCFADDEN, D. L. (2005): "Revealed Stochastic Preference: A Synthesis," *Economic Theory*, 26, 245–264.[1884,1886,1890,1906]

PERLMAN, M. D. (1969): "One-Sided Testing Problems in Multivariate Analysis," *Annals of Mathematical Statistics*, 40, 549–567.[1886]

ROBINSON, R. (1973): "Counting Labeled Acyclic Digraphs," in *New Directions in the Theory of Graphs*, ed. by F. Harary. New York: Academic Press.[1893]

ROMANO, J., AND A. SHAIKH (2010): "Inference for the Identified Set in Partially Identified Econometric Models," *Econometrica*, 78 (1), 169–211.[1886,1907]

ROSEN, A. M. (2008): "Confidence Sets for Partially Identified Parameters That Satisfy a Finite Number of Moment Inequalities," *Journal of Econometrics*, 146 (1), 107–117.[1886]

SAMUELSON, P. A. (1938): "A Note on the Pure Theory of Consumer's Behaviour," *Economica*, 5 (17), 61–71.[1884]

SHAPIRO, A. (1988): "Towards a Unified Theory of Inequality Constrained Testing in Multivariate Analysis," *International Statistical Review*, 56 (1), 49–62.[1886]

STOYE, J. (2009): "More on Confidence Intervals for Partially Identified Parameters," *Econometrica*, 77, 1299–1315.[1886]

TAKEMURA, A., AND S. KURIKI (1997): "Weights of $\overline{\chi}^2$ Distribution for Smooth or Piecewise Smooth Cone Alternatives," *The Annals of Statistics*, 25 (6), 2368–2387.[1886]

TALLA NOBIBON, F., B. SMEULDERS, AND F. C. R. SPIEKSMA (2015): "A Note on Testing Axioms of Revealed Preference," *Journal of Optimization Theory and Applications*, 166 (3), 1063–1070.[1893]

WOLAK, F. A. (1991): "The Local Nature of Hypothesis Tests Involving Inequality Constraints in Nonlinear Models," *Econometrica*, 59, 981–995.[1886]

ZIEGLER, G. M. (1995): *Lectures on Polytopes*. Graduate Texts in Mathematics. New York: Springer.[1896]