



Classifying bounded rationality in limited data sets: a Slutsky matrix approach

Victor H. Aguiar¹ · Roberto Serrano²

Received: 10 September 2017 / Accepted: 23 June 2018 / Published online: 5 July 2018
© The Author(s) 2018

Abstract

Given any observed finite sequence of prices, wealth, and demand choices, we propose a way to measure and classify the departures from rationality in a systematic fashion, by connecting violations of the underlying Slutsky matrix properties to the length of revealed demand cycles. The approach complements our previous study (Aguiar and Serrano in *J Econ Theory* 172:163–201, 2017), which is based on the entire demand function. The methodology can be easily applied in experimental demand data sets.

Keywords Consumer theory · Rationality · Slutsky matrix norm · Revealed preference approach · Bounded rationality

JEL Classification C60 · D10

1 Introduction

The Strong Axiom of Revealed Preference (SARP) completely characterizes whether a finite set of prices and demand choices can be rationalized. However, this binary approach is incomplete, in the sense that we do not know by how much behavior departs from rationality. Furthermore, we learn little about the demand data that fails

Some of the results here without their proofs have appeared as an online supplement to Aguiar and Serrano (2017). This paper is a revised version of Aguiar and Serrano (2016). The co-editor and two anonymous referees had useful comments that improved the paper. We thank Mark Dean, Xavier Gabaix, John Quah, Jesse Shapiro, and the participants at the Brown University Theory Lunch and SAET 2015 in Cambridge for useful comments and encouragement..

✉ Victor H. Aguiar
vhaguiar@uwo.ca

Roberto Serrano
roberto_serrano@brown.edu

¹ Department of Economics, University of Western Ontario, London, ON, Canada

² Department of Economics, Brown University, Providence, RI, USA

to satisfy SARP. This has motivated an important body of literature that has focused on how to measure the size of rationality violations. The revealed preference approach (henceforth RP) is the preferred method to test the rationality hypothesis in consumer behavior, due to its nonparametric nature. The leading contributions of this approach to testing rationality are Afriat (1973) and Varian (1983), which revolve around Afriat Efficiency Index (henceforth AEI). The prevalence of the RP approach is due to the simplicity of its implementation for the commonly used data sets, and to its flexibility, explained by its nonparametric nature. However, in the words of Varian (1983) “the RP approach fails to *summarize* the data in a useful way.”

In a recent contribution, Aguiar and Serrano (2017) propose an index of violations of rationality based on the Slutsky matrix. Their Slutsky matrix norm (SMN) approach helps to answer the question of how far is a given behavior (demand function) from rationality, using the entire observed Slutsky matrix function distance to its closest rational Slutsky matrix function. Moreover, the methodology provides a useful classification of the violations of the classical axioms of revealed demand. The main limitation of the approach, however, is its reliance on infinite data, making its applicability questionable.

In this paper, we demonstrate how the approach in Aguiar and Serrano (2017) can be adapted to finite data sets. In doing so, we attempt to provide a complementary approach to the RP methodology, stemming from the tradition of Antonelli (1951), Slutsky (1915), and Hurwicz and Uzawa (1971) by using the Slutsky matrix to test the empirical implications of consumer behavior in a finite data set environment. In particular, we build and implement a methodology that may help to measure and classify boundedly rational demand data by means of an index of deviations from rationality that has a useful decomposition of SARP-inconsistent behavior.

We use classical results from the RP approach to show that nonrational behavior will be accounted for by a Slutsky matrix function that fails to satisfy the Slutsky regularity conditions for *any* extension of the data set. Conversely, if a data set can be rationalized in the sense of Afriat (1973), there exists an extension thereof that offers a rationalization in the sense of Hurwicz and Uzawa (1971). An implication of this observation is that testing rationality in the RP approach and our modification of the SMN approach with finite data is equivalent. In particular, a data set that violates SARP will have a positive Slutsky norm for *any* extension of the data set, with an informative decomposition of SARP-inconsistencies. These results are obtained because the length of revealed demand cycles, revealed demand axioms, and properties of the Slutsky matrix are neatly connected. To link finite and infinite environments, we exploit the fact that any sequence of data points can be interpolated by a continuously differentiable function.

Thus, our main contribution is to propose a way to classify consumer behavior according to its Slutsky matrix norm in limited data sets. We adapt the decomposition of the Slutsky norm into the different violations of axioms of revealed demand for finite data. Our results seem to hold well in simulations.¹

¹ In the online supplement to Aguiar and Serrano (2017), we have applied the methodology of the current paper to a well-known experimental data set from Ahn et al. (2014). This data set reports consumer demand at the level of the individual. We find that the different kinds of violations of rationality appear to correlate

To the best of our knowledge, this is the only measure of departures from rationality that allows such a decomposition analysis. In addition, the Slutsky matrix norm differs from the AEI (Afriat 1973), the most frequently used measure of violations of rationality, in that the former is a positive index of rationality, whereas the latter is normative. Indeed, the Slutsky matrix norm measures the error a modeller would make when predicting the correct elasticity behavior of a consumer, if she makes the rationality assumption. In contrast, the AEI quantifies the minimum wasted income that the consumer has incurred without improving his own welfare. Other well-known indices of departures from rationality that can be called normative are the Money Pump Index (Echenique et al. 2011) and the Minimum Cost Index (Dean and Martin 2015). For other positive indices, see Varian (1983) and any other goodness of fit index, concerning failures in predicting out-of-sample demand quantities (behavioral nearness), under the rationality assumption. Relating our approach to some of these, we also provide theoretical results that connect quantitatively the SMN with the AEI and the Behavioral Nearness problem (i.e., the least distance from any given demand to its best rational demand approximation).

This is the paper’s outline. Section 2 presents the primitives of our analysis and the model. Section 3 presents an auxiliary equivalence result between the RP and the SMN approaches that provides a logical bridge between the finite data set environment and the Slutsky matrix function approach. Section 4 contains our main results and develops a new class of interpolators of consumer choice data sets that have a minimal Slutsky norm and have desirable properties. Section 5 implements simulation exercises to study the numerical behavior of the minimal Slutsky norm interpolators and a comparison with Afriat’s index. Section 6 discusses formally the connection of the SMN with other measures of bounded rationality. Finally, we present a brief literature review in Sect. 7 and conclude in Sect. 8. The proofs are collected in an “Appendix”.

2 The model: testing rationality with limited data sets

Our primitive is a finite array of prices, wealth levels, and demand choices $O^K = \{(p^k, w^k), x^k\}_{k=1}^K$ for an individual or decision unit. The observation $x^k \in \mathbb{R}_+^L$ is a L dimensional vector of quantities, corresponding to L goods, that we observed being chosen by the consumer at the price-wealth pair $(p^k, w^k) \in \mathbb{R}_{++}^L \times \mathbb{R}_{++}$. These consumption choices may or may not be rational.

The consumption set is \mathbb{R}_+^L so the finiteness of the observed data does not mean that the consumer chooses from finitely-many options. Rather, our problem is that, as analysts, we can observe only a finite number of price-wealth pairs with their corresponding choices.

Consider a demand function $x : P \times W \mapsto \mathbb{R}_+^L$, going from $\mathcal{Z} = P \times W$, a compact space of price-wealth pairs (p, w) , $P \subseteq \mathbb{R}_{++}^L$, $W \subseteq \mathbb{R}_{++}$, to the consumption set. We assume x satisfies Walras’ law: $\sum_{l=1}^L p_l x_l(p, w) = w$ for all price-wealth pairs.

Footnote 1 continued

with one another, and that failures of symmetry (Ville axiom/Path Independence) are more prevalent than failures of negative semidefiniteness (WARP).

This demand is a deterministic choice rule that represents the behavior of a consumer at each given price-wealth pair. The demand functions considered can be thought (without loss of generality) to belong to a closed and bounded set of demand functions $\mathcal{X}(\mathcal{Z}) \subset \mathcal{C}^1(\mathcal{Z}, \mathbb{R}_+^L)$. We assume that this compact space has the $\mathcal{C}^1(\mathcal{Z}, \mathbb{R}_+^L)$ norm.² When we have two vectors of the same length say $v_1, v_2 \in \mathbb{R}^L$ we write its inner product as $v_1 v_2$.

Definition 1 Data Generating Demand Function (DGDF). We say that a demand function $x \in \mathcal{X}$ is a DGDF of $O^K = \{p^k, w^k, x^k\}_{k=1}^K$ if $x^k = x(p^k, w^k)$ for all $k \in \{1, \dots, K\}$.

Given that our interest is to connect with approaches based on the entire (infinite) functional data, we introduce the following theoretical construct. We say that $x \in \mathcal{X}(\mathcal{Z})$ is the “true” DGDF if for any data set O^K or “sample” of individual choices for a given price-wealth situation $(p^k, w^k) \in P \times W$, it is the case that $x \in \mathcal{X}(\mathcal{Z})$ is the unique function $x : P \times W \rightarrow \mathbb{R}_+^L$ such that $x^k = x(p^k, w^k)$ for all $k \in \{1, \dots, K\}$ for all O^K . For example, a true DGDF does not exist if we observe two different choices for the same price-wealth situation, hence excluding the case of choice correspondences.

We also assume that Walras’ law holds throughout, for all O^K , $p^k x^k = w^k$ and all $x \in \mathcal{X}$ (i.e., $px(p, w) = w$ for all $(p, w) \in P \times W$). Behaviorally, when we face deterministic choices that fulfill Walras’ law, we are concerned with testing whether O^K can be rationalized by a locally non satiated strictly convex preference relation.

In addition, we need a second ingredient, which can be expressed as the extensions of a finite data set O^K . Every finite data set O^K has an associated set of functions that describe all possible extensions, which are also elements of $\mathcal{X}(\mathcal{Z})$.

Definition 2 Extensions of a consumer’s data set. We say that a subset of demand functions $\mathcal{X}^K(\mathcal{Z}) \subseteq \mathcal{X}(\mathcal{Z})$ is the set of extensions of a data set O^K when $\mathcal{X}^K(\mathcal{Z}) = \{x \in \mathcal{X}(\mathcal{Z}) \mid x^k = x(p^k, w^k) \text{ and } (p^k, w^k) \in \mathcal{Z} \ \forall \{p^k, w^k, x^k\} \in O^K\}$.

The following claim is trivial:

Claim 1 For finite K and any O^K its set of extensions $\mathcal{X}^K(\mathcal{Z}) \neq \emptyset$ if for any $k, m \in \{1, \dots, K\}$, whenever $p^k = p^m$ and $w^k = w^m$, it is the case that $x^k = x^m$.

For instance, we can always find a smooth $x \in \mathcal{X}(\mathcal{Z})$ using a number of interpolation techniques for O^K , under the condition above (such as an interpolator of polynomial degree or one based on B-splines, among others).

3 Connecting the Slutsky matrix, revealed demand cycles, and axioms

3.1 Preliminaries

3.1.1 RP and Slutsky matrix approaches to test rationality

The RP approach to testing and measuring departures from rationality in limited data sets relies on the SARP. The main concept needed is that of a revealed demand cycle.

² The $\mathcal{C}^1(\mathcal{Z}, \mathbb{R}^L)$ is the space of continuously differentiable mappings from \mathcal{Z} to \mathbb{R}^L .

Definition 3 (*Revealed demand cycle*) A Revealed demand cycle $C^{D,T}$ is a pair of sequences $\{p^t, x^t\}_{t=0}^T$ where $\{x^t\}_{t=0}^T$ is a sequence of choices and $\{p^t\}_{t=0}^T$ is a sequence of prices such that $x^0 = x^T$ and $p^t[x^t - x^{t-1}] \geq 0$ for all $t \in \{1, \dots, T\}$.³

If we can find a revealed demand cycle $C^{D,T}$ for $T \geq 2$, then we have a violation of SARP. Namely, we define the relation R as follows: $x^t R x^{t-1}$ whenever $p^t x^t \geq p^t x^{t-1}$, which is read as x^t being “directly revealed preferred” to x^{t-1} . Also, define the relation \bar{R} as follows: $x^t \bar{R} x^s$ whenever one can find a sequence $x^t R x^{t-1} R x^{t-2} R \dots R x^s$, which we read as x^t being “revealed preferred” to x^s . It is clear that in $C^{D,T}$ we have $x^T R x^{T-1} R x^{T-2} \dots R x^1 R x^0$, which implies that $x^T \bar{R} x^1$, but since $x^0 = x^T$, we also have $x^1 R x^T$, thus violating SARP. We say that a revealed demand cycle is strict whenever $C^{D,T}$ is such that for some $t \in \{1, \dots, T-1\}$, it holds that $p^t[x^t - x^{t-1}] > 0$.

Another characterization of rationality for demand functions $x \in \mathcal{X}(\mathcal{Z})$ was provided by Hurwicz and Uzawa (1971) in the form of restrictions over the Slutsky matrix function. This characterization requires the knowledge of the whole demand function (i.e., infinite data).

Definition 4 Let $\mathcal{Z} \subset P \times W$ be given, and denote by (p, w) an arbitrary price-wealth pair in \mathcal{Z} .⁴ Then the Slutsky matrix function $S \in \mathcal{M}(\mathcal{Z})$ ⁵ is defined pointwise: $S(p, w) = D_p x(p, w) + D_w x(p, w)x(p, w)' \in \mathbb{R}^{L \times L}$, with entry $s_{l,k}(p, w) = \frac{\partial x_l(p, w)}{\partial p_k} + \frac{\partial x_l(p, w)}{\partial w} x_k(p, w)$ for a given demand $x \in \mathcal{X}(\mathcal{Z})$.

The Slutsky matrix function S of a rational demand $x \in \mathcal{X}(\mathcal{Z})$ must satisfy the Slutsky regularity conditions. Namely, S is symmetric (property σ) or $s_{ij}(p, w) = s_{ji}(p, w)$ for all $(p, w) \in \mathcal{Z}$, S is singular in prices (property π) or $S(p, w)p = 0$ and $S(p, w)'p = 0$ for all $(p, w) \in \mathcal{Z}$, and S is negative semidefinite (property ν) or $S(p, w)$ is negative semidefinite for all $(p, w) \in \mathcal{Z}$. Hurwicz and Uzawa (1971) proved that the converse is also true, that is if $x \in \mathcal{X}(\mathcal{Z})$ has an associated Slutsky matrix function that satisfies σ, π , and ν , then it can be generated by maximizing a continuous utility function.⁶

3.1.2 Ville cycles, weak WARP, and revealed demand cycles

Note that when attempting to evaluate by how much a data set O^K departs from being rationalized, multiple paths can be followed. In fact, there is a large literature of measuring departures from rationality in finite data sets. Most of them share the common feature that their measures are a function of the revealed demand cycles $C^{D,T}$ for $T \geq 2$. One central question that we tackle is how the properties of the revealed

³ We sometimes write $x(t) = x^t$ and $p(t) = p^t$. Recall also that we assume Walras’ law: $p^t x^t = w^t$ for all t .

⁴ Since Z is closed, we use the definition of differential in Graves (1956) that is defined not only in the interior but also on the accumulation points of Z .

⁵ Let $\mathcal{M}(\mathcal{Z})$ be the complete metric space of matrix-valued functions, $F : Z \mapsto \mathbb{R}^L \times \mathbb{R}^L$, equipped with the weighted inner product $\langle F, G \rangle_W = \int_{z \in Z} Tr([W(z)F(z)]'[W(z)G(z)])dz$.

⁶ Hurwicz and Uzawa (1971) require other regularity conditions to prove the sufficiency of the Slutsky regularity conditions for the existence of a utility function that generates the demand function. However, recent work has shown the necessity and sufficiency of those conditions in our environment Hosoya (2017).

demand cycle relate to the Slutsky matrix of a demand function. The key to do this is to draw a connection between the revealed demand cycles from finite data and both the Ville cycles and the Weak WARP defined for infinite data.

Roughly, recall that the symmetry of the Slutsky matrix function is equivalent to the Ville Axiom of Revealed Preference (VARP)/Path Independence (Hurwicz and Richter—HR henceforth—Hurwicz and Richter 1979). To state this axiom, we need to define a real income path.

Definition 5 (*Real income path*). A real income path consists of both a wealth path $w : [0, b] \mapsto W$, and a price path $p : [0, b] \mapsto P$, having that $(w(\tau), p(\tau))$ is a piecewise continuously differentiable path in \mathcal{Z} .

Thus, we are allowing for continuity of the derivative to fail at a countable subset of points. Jerison and Jerison (1992) define a rising real income path whenever $(\frac{\partial w}{\partial \tau}(\tau), \frac{\partial p}{\partial \tau}(\tau))$ exist, leading to $\frac{\partial w}{\partial \tau}(\tau) > \frac{\partial p}{\partial \tau}(\tau)'x(p(\tau), w(\tau))$. A Ville cycle is a rising real income path such that $(w(0), p(0)) = (w(b), p(b))$.

We define a Ville Cycle in our setting as follows:

Definition 6 (*Ville cycle*) A Ville Cycle $C^{V(S),b}$ is a pair of functions $(p(\tau), x(\tau))$ for $\tau \in [0, b]$ for some $b > 0$ where x is a S continuously differentiable commodity path $x : [0, b] \rightarrow \mathbb{R}_+^L$ such that $x(0) = x(b)$ and $x \in C^S([0, b]; \mathbb{R}_+^L)$ for $S \geq 1$ and $p(\tau) \frac{\partial x(\tau)}{\partial \tau} > 0$ almost everywhere in $\tau \in [0, b]$, for any piecewise continuous price path $p : [0, b] \rightarrow \mathbb{R}_{++}^L$.

Observe that if there is a rising real income situation $\frac{\partial w}{\partial \tau}(\tau) > \frac{\partial p}{\partial \tau}(\tau)x(p(\tau), w(\tau))$, and the price-wealth path forms a cycle $(p(0), w(0)) = (p(b), w(b))$, then the situation is equivalent (almost everywhere) to $p(\tau) \frac{\partial x(\tau)}{\partial \tau} > 0$ for all $\tau \in [0, b]$ since $\frac{\partial w}{\partial \tau}(\tau) = p(\tau) \frac{\partial x(\tau)}{\partial \tau} + \frac{\partial p}{\partial \tau}(\tau)x(p(\tau), w(\tau))$.

Axiom 1 *Ville Axiom of revealed preference (VARP)*. A demand function $x \in \mathcal{X}$ is said to satisfy VARP if it does not have a Ville cycle $C^{V(S),b}$ for all $S \geq 1$ and $b > 0$.

Next, we state the weak version of WARP, which is equivalent to the NSD of the Slutsky matrix.

Axiom 2 *Weak version of the weak axiom of revealed preference (Weak WARP)*.⁷ We say that a demand function $x \in \mathcal{X}$ satisfies Weak WARP if when we have $\bar{p}x(p, w) \leq \bar{w}$ then it follows that $px(\bar{p}, \bar{w}) \geq w$.

Observe that, under the restrictions on the sampling scheme (which rules out non-deterministic choice), Weak WARP is simplified. In fact, for our purposes, we take Weak WARP to require this: if $\bar{p}x(p, w) < \bar{w}$ then $px(\bar{p}, \bar{w}) > w$ when $(\bar{p}, \bar{w}) \neq (p, w)$.

An additional condition that rational consumers fulfill is the property of “no money illusion” or homogeneity of degree zero.

Axiom 3 *Homogeneity of degree zero (HD0)*. We say that a demand function $x \in \mathcal{X}$ satisfies HD0 if $x(\alpha p, \alpha w) = x(p, w)$ for all $\alpha > 0$.

⁷ The strict version of WARP says that $p^l x^l \geq p^l x^s$ then $p^s x^l > p^s x^s$.

Finally, we state an additional axiom, that is equivalent to Weak WARP under homogeneity of degree zero, and that is the Wald Axiom. The Wald axiom requires that if we have a sequence (possibly a pair) of observations such that $p^k x^k = \bar{w}$ for all $k \in \{1, \dots, K\}$, then: $p^k x(p^l, \bar{w}) < \bar{w}$ implies $p^l x(p^k, \bar{w}) > \bar{w}$.⁸

3.2 Preliminary results

We are ready to make the connection between the finite and the infinite data approaches to characterize rationality. The RP approach is concerned with measuring departures from SARP when there is finite data, while the SMN approach, measuring whether the Slutsky matrix norm is positive, uses functional data (i.e., the actual DGDF). It follows that a global connection between both approaches must rely on how to extend a data set $O^K = \{p^k, w^k, x^k\}_{k=0}^K$ to the whole region \mathcal{Z} . The set of extensions can be very large, generally infinite. So our first interest is to know if a data set that violates SARP, when extended, produces a Slutsky matrix function that is nonrational. Also, we are interested in the question of whether a demand function with a nonrational Slutsky matrix will be able to generate a data set that violates SARP. The answer in the affirmative is provided by Proposition 0, which follows from three well-known characterizations of rationality, the ones by Afriat, Hurwicz-Richter, and Hurwicz-Uzawa, when adapted to a limited data set O^K .

Proposition (0) *The next three statements are equivalent:*

1. *A given data set O^K with $K \geq 2$ has at least a revealed demand cycle.*
2. *All elements $x \in \mathcal{X}^K(\mathcal{Z})$ of the set of extensions of the data set either violate Weak WARP or violate VARP.*
3. *All elements $x \in \mathcal{X}^K(\mathcal{Z})$ in the set of extensions of the data set have an associated Slutsky matrix that fails at least one Slutsky regularity condition (σ , π or ν).*

We will show only that (1) is equivalent to (3), in the following lemma.

Lemma 1 *Any given data set $O^K = \{p^k, w^k, x^k\}_{k=1}^K$ for $K \geq 2$ contains at least one revealed demand cycle if and only if all elements $x \in \mathcal{X}^K$ in the set of extensions have an associated Slutsky matrix function that fails σ , π or ν .*

The equivalence of (2) and (3) follows from HR and Kihlstrom et al. (1976).

Next, we draw connections between some properties of the Slutsky matrix function for the set of extensions of a finite data set O^K and the length of the revealed demand cycles.

Lemma 2 *For the Slutsky matrix function, we have:*

- (i) *If there are strict $C^{D,T}$ for $T \geq 3$ and no $C^{D,T}$ for $T = 2$, then for all extensions, σ fails.*
- (ii) *If there are $C^{D,T}$ for $T = 2$, then for all extensions, ν fails.*

⁸ We note that HR provides a third characterization of rationality based on revealed demand axioms: a demand function $x \in \mathcal{X}(\mathcal{Z})$ can be generated by maximization of a continuous utility subject to a linear budget constraint if and only if the demand satisfies VARP and Weak WARP.

We proceed next to draw connections between revealed demand cycles and properties of the Slutsky matrix function:

Corollary 1 *A necessary and sufficient condition for some data set $O^K = \{p^k, w^k, x^k\}_{k=1}^K$ for $K \geq 2$ observations to contain at least one revealed demand cycle is that the true DGDF x has an associated Slutsky matrix function that fails σ , π , or ν .*

Several brief remarks are in order:

Remark 1 Our argument in the foregoing corollary applies the reasoning of Afriat's and Varian's theorems in a slightly different manner by showing that there cannot exist a Walrasian demand completion $x \in \mathcal{X}^K(\mathcal{Z})$ that at the same time can be rationalized by a utility function and made compatible with a data set that contains a revealed demand cycle. This result implies that, when there is a data set that can be rationalized in the sense of Afriat, we can find at least one demand function completion that can be rationalized in the sense of Hurwicz and Uzawa (1971).

Remark 2 The proofs of Lemma 1 and Corollary 1 use Hurwicz and Uzawa (1971) integrability theorem to define a utility function for an arbitrary compact set $\mathcal{Z} \subseteq P \times W$ instead of locally relying on a specific cycle. In other words, for the case of a data set that can be rationalized, there is at least one element $x \in \mathcal{X}^K(\mathcal{Z})$ that has a Slutsky matrix function that satisfies the regularity conditions σ , π , and ν . Therefore, we can identify such an element $x \in \mathcal{X}^K(\mathcal{Z})$ as the outcome of maximizing a unique quasiconcave utility function subject to a linear budget constraint.

4 Slutsky matrix norms for finite data: classifying consumer behavior

There is a need to classify experimental data sets about consumer behavior in a useful way, in order to understand how rationality is being violated. The SMN (Slutsky matrix norm) approach, developed in Aguiar and Serrano (2017), provides a natural way to do this since the SMN can be decomposed additively in terms of the intensity of violations of the Slutsky regularity conditions.

The SMN measures the departures from rationality by computing a norm of the smallest perturbing error matrix function E such that $S - E$ is a rational Slutsky matrix. This approach requires functional data (i.e., the knowledge of the true DGDF). However, as will be detailed in this section, the SMN can be modified to be applied in cases where we have access only to a finite data set.

Definition 7 (*SMN approach*) The measure of bounded rationality for a given Slutsky matrix function $S = D_p x + D_w x x'$ is

$$d(S) = \min\{\|E\|_W : S - E \text{ satisfies the Slutsky regularity conditions}\},$$

where $\|F\|_W^2 = \int_{z \in \mathcal{Z}} \|W(z)F(z)W(z)\|_{\mathbb{R}^{L \times L}}^2 \mu(z) dz$ is the weighted Frobenius norm with $W(z)$ a (positive semidefinite and symmetric) weighting matrix and μ a probability measure on \mathcal{Z} that is assumed to be measurable. The two main results of Aguiar and Serrano (2017) are that any Slutsky matrix function can be decomposed into two parts, a rational one and a behavioral error part. In fact, $S = S^r + E$

where $S^r = S^{\sigma,\pi,\nu}$ is the projection of S on the space of matrix functions that satisfy all three Slutsky conditions and E is the sum of three orthogonal complement matrix functions, namely, $E = E^\sigma + E^\pi + E^\nu$. Here, E^σ measures the violations of symmetry, E^π corresponds to the violations of π , and E^ν captures the violations of NSD. Finally, the Slutsky matrix norm is decomposed in three parts $\|E\|_W^2 = \|E^\sigma\|_W^2 + \|E^\pi\|_W^2 + \|E^\nu\|_W^2$.

Recall that we have the following relations:⁹

- (i) If Weak WARP holds, then $\|E^\pi\|_W = 0$ and $\|E^\nu\|_W = 0$ (i.e., the Slutsky matrix is negative semidefinite and singular in prices).¹⁰
- (ii) If the Ville Axiom holds, then $\|E^\sigma\|_W = 0$ (i.e., the Slutsky matrix is symmetric).
- (iii) If homogeneity of degree zero holds, then $\|E^\pi\|_W = 0$ (i.e., the Slutsky matrix is singular in prices).
- (iv) If the Wald axiom holds, then $\|E^\nu\|_W = 0$ (i.e., the Slutsky matrix is negative semidefinite).

The decomposition in Aguiar and Serrano (2017) remains meaningful in finite data sets. Proposition 0 has the strong implication that, if the true DGDF violates one of the axioms of revealed preference, then *all* of the extensions of O^K for any K have a positive value for the corresponding Slutsky matrix norm. We shall pursue such extensions as an interpolation exercise given the limited data set, and we shall typically seek interpolators that minimize the Slutsky norm. In doing so, the actual value of the decomposition is also exhibited. We shall close the section with a different approximation exercise: instead of seeking extensions in the set of all allowable demands, we shall approximate this functional space by means of Sieves spaces, of help for computational purposes. In what follows we will fix $W(z) = 1$, without loss of generality, and we denote the corresponding matrix function norm as $\|\cdot\|$.

4.1 The minimal Slutsky matrix norm of a limited data set

We first provide a way to measure by how much a limited data set departs from rational behavior. We define the least distance from any given data set extension to rationality under the SMN:

Definition 8 (Minimal SMN) The measure of bounded rationality for a given data set O^K is given by

$$\alpha^*(O^K) = \min_{x \in X^K(\mathcal{Z})} \|E^x\|^2 \geq 0,$$

⁹ The converse statements of each bullet point hold under additional technical regularity conditions on \mathcal{Z} , requiring no additional assumptions on O^K . In particular, if we let $\mathcal{X}(\mathcal{Z})$ be a space of smooth, Lipschitz continuous (at least in wealth) demand functions, and the space \mathcal{Z} be path-connected, convex, and open, then we have: (i) $\|E^\sigma\|_W = 0$ implies the Ville Axiom (Hurwicz and Richter 1979), (ii) $\|E^\pi\|_W = 0$ implies HD0 (this follows from the Euler’s Homogeneous Function Theorem, since $E^\pi(p, w) = 0 \iff D_p x(p, w)p + D_w x(p, w)w = 0$), (iii) $\|E^\nu\|_W = 0$ and $\|E^\pi\|_W = 0$ imply that Weak WARP holds because of (ii) and Kihlstrom et al. (1976), and (iv) $\|E^\nu\|_W = 0$ implies the Wald axiom (John 1995).

¹⁰ John (1995) established that Weak WARP implies HD0 under Walras’ law. Hence, a demand function that satisfies Weak WARP must be associated with a Slutsky matrix with both the ν and π properties.

where E^x is the Slutsky error matrix corresponding to an extension $x \in X^K(\mathcal{Z})$.

We denote $\alpha^*(O^K)$ as α^{*K} for short, as we fix the data set O^K . This is the smallest SMN-distance possible to rationality from any demand function that interpolates a data set. With the minimal Slutsky matrix norm, a classification of nonrational behavior is provided on the basis of the properties of the set of extensions. The minimal SMN is additively decomposable $\alpha^{*K} = \alpha^{\sigma,*K} + \alpha^{\pi,*K} + \alpha^{\nu,*K}$, with $\alpha^{i,*K} = \|E^{i,x^*}\|^2$ for $i \in \{\sigma, \pi, \nu\}$ and x^* such that $\alpha^{*K} = \|E^{x^*}\|^2$.

In the rest of the section we provide the statements and discussion of our main results.

Proposition 1 *A test for rationality in limited data sets corresponds to the minimal SMN $\alpha^{*K} = \min_{x \in X^K(\mathcal{Z})} \|E^x\|^2$, where $\alpha^{*K} = 0$ if and only if the data set can be rationalized (i.e. O^K satisfies SARP). Moreover, $\alpha^{\sigma,*K} = \|E^\sigma\|^2 > 0$ when there is a strict $C^{D,T}$ for $T \geq 3$ and no $C^{D,T}$ for $T = 2$, and $\alpha^{\nu,*K} = \|E^\nu\|^2 > 0$ when there is a $C^{D,T}$ for $T = 2$.*

This result provides the first measure of bounded rationality that is additively decomposable into necessary conditions for rational behavior.

Notice also the following key (but direct) property of the Minimal SMN:

Corollary 2 *The minimal SMN α^{*K} is a lower bound of the true value of $\|E^{x^*}\|^2$ for the underlying DGDF $x^* \in X^K(\mathcal{Z})$.*

In addition, the minimizers of the problem posed in Proposition 1 are interpolators of demand data that have interesting properties in their own right. Specifically, we can understand Proposition 1 as a generalized integrability result in the following sense:

Proposition 2 *If $x^{K,*} \in \operatorname{argmin}_{x \in X^K(\mathcal{Z})} \|E^x\|^2$ then:*

- $x^{K,*}$ can be rationalized (by a twice continuously differentiable strictly concave locally non satiated utility function) when $\alpha^{*K} = 0$.
- $x^{K,*}$ can be generated by maximizing a complete, regular Quah (2005) but nontransitive preference subject to $px(p, w) = w$ for all $(p, w) \in \mathcal{Z}$ when $\alpha^{*K} = \|E^\sigma\|^2 \geq 0$.
- $x^{K,*}$ can be generated by a scalar function $\phi : \mathcal{Z} \mapsto \mathbb{R}$ such that $\nabla_p \phi(p, w) = x(p, \phi(p, w))$ and $px(p, w) = w$ for all $(p, w) \in \mathcal{Z}$ when $\alpha^{*K} = \|E^\pi\|^2 + \|E^\nu\|^2 \geq 0$.

The first two cases considered in Proposition 2 provide two different notions of rationalization of a finite data set. The third case is different because it does not have an explicit notion of maximization. In fact, it can be compatible with utility minimization subject to a linear budget constraint (Hurwicz and Richter 1979). One could view the scalar function ϕ as a mapping that is sufficient for describing demand behavior locally. In particular, it can be interpreted as a function describing local preferences in the sense of Balasko and Tvede (2010) by means of a vector field.

The classification of consumer behavior on the basis of the Slutsky properties is presented in Table 1. Observe that not all possible combinations have been studied. Those that have a full characterization were presented in Proposition 2. For the rest,

Table 1 Consumer behavior classification

Slutsky matrix	Satisfies σ	Fails σ
Satisfies π, ν	Rational	Regular (Quah)
Satisfies π /Fails ν	Scalar ϕ HD1	HD0 Non-Rational (Shafer ^a)
Satisfies ν /Fails π	(Impossible) ^d	Gabaix Sparse Max ^b
Fails π /Fails ν	(Impossible) ^c	Strongly Bounded Rational ^c

only specific examples are provided. It is an open question to fully characterize them, but our methodology informs about which of them is missing.

The following are examples of models that have the combination of properties of the Slutsky matrix:

- (a) Shafer (1977) provides an example of a demand that is homogeneous of degree zero -HD0- but fails σ, ν . In particular, he proposes the following demand system for $L = 3$ and $c > 1$: $x^{sf}(p, w) = \left(\frac{w(p_1+(1-c)p_2)}{2p_1(p_1+p_2)}, \frac{w(p_1+(1-c)p_2)}{2p_2(p_1+p_2)} \frac{wcp_2}{p_3(p_1+p_2)} \right)'$,
- (b) Gabaix sparse max consumer (2014) with a Cobb-Douglas/CES utility and linear attention is an example of failures of σ, π while satisfying ν (Aguiar and Serrano 2017). In particular, the following Cobb-Douglas sparse max satisfies this property: $x^G(p, w) = \frac{1}{\alpha p_1/p_1^G + (1-\alpha)p_2/p_2^G} \left(\frac{\alpha w}{p_1^G}, \frac{(1-\alpha)w}{p_2^G} \right)'$, with a psychological price $p_l^G = m_l p_l + (1 - m_l) p_l^d$, an attention parameter $m_l \in [0, 1]$, $p_l^d \in P$ a default price, and $\alpha \in (0, 1)$ (the Cobb-Douglas share parameter).
- (c) Gabaix sparse consumer (2014) with general limited attention can generate any boundedly rational model.
- (d) John (1995) proved that Weak WARP implies HD0.
- (e) If $E^\sigma = 0$ under Walras' law it implies $E^\pi = 0$ because $p'S(p, w) = S(p, w)p$ then HD0 must hold when \mathcal{Z} is connected due to Euler's Homogeneous Function Theorem.

4.2 Computing the minimal Slutsky matrix norm: limiting behavior and sieves spaces

We next provide an upper bound (possibly not the sharpest one) that helps to compute the error of the Slutsky matrix norm associated with the error of the extension of a data set. Notice that these extensions are deterministic and they pass through all observed points; the error is with respect to the true DGDF in a domain of interest, assuming that the true DGDF is sufficiently smooth.

Here we analyze how to provide an upper bound for the distance between the true value of $E = a(x)$ for the DGDF and the finite sample values that we could get from any extension of O^K , say $x^K \in \mathcal{X}^K(\mathcal{Z})$. For a given sample and an extension (e.g., if x^K is a cubic spline interpolator), we usually have some information about the upper bound of the error $\|x - x^K\| \leq \epsilon^K$. That is, while the error is zero at the points in O^K , it may be positive in other points in the region \mathcal{Z} . Under mild assumptions, such as smoothness of the DGDF, there are known upper bounds for

ϵ^K for several interpolation techniques. In the next result we use this information to bound the distance of the values of the Slutsky error map $a(\cdot)$ for any given extension x^K .

Proposition 3 *For every O^K , with x denoting its true DGDF and any extension $x^K \in \mathcal{X}^K(\mathcal{Z})$, with $\|x - x^K\| \leq \epsilon^K$, there is a constant $\kappa' > 0$ such that $\|a(x) - a(x^K)\| \leq \kappa' \cdot \epsilon^K$, where $\kappa' = \|x\| + \|D_w[x]\|$.*

In other words, if we have enough observations such that ϵ^K is small enough, the Slutsky matrix norm of any interpolator for that data is close enough to the value of the Slutsky matrix norm of the true DGDF. Continuity of the Slutsky matrix norm error is established in Claim 2 in the ‘‘Appendix’’.

Remark 3 We can allow a family of interpolators based on the SMN. For each set O^K , there is a bracket of SMN values formed by $\alpha^{*K} = \alpha_{min}^K = \min_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$ and $\alpha_{max}^K = \max_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$. Observing that $\mathcal{X}^K \supseteq \mathcal{X}^{K+1}$ for any $K, K + 1$ with strict inclusion relation when the new observation $K + 1$ is not redundant, we have that $[\alpha_{min}^K, \alpha_{max}^K] \supseteq [\alpha_{min}^{K+1}, \alpha_{max}^{K+1}]$. We say that $x^* \equiv \lim_{K \rightarrow \infty} \mathcal{X}^K$ is the true DGDF, the unique limit of the sequence of extensions. Therefore, $\lim_{K \rightarrow \infty} [\alpha_{min}^K, \alpha_{max}^K] = \{\alpha^*\}$, the measure of bounded rationality at the true DGDF.

Next, we present a result that differs from the previous ones in one important respect. Namely, keeping the data set O^K fixed, we do not consider all its unrestricted extensions. Rather, we approximate the space of functions where we search for the minimal Slutsky norm interpolator demand function. Such an approximation result helps in the practical implementation of the results provided in Proposition 1. The test provided in that result is a variational problem that has no closed-form solution. To be able to implement this test numerically, we have to ensure that we can approximate the result with a limited functional space. In fact, consider an increasing sequence of sets of demand functions $\{\mathcal{X}_h\}_{h \in \mathbb{N}}$ with $\mathcal{X}_h \subset \mathcal{X}$ and $\mathcal{X}_{h_1} \subset \mathcal{X}_{h_2}$ if $h_2 > h_1$; here, for any $x \in \mathcal{X}$, there is a (projection) mapping $\phi_h : \mathcal{X} \rightarrow \mathcal{X}_h$ with $\phi_h \circ x \in \mathcal{X}_h$ such that $\phi_h \circ x \rightarrow x$ as $h \rightarrow \infty$. This is usually called a Sieves space (e.g., think of a space of polynomial functions approximating our space of demand functions, in which h is the highest degree of the polynomial). We define analogously the allowable set of extensions of a data set $\mathcal{X}_h^K(\mathcal{Z}) = \mathcal{X}^K(\mathcal{Z}) \cap \mathcal{X}_h$ under the h^{th} approximating space. We only consider the case when $\mathcal{X}_h^K(\mathcal{Z})$ is nonempty, in other words, h is sufficiently large to guarantee nonemptiness given the data set. The next result shows that the correspondence of Slutsky norm minimizers is ‘‘continuous at $h = \infty$ ’’, or more formally, a consistent approximation.

Proposition 4 *Fix a data set O^K and the corresponding set of extensions $\mathcal{X}^K(\mathcal{Z})$. An approximate SMN test for rationality in such a limited data set corresponds to the solution to the optimization problem $\alpha_h^* = \min_{x \in \mathcal{X}_h^K(\mathcal{Z})} \|E^x\|^2$, with $\alpha_h^* \rightarrow \alpha^{*K}$ as $h \rightarrow \infty$, where $\alpha^{*K} = \min_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$. Moreover, for any $x^* \in \operatorname{argmin}_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$, there exists a sequence $x^{*,h} \rightarrow x^*$ as $h \rightarrow \infty$ where $x^{*,h} \in \operatorname{argmin}_{x \in \mathcal{X}_h^K(\mathcal{Z})} \|E^x\|^2$.*

Remarks on computational complexity

The results in this section provide guidelines for practical work. If we have a large data set (with K large enough) we could use any interpolator, possibly a fast one for computational reasons, and it will be close to the true DGDF and the distance to the true Slutsky matrix norm will be small as well. In contrast, if K is small, we could use a large Sieves space (h large enough) to find an interpolator that has a Slutsky matrix norm that is close enough to the lower bound of the true value.

We emphasize that the computational complexity of Spline Interpolation for order of polynomials h , which is an example of a sieve space, and for K interpolation points is $O(h^2 K)$ for the univariate case (Toraichi et al. 1987). In the multivariate case, we consider tensor products of univariate splines that depend on the dimension of the consumption space L . Since the derivatives of the interpolator can be obtained in a recursive way from the splines basis, in practice the derivation step adds no substantial computational time, other than evaluation, which can be vectorized in the algorithmic implementation.

The computation of the Slutsky matrix norm is trivial, since Aguiar and Serrano (2017) provide a close form expression for it. The only object needed to compute the Slutsky matrix norm of a given interpolator is the computation of the negative semidefinite part of a matrix and its residual E^v at all points in the data set. The complexity of the pointwise eigenvalue decomposition of the Slutsky matrix of size $L \times L$ is $O(L^3)$ (Pan and Chen 1999). In this sense, our approach permits the user to adjust the computational complexity of its computation with a sacrifice in numerical precision. By choosing low degree spline interpolation we are effectively close to a parametric exercise, close in spirit to Halevy et al. (2014). In contrast, by increasing the size of h we approximate a fully nonparametric approach to bounded rationality. Our approach will become computationally more complex when the consumption space dimension grows. In this regard, the computational flexibility of our approach ought to be contrasted with the computational complexity of other methods to measure departures from rationality. Namely, the Money Pump Index is known to be NP-hard (Dean and Martin 2015), and the Minimum Cost Index is also NP-hard (Shiozawa et al. 2015). The AEI does not suffer from the computational issues of the indices just mentioned.

5 Simulation study

In order to complement our results for the computation of the minimal Slutsky norm interpolators in finite samples with restricted functional spaces (e.g., piecewise polynomial interpolators of finite degree), we provide a simulation study of the behavior of the testing procedures provided in Sect. 4 for a small number of observations and for interpolators chosen from small functional spaces. The results are encouraging. They exhibit fast convergence when the true DGDF corresponds to a classical boundedly rational model proposed by Shafer (1977):

- $x_1^{sf}(p, w) = \frac{w(p_1 + (1-c)p_2)}{2p_1(p_1 + p_2)},$

- $x_2^{sf}(p, w) = \frac{w(p_1+(1-c)p_2)}{2p_2(p_1+p_2)},$
- $x_3^{sf}(p, w) = \frac{wc p_2}{p_3(p_1+p_2)}.$

The parameter $c \in \mathbb{R}$ controls the violations of rationality. When $c = 0$, this demand system corresponds to a rational Cobb-Douglas consumer with the same preferences over the first two goods and no interest in consuming the third one. Homogeneity of degree 0 and Walras’ law hold for all values of c . For $0 < c < 1$, the only property of the Slutsky matrix being violated is symmetry (σ). For $c > 1$, both symmetry (σ) and NSD (v) fail.

We implement our simulation study in the budget-share form of this demand function, and we focus on computing the lower bound of the budget-share elasticities version of the Slutsky matrix norm (Aguilar and Serrano 2017). Formally, the budget-share elasticity Slutsky matrix norm is the Slutsky matrix norm with a weight matrix $W(p, w) = \frac{1}{\sqrt{w}}diag(p)$. This variant of the Slutsky matrix norm is unit-free. The Slutsky matrix norm is the unknown quantity of interest that we want to bound below, and in general, to obtain a good approximation, we use $\|E^{sf}\|_W = (\int_{z \in \mathcal{Z}} \|E^{x^{sf}}(p, w)\|_{\mathbb{M}, W}^2 \mu(z) dz)^{1/2}$, where $\|\cdot\|_{\mathbb{M}, W}$ is the weighted Frobenius matrix norm in the space $\mathbb{R}^{L \times L}$, and $\mu : \mathcal{Z} \mapsto [0, 1]$ is a probability density function that corresponds to the sampling scheme chosen by an observer of prices and wealth. For us, $\mu(z)$ corresponds to a probability density such that the vector $log(z)$ has a multivariate normal distribution probability formed as the product of $L + 1$ normal distributions with zero mean and standard deviation $1/20$.

Notice that μ is known to the experimenter because it is the sampling scheme generator process. The choice of this particular distribution is irrelevant and we do not need to estimate it. We note that our exercise is deterministic.

In practice, the observer obtains only a finite sample $O^K = \{p^k, x^k\}_{k=1}^K$ such that $w^k = p^k x^k$ and in this example $x^k = x^{sf}(p^k, w^k)$. We obtain the budget shares of the data $b_l^k = p_l^k x_l^k / w^k$.

The budget-share form of the demand system expressed in terms of logarithms of prices and logarithm of wealth, which we take as the true DGDF is:

- (i) $b_1^*(ln(p), ln(w)) = \frac{p_1+(1-c)p_2}{2(p_1+p_2)}.$
- (ii) $b_2^*(ln(p), ln(w)) = \frac{p_1+(1-c)p_2}{2(p_1+p_2)}.$
- (iii) $b_3^*(ln(p), ln(w)) = 1 - b_1^*(ln(p), ln(w)) - b_2^*(ln(p), ln(w)).$

Here, the vector entry $ln(p)_i = ln(p_i)$. The Slutsky matrix at a point (p, w) , in its budget-share elasticity form, can be computed for numerical stability purposes from the budget shares expressed in terms of logged prices and wealth as follows:

$$s_{ij}(p, w) p_i p_j / w = \frac{\partial b_i(ln(p), ln(w))}{\partial ln(p_j)} + \frac{\partial b_i(ln(p), w)}{\partial ln(w)} b_j(ln(p), ln(w)) + b_j(ln(p), ln(w)) b_i(ln(p), ln(w)) - \delta_{ij} b_i(ln(p), ln(w)),$$

where $\delta_{ij} = 1$ if $i = j$ and zero otherwise.

Using corollary (1) we approximate the lower bound of $\|E^{sf}\|_W$ for the fixed sample O^K . The lower bound is $\alpha^K = \min_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|_W$ where $\mathcal{X}^K(\mathcal{Z})$ is the set

Table 2 Numerical approximation of the square of the minimal Slutsky norm of a set of extensions, for different sample size and different size of functional space of interpolators

Sample	Degree	$(\gamma^{K,h})^2 = \min_{b \in \mathcal{X}_h^{K,b}(Z)} \ E^b\ _W^2$		
		$c = 0$ ($\ E\ _W^2 = 0$)	$c = 1/2$ ($\ E\ _W^2 = 0.0241$)	$c = 2$ ($\ E\ _W^2 = 0.511$)
$K = 20$	$h = 5$	0.003	0.026	0.487
	$h = 7$	0.002	0.025	0.454
	$h = 11$	0.000	0.025	0.465
$K = 50$	$h = 5$	0.004	0.0239	0.5122
	$h = 7$	0.004	0.0246	0.5102
	$h = 11$	0.005	0.0240	0.5112
$K = 100$	$h = 5$	0.0002	0.0241	0.5105
	$h = 7$	0.0002	0.0241	0.5106
	$h = 11$	0.0002	0.0241	0.5106

of extensions of O^K that belongs to $\mathcal{X}(Z)$. We expand on the methodological details in “Appendix” (9.8).

The effects of the number of observations and polynomial degrees on the accuracy of the estimation We obtain a random sample of size $K \in \{20, 50, 100\}$ i.i.d. draws from $\mu(z)$ and x^{sf} with its budget share b^{sf} such that we obtain: $O^K = \{p^{m,k}, b^{m,sf}(p^{m,k})\}_{k=1}^K$ for each value of K . We fix a budget-share space of extensions of O^K , $\mathcal{X}_h^{K,b}(Z)$ for $h \in \{5, 7, 11\}$.

The key parameter c is fixed at three different values $c \in \{0, \frac{1}{2}, 2\}$, to generate different intensities of both types of violations (i.e., of σ and ν). We estimate the Budget-Share Slutsky matrix, The results of the simulation are provided in Table 2 and show that even for $K = 50$, which is a common data set size (see Ahn et al. 2014) the SMN test works as well as the revealed preference test. In addition, the quality of the approximation to the true value increases in both the sample size K and the degree of the polynomial h .

The effects of sample size on the decomposition of violations A key feature of our methodology is the decomposition of the size of bounded rationality into its components (σ, π, ν) . By construction, we can obtain the lower bounds of each type of violation. In this example the approximated values will be contrasted with the true values to show that the performance of our procedure is encouraging here as well. For concreteness, we present only the results for $h = 5$ for each element of the decomposition in a vector.

The results of the simulation are provided in Table 3 and show that even for $K = 50$ we can get small absolute errors when estimating the SMN using standard non parametric estimators of the conditional expectation. In any case, the results of these simulations are encouraging in terms of their speed of convergence of an out-of-the-box estimator of the SMN to its true value.

Table 3 Numerical approximation of the square of the minimal Slutsky norm of a set of extensions, for different values of c

Sample	$(\alpha^{K,h})^2 = \min_{b \in \mathcal{X}_h^{K,b}(\mathcal{Z})} \ E^b\ _W^2$			
	Degree	$c = 0$	$c = 1/2$	$c = 2$
True	$\ E^{b,\sigma}\ _W^2$	0	0.0241	0.381
	$\ E^{b,\pi}\ _W^2$	0	0	0
	$\ E^{b,v}\ _W^2$	0	0	0.130
$K = 20$ $h = 5$	$\ E^{b,\sigma}\ _W^2$	0.00120	0.02534	0.36140
	$\ E^{b,\pi}\ _W^2$	0.00210	0.00136	0.00292
	$\ E^{b,v}\ _W^2$	0.00022	0.00000	0.12299
$K = 50$ $h = 5$	$\ E^{b,\sigma}\ _W^2$	0.00004	0.00233	0.38146
	$\ E^{b,\pi}\ _W^2$	0.00036	0.00006	0.00070
	$\ E^{b,v}\ _W^2$	0.00004	0.00000	0.13006
$K = 100$ $h = 5$	$\ E^{b,\sigma}\ _W^2$	0.00000	0.00237	0.38042
	$\ E^{b,\pi}\ _W^2$	0.00020	0.00004	0.00056
	$\ E^{b,v}\ _W^2$	0.00000	0.00000	0.12677

Relation to AEI

We continue our simulation study, looking at the comparison between the popular AEI and the minimal SMN. We consider only $h = 5$ and $K = 50$, while the data generation of prices and wealth is the same as in the previous sections.

We allow the bounded rationality parameter in the Shaffer model, c to take values from 2 to 20. We chose the c parameter numbers to get a range in the AEI roughly from 0.99 to 0.95. We observe that both the AEI and the minimal SMN grow monotonically in the c parameter. In this sense, both the AEI and SMN are ordinally equivalent for this example. We provide a formal study of the connection between the two indices in the sequel. Here we confine ourselves to the numerical properties of the measures of departures from rationality.

Afriat (1973) considers that a high AEI (very close to 1) is evidence of a high goodness of fit of the rationality consumer. Although there is no formal threshold of when to consider that the AEI is high enough to be deemed close to rationality, it is common to use a threshold of 0.95 or 0.99.¹¹

Our results presented in Table 4 provide numerical evidence that very high levels of the AEI (0.965), which are usually interpreted as evidence in favor of the rationality hypothesis, can be associated to data sets with a very high SMN. Recall that a high SMN implies that the modeller is making a high prediction error. In fact, the α minimal SMN measures the minimal error a modeller would make in predicting the elasticity behavior (for price changes when they are compensated in the Slutsky sense) under the assumption of rationality. Indeed, we observe that in the worst case, which corresponds

¹¹ See Ahn et al. (2014), Andreoni et al. (2005), Varian (1990).

Table 4 Numerical approximation of the square of the minimal Slutsky norm of a set of extensions, for different sample size and different size of functional space of interpolators

Sample	Degree	$(\alpha^{K,h})^2 = \min_{b \in \mathcal{X}_h^{K,b}(z)} \ E^b\ _W^2$			
		$c = 2$	$c = 5$	$c = 10$	$c = 20$
True	AEI	0.9999	0.9947	0.9829	0.9650
	$\ E^b\ _W^2$	0.511	34.895	957.25	19605.5
	$\ E^{b,\sigma}\ _W^2$	0.381	2.3988	9.5952	38.381
	$\ E^{b,\pi}\ _W^2$	0	0	0	0
$K = 50$ $h = 5$	$\ E^{b,\nu}\ _W^2$	0.130	32.496	947.66	19567.1
	$\ E^b\ _W^2$	0.5122	34.9293	957.2055	19602.243
	$\ E^{b,\sigma}\ _W^2$	0.38146	2.4069	9.6125	38.408
	$\ E^{b,\pi}\ _W^2$	0	0	0	0
	$\ E^{b,\nu}\ _W^2$	0.13006	32.5224	947.5930	19563.835

to a parameter $c = 20$, the minimal SMN calculated using our methodology is $\alpha = \sqrt{1960.243} = 140.01$. Our numerical computation of the minimal SMN is very close to the true value ($\alpha^{50,5} = 140.008$). When the AEI is very high (0.9999) we have an associated SMN index of $\alpha = \sqrt{0.511} = 0.71$, which is smaller, but still sizeable. Our numerical computation of the SMN is very close to the true value ($\alpha^{50,5} = 0.716$).¹²

We emphasize that, beyond the obvious benefits of having different measures of rationality to evaluate whether we can consider a data set close to the null of utility-maximizing behavior, the SMN provides a useful decomposition of the measure into the Slutsky regularity conditions. We observe that for most values of the parameter c , the error that corresponds to violations of the law of demand (ν) are more important than the violations that correspond to the property σ . We also observe that higher values of the bounded rationality parameter c are associated with a higher relative weight of the violations of the law of demand in terms of the total SMN. We can observe how this information provides us with an additional insight about the nature of bounded rationality in this model. If a modeller wants to effectively predict the elasticity behavior of this boundedly rational consumer, she should drop the regularity conditions ν and σ , with a priority on the former.

Applications to experimental data sets

We used the methods developed in this paper in the online Supplement to Aguiar and Serrano (2017). In particular, we compute the Minimal SMN for the experimental data

¹² To further clarify what we mean by a high level of the SMN, let us recall that $\alpha = 0$ if and only if the data is rationalizable. In addition, α is a unitless index that measures the quadratic norm of the errors in the prediction of the elasticity of the expenditure on all goods, as a fraction of total wealth at any given experimental trial. The expenditure elasticity is taken with respect to a change of 1 percent of all prices, when this change is compensated in the sense of Slutsky.

set of Ahn et al. (2014), which consists of $K = 50$ budget allocation tasks regarding the purchase of Arrow securities.

In that experimental data set, we find evidence that the violations of ν are less prevalent in the sample than the violations of σ . The violations of π are not possible due to the experimental design. We refer the reader to the online Supplement to Aguiar and Serrano (2017) for details.

6 Generalized SMN and its connections to other measures of bounded rationality

6.1 The SMN and its local convergence to the afriat approach through the Jerison-Jerison index

The most well-known measure in the RP approach is the AEI. An important question is how the Slutsky matrix norm relates numerically to this index. We provide its definition. :

Definition 9 (*Afriat Efficiency Index, AEI*) The AEI of a data set O^K , with at least one $C^{D,T}$ with $T \geq 2$ is:

$$\mathcal{G} = \sup_{C^{D,T}} \{ \min_{t \in \{0, \dots, T\}} \{ p^t [x^t - x^{t-1}] / [p^t x^t] \} | (p^t, x^t) \in C^{D,T} \},$$

and zero otherwise.

AEI is such that $\mathcal{G} \in [0, 1]$ and researchers interpret its measure as a “loss of efficiency” or distance from a rational behavior benchmark: the larger it is, the larger gap between actual expenditure and expenditure on a nonchosen bundle that becomes chosen for different prices, suggesting strong swings of (irrational) behavior. For a fixed cycle $C^{D,T}$ we call $\min_{t \in \{0, \dots, T\}} \{ p^t [x^t - x^{t-1}] / [p^t x^t] \}$ $(p^t, x^t) \in C^{D,T}$ the slope of the cycle.

The connection between the SMN approach and AEI is established through the Jerison-Jerison modified Slutsky matrix index. Indeed, an earlier approach, also based on the Slutsky matrix, is proposed in the work of Jerison and Jerison (2012), where a local measure of departures from rationality is proposed in the form of a family of indices.

Definition 10 (*JJ-Index*) The Jerison and Jerison index is defined as

$$\gamma(\{v^t\}_{t=0}^T, S(\bar{p}, \bar{w})) = \frac{1}{T} \sum_{t=1}^T v^t S(\bar{p}, \bar{w}) [v^t - v^{t-1}],$$

where $S(\bar{p}, \bar{w})$ is the Slutsky matrix evaluated at a reference point $(\bar{p}, \bar{w}) \in P \times W$ and $\{v^t\}_{t=0}^T$ is a collection of $v^t \in \mathbb{R}^L$ vectors such that $v^0 = v^T$.

Jerison and Jerison (2012) establish a remarkable connection between Afriat’s inefficiency cost measure of irrationality and the JJ index. Both coincide in a local ball

around some price and wealth reference point. Here we will revisit this result, showing that their measure can be seen as a quadratic form associated with the minimal correcting matrix E from Aguiar and Serrano (2017).

6.2 A generalization of the Slutsky matrix norm

Our intent in this section is to focus on a generalized SMN that has as a special case the index of Jerison and Jerison (2012), and relate it to the Afriat index. The space of matrix-valued functions $\mathcal{M}(\mathcal{Z})$ has some nice properties associated with the Slutsky regularity conditions σ, π, ν . As shown in Aguiar and Serrano (2017) the spaces $\mathcal{M}(\mathcal{Z})_\sigma$ of symmetric matrices, $\mathcal{M}(\mathcal{Z})_{\sigma,\pi}$ of symmetric matrices that are singular in prices, and $\mathcal{M}(\mathcal{Z})_{\sigma,\pi,\nu}$ of symmetric, singular in prices, and negative semidefinite matrices are closed and convex subspaces under the topology induced by the Weighted Frobenius norm. Using the weighted inner product $\langle F, H \rangle_W$ in the space, we define the projections of any Slutsky matrix $S \in \mathcal{M}(\mathcal{Z})$ on the spaces of interest. With it, we have the same unique orthogonal decomposition as in Aguiar and Serrano (2017) with $S = S^r + E$ with $E = E^\sigma + E^\pi + E^\nu$. However, it is possible to change the norm, in fact we can use any pseudonorm or quasinorm to compute our measure of bounded rationality.

Definition 11 (*Generalized Slutsky Error Seminorm*) A generalized Slutsky error seminorm is a mapping $\Gamma = n \circ a$ such that is the composition of a $n : \mathcal{M}(\mathcal{Z}) \mapsto \mathbb{R}_+$ seminorm on the space $\mathcal{M}(\mathcal{Z})$ and the Slutsky error mapping $a : \mathcal{X}(\mathcal{Z}) \mapsto \mathcal{M}(\mathcal{Z})$ (i.e., $a(x) = E$).

We establish now that the JJ index can be obtained as a subcase of the Generalized Slutsky error seminorm. The JJ index definition depends on the Slutsky matrix $S = S^r + E$, not only its error matrix E . Ideally, we want a measure of violations of rationality that depends only on the departures of rationality and on nothing else (i.e. on the matrix E and not on the matrix S^r). We show this is the case for the supremum of the JJ index. We next prove a nice feature of the discrete JJ index that, to the best of our knowledge, had not been previously pointed out.

Recall from Aguiar and Serrano (2017) that $E = E^\sigma + E^\pi + E^\nu$ is defined as the minimal perturbation matrix function that yields a rational Slutsky matrix function, decomposed in its symmetry, singularity, and NSD error parts. Then:

Lemma 3 *The JJ index is equal to*

$$\gamma(\{v_t\}_{t=1}^T, S^x(\bar{p}, \bar{w})) = \gamma(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w})) + \sum_{j \in \{\sigma, \pi, \nu\}} \gamma(\{v_t\}_{t=1}^T, E^{x,j}(\bar{p}, \bar{w})).$$

This result shows that the JJ index can be decomposed additively into the JJ index of a rational Slutsky matrix, which is always nonpositive $\gamma(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w})) \leq 0$, because the eigenvectors of S^r are all nonpositive and the JJ index is an average of quadratic forms, and an unsigned, sometimes nonnegative, part (i.e., it may be nonnegative for some cycle $\{v_t\}_{t=1}^T$).

Looking at the case of infinite data sets, we can begin with a local result, and then provide a global connection as well. Jerison and Jerison connect the JJ index with the Afriat index. Here we establish the connection with the SMN using our previous result. We define $V^{T(\epsilon_p)}$ as the set of sequences of the type $\{v^t\}_{t=0}^T$ where $v^t \in V \subset \mathbb{R}^L$. The set V is compact, $v^0 = v^T$ and $T \geq 1$ is fixed. Additionally, defining $p^t = \bar{p} + v^t$ and $|v^t| = |\bar{p} - p^t| \leq \epsilon_p$ for all $t \in \{0, \dots, T\}$, we require that the related cycle $C^{D,T} = \{p^t, w^t, x(p^t, w^t)\}_{t=0}^T$ is a revealed demand cycle. Moreover, we assume that $|w^t - \bar{w}| \leq \epsilon_p \eta$ where $\eta > \max\{v \in V \mid |v^T x(\bar{p}, \bar{w})|\}$. Let $V_w^{T(\epsilon_p, \eta)}$ be the set of cycles of the type $\{v^t, u^t\}_{t=0}^T$ where $u^t = w^t - \bar{w}$ such that $|u^t| \leq \epsilon_p \eta$. Using this notation, we can rewrite the Afriat index of an RP cycle as $\mathcal{G}((\bar{p}, \bar{w}) + V_w^{T(\epsilon_p, \eta)}) = \sup_{\{v^t, u^t\}_{t=0}^T \in V^{T(\epsilon_p, \eta)}} \{\min_{t \in \{0, \dots, T\}} \{p^t[x^t - x^{t-1}]/[p^t x^t]\} (\bar{p} + v^t, \bar{w} + u^t x^t)\} = C^{D,T}(\{v^t, u^t\}_{t=0}^T \in \{v^t, u^t\}_{t=0}^T)$ (i.e., the supremum over the Afriat inefficiency index if we are allowed to sample price and wealth at will). Now we establish the desired result regarding the supremum of the JJ index in $V^{T(\epsilon_p)}$ and its relation with the Generalized Slutsky error matrix norm.

Lemma 4 *Given a set of cycles $V^{T(\epsilon_p)}$ and for a fixed $T \geq 1$ we have the supremum of the JJ index $\sup\{\gamma(\mathbf{v}, S(\bar{p}, \bar{w})) : \mathbf{v} \in V^{T(\epsilon_p)}\} = \sup\{\gamma(\mathbf{v}, E(\bar{p}, \bar{w}))\}$, $\mathbf{v} \in V^{T(\epsilon_p)}$, where $S(\bar{p}, \bar{w}) = S^r(\bar{p}, \bar{w}) + E(\bar{p}, \bar{w})$.*

The statement says that the supremum of the JJ index is the composition of a pseudonorm on the space of matrix functions and the mapping a , because it only depends on the matrix E . This holds only if $V^{T(\epsilon_p)}$ is a compact neighborhood. Evidently, the SMN is a special case of the generalize definition that we have proposed in this subsection. The SMN is essentially the only norm that is associated with the inner product of the space $\mathcal{M}(\mathcal{Z})$, and for that reason, it provides the orthogonal decomposition that allows us to classify violations of the Slutsky regularity properties, but in some cases like the present one we can investigate other norms to understand its connections with other notions of “distance” from rationality.

6.3 Convergence results

With the previous results in hand, the next result requires that the analyst be able to sample at will from prices and wealth and that the Slutsky matrix be known at the reference point, which requires infinite data. In contrast, we have been assuming in this paper that we are given a data set O^K so the sampling process is given. Hence, the next proposition should be seen mainly as a way to connect theoretically the traditional RP Afriat inefficiency measure of rationality and the SMN approach in arbitrarily large data sets. In words, it says that when the ratio $\epsilon_p \rightarrow 0$, the Afriat inefficiency index converges to zero at the rate of ϵ_p^{-2} , in proportion to the supremum of the JJ index in the set of cycles considered. The key assumption is that wealth across the path is such that $|w^t - \bar{w}| \leq \epsilon_p \eta$:

Proposition 5 *Given a set of cycles $V^{T(\epsilon_p)}$ and $V_w^{T(\epsilon_p, \eta)}$, for a fixed $T \geq 1$, the Afriat inefficiency index converges to zero proportionally to the supremum of the JJ index of the Slutsky matrix error norm $E(\bar{p}, \bar{w})$:*

$$\lim_{\epsilon_p \rightarrow 0} \epsilon_p^{-2} \mathcal{G}((\bar{p}, \bar{w}) + V_w^{T(\epsilon_p, \eta)}) = \sup\{\gamma(v, E(\bar{p}, \bar{w}))/\bar{w} : v \in V^{T(\epsilon_p)}\}.$$

6.4 The behavioral nearness problem and the “Almost Implies Near” principle

We refer to the problem of trying to find the closest rational demand to a given demand as the “behavioral nearness” problem (Anderson 1986). At a general level, as discussed in Aguiar and Serrano (2017), the behavioral nearness problem presents several serious difficulties, which was our motivation to take the Slutsky matrix route. Intuitively, there should be a close relationship between the two approaches. In order to make this relationship explicit, we will make extensive use of Anderson (1986) “almost implies near” (AN) principle and its recent elaboration, developed by Boualem and Brouzet (2012). At the end of the subsection we adapt the principle to finite data sets.

Assume that the set of demand functions $\mathcal{X}(\mathcal{Z})$ is equipped with the norm $\|f\|_{C1} = \max(\{\|f_l\|_{C1,1}\}_{l=1,\dots,L})$, with $\|f_l\|_{C1,1} = \max(\|f_l\|_{\infty,1}, \|\nabla f_l\|_{\infty,L+1})$ where $f(z) = [f_1(z) \dots f_L(z)]' \in \mathbb{R}^L$.¹³ Let $\mathcal{R}(\mathcal{Z}) \subset \mathcal{X}(\mathcal{Z})$ be the set of rational demand functions (i.e., $x^r \in \mathcal{R}(\mathcal{Z})$ is the solution to maximizing a complete, locally nonsatiated and transitive preference over a linear budget constraint).

Definition 12 Define the distance of $x \in \mathcal{X}(\mathcal{Z})$ to the set of rational demands $\mathcal{R}(\mathcal{Z})$ by the “least” distance from an element to a set: $\|e\|_{C1} = \inf\{\|x - x^r\|_{C1} \mid x^r \in \mathcal{R}(\mathcal{Z})\}$.

The matrix nearness problem allows us to represent property \mathfrak{R} —a shorthand to refer to the Slutsky regularity properties—by the Slutsky correction map a with the AN property, as defined next.

Definition 13 (Boualem and Brouzet 2012) A function $a : G \mapsto Y$ (with G and Y metric spaces) satisfies the “almost implies near” (AN) property at $C \in Y$, if for all $\epsilon > 0$, there exists $\delta > 0$ such that for every $g \in G$, the inequality $\|a(g) - C\|_Y < \delta$ implies the existence of an element $g_0 \in G$ satisfying $a(g_0) = C$ and $\|g_0 - g\|_G < \epsilon$.

The Slutsky correction mapping $a : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$, with $\mathcal{X}(Z)$ and $\mathcal{M}(Z)$ as defined above, represents property \mathfrak{R} when for every $x \in \mathcal{R}(Z)$:

$$a(x) = E = 0,$$

where 0 represents the zero matrix function of $L \times L$ dimension in the metric space $\mathcal{M}(Z)$, and $E = S^r - S$ denotes a solution of the program in the definition of the Slutsky norm. The analytical expression of the (unique) solution to such a problem, as well as its properties, were derived previously (Aguiar and Serrano 2017).

The “almost implies near” (AN) principle allows us to assert that for all $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in \mathcal{X}(Z)$, whenever $a(x) = E$ and $\|E\| < \delta$, one can find $x^r \in \mathcal{X}(Z)$ such that $a(x^r) = 0$ and $\|x - x^r\|_{C1} < \epsilon$. In the name of

¹³ In fact, any norm that makes $\mathcal{X}(Z)$ complete works. We use also the related normed space of real-valued functions with the norm $\|\cdot\|_{\infty,m} = \sup_{z \in Z} |g(z)|$ for $g : Z \mapsto \mathbb{R}^m$, for finite $m \geq 1$ and $|\cdot|$ the absolute value. This norm will come in handy when dealing with some technical proofs in the sequel. By the set $\mathcal{C}^1(Z)$ we mean the set of functions that have \mathcal{C}^1 extensions to an open set containing the compact domain Z .

the principle, the “almost” part refers to $\delta > 0$ or property \mathfrak{R} being δ -satisfied (matrix nearness), and the “near” part to $\epsilon > 0$ (behavioral nearness).¹⁴

Given the additional restriction that $\mathcal{X}(Z)$ is compact, we have the following result, which links the Slutsky matrix norm approach with the Varian and Jerison-Jerison measure of the “least distance” to the closest rational demand function in the compact subset $\mathcal{X}(Z)$.

Proposition 6 *For all $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in \mathcal{X}(Z)$, the inequality $\|E\| \leq \delta$ implies the existence of a rational demand function $x_1^r \in \mathcal{X}(Z)$ satisfying $a(x_1^r) = 0$ and $\|x - x^r\|_{C1} \leq \epsilon$. Moreover if $\delta \rightarrow 0$ then $\epsilon(\delta) \rightarrow 0$.*

Thus, there is an ordinal equivalence between the two problems, in the sense that if one takes two degrees of approximation $\epsilon_1 \leq \epsilon_2$, we have that $\delta_1 \leq \delta_2$. We underscore the fact that the AN property is stated for every $\epsilon > 0$, not necessarily arbitrarily small, and therefore, we are able to account for violations of rationality of any “size”. In particular, we have established an equivalence of testing rationality in the behavioral nearness setting as in Epstein and Yatchew (1985) and in our Slutsky matrix norm framework.

For the case of finite data sets, this result can be used to establish the equivalence between the behavioral nearness problem and the Minimal SMN.

Definition 14 (*Finite data behavioral nearness*) The behavioral nearness measure of bounded rationality for a given data set O^K is given by

$$\beta^*(O^K) = \inf_{x^r \in \mathcal{R}^K(Z)} \|x - x^r\|_{C1}^2 \geq 0,$$

where $\mathcal{R}^K(Z) = \mathcal{R}(Z) \cap \mathcal{X}^K(Z)$.

We can establish the following corollary.

Corollary 3 *For all $\epsilon > 0$, there exists $\delta > 0$ such that for all O^K , the inequality $\alpha^*(O^K) \leq \delta$ implies that $\beta^*(O^K) \leq \epsilon$. Moreover, if $\delta \rightarrow 0$ then $\epsilon(\delta) \rightarrow 0$.*

7 Literature review

The current research is related to the work initiated by Chiappori and Rochet (1987), Matzkin and Richter (1991) and culminated by Lee and Wong (2005) that prove the equivalence of SARP and rationalization by a smooth demand generated by maximizing a smooth utility function. Their work unifies the binary tests of rationality from both RP and smooth approaches. Our results focus on the realm of quantifying departures from rationality. In particular, we provide a minimal Slutsky norm interpolator of any finite data set, in the form of a continuously differentiable demand. When the data set satisfies SARP, then the minimal Slutsky norm is zero and the interpolator can

¹⁴ Instead of relying on nonstandard analysis, Boualem and Brouzet (2012) use functions between metric spaces to represent a property in Anderson’s language, and a metric to represent his formulas. This treatment is also useful because it allows us to adapt our results in order to derive an explicit expression for $\epsilon(\delta)$ for an arbitrary $Z \subset P \times W$ when dealing with compact subsets of $\mathcal{X}(Z)$.

be generated by maximizing a twice continuously differentiable utility function as in Hurwicz and Uzawa (1971). Our results provide a more general sense of integrability, since if the data set fails SARP but satisfies WARP (known also as Weak WARP), the minimal Slutsky norm interpolator can be generated by regular preferences in the sense of Quah (2005), that are nontransitive. Also related is the recent work of Nishimura et al. (2013), which provides a bridge between the revealed preference approach for limited data sets and the Richter congruence axiom (1996) for complete data sets.

Jerison and Jerison (2012) is the closest work to ours. They provide a study of the relation of the Slutsky matrix and revealed demand cycles that generalizes the findings of Shafer (1977). However, our work differs fundamentally from Jerison and Jerison (2012), among other things, in that we assume a finite data set of prices and observed choices, while these authors assume an infinite amount of observations. Because we never have access to infinite data, our work is the first to allow the use of the Slutsky matrix in a laboratory data set environment.

We have provided a complementary framework to the RP approach to testing consumer behavior, as initiated by Varian (1983), Afriat (1973), and continued in the works of Echenique et al. (2011), Dean and Martin (2015), Apesteguia and Ballester (2015) and others. Our measure differs from this tradition in that ours is a positive measure of bounded rationality, in contrast to the more common welfare or normative approaches to measure bounded rationality, which implicitly assume that the consumer wants to optimize, failing in his goal. In contrast, we are concerned with failures in doing comparative statics analysis for any given consumer behavior when we assume rationality (possibly contrary to the fact).

We leave as an open avenue of research to relate our results with the contribution of Halevy et al. (2014), which proposes a parametric procedure to recover preferences based on minimizing a measure of behavioral closeness that considers the specification error (due to the choice of a parametric family) and an inconsistency index that uses the money-metric notion (a normative index of distance from rationality). We conjecture that we can replace the money metric by a parametric Slutsky matrix norm to produce a demand function that minimizes the comparative statics errors in prediction given the rationality restriction and the data.

8 Conclusion

We have provided a framework to adapt the results in our companion paper (Aguar and Serrano 2017) to finite data. In doing so, we have presented a unification of the RP and Hurwicz-Uzawa characterizations of rationality in consumer theory. In particular, we have concluded that: (i) a limited data set that satisfies SARP admits at least one continuous differentiable extension (demand function that interpolates the data) that is integrable in the sense of Hurwicz-Uzawa; and (ii) the violations of the Slutsky regularity conditions by the set of extensions of a limited data set correspond to demand cycles of different length in the RP approach. One advantage of our unified approach is that it allows a classification of consumer behavior. Finally, the modified SMN measure of departures from rationality is a positive or objective index, and as such, it allows interpersonal comparisons. We have tested it in simulations, and

applied it elsewhere to experimental data. Our empirical findings are that insisting on the Slutsky matrix symmetry hypothesis leads to making a higher error than assuming its NSD. To the best of our knowledge, this is the first evidence brought forward on the claims of the higher robustness of WARP vis-à-vis the Ville axiom, as already suggested in Kihlstrom et al. (1976).

Compliance with ethical standards

Ethical Statement (i) We declare that our research article is in compliance with the Ethical Standards of the SERIEs and our home institutions The University of Western Ontario and Brown University. (ii) We have not received any funding for this research. (iii) We declare we have not conflict of interest regarding this research. (iv) We do not need an ethical approval as our paper is a theoretical piece. (v) We do not need an informed consent as our paper is a theoretical piece.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

Proof of Lemma 1

Proof Suppose that all elements $x \in \mathcal{X}^K$ in the set of extensions have an associated Slutsky matrix norm error that is non zero. Then the associated data set O^K contains at least one cycle. To show this, we use the contrapositive. When there are no revealed demand cycles $C^{D,T}$ contained in O^K , then O^K satisfies SARP. By Chiappori and Rochet (1987) and Lee and Wong (2005)¹⁵, we conclude that there is at least one extension $x \in \mathcal{X}^K$ that can be generated by utility maximization and thus has a Slutsky matrix function with property σ , π and ν , a contradiction.

For the converse, if we observe a data set of $K \geq 2$ observations we define the set of walrasian demands passing through each of the observations, a subset of the set $\mathcal{X}(\mathcal{Z})$:

$$\mathcal{X}^K(\mathcal{Z}) = \{x \in \mathcal{X}(\mathcal{Z}) \mid x^k = x(p^k, w^k) \quad \forall k \in \{1, \dots, K\}\}$$

We argue by contradiction. Assume that the data set contains a revealed demand cycle, yet at least one $x \in \mathcal{X}_L(\mathcal{Z})$ has a Slutsky matrix that satisfies σ , π and ν . Then, pick any of such $x \in \mathcal{X}_L(\mathcal{Z})$ with Slutsky matrix satisfying σ , π and ν . Then, by Hurwicz and Uzawa (1971), we know there exists a locally nonsatiated continuous utility function defined on the region $x(\mathcal{Z}) \subseteq X, u : X(\mathcal{Z}) \mapsto \mathbb{R}$ such that $x(z)$ for $z \in \mathcal{Z}$, $x \in \mathcal{X}_L$ is the unique maximizer of u subject to the budget constraint $\{px(p, w) \leq w \mid (p, w) \in \mathcal{Z} \subseteq P \times W\}$. By Kim and Richter (1986) and Mas-Colell (1974), this implies that there is a strictly concave locally non satiated utility function

¹⁵ Lee and Wong (2005) showed that SARP is necessary and sufficient to obtain a smooth demand rationalization of a finite data set.

such that $u(x^{t-1}) < u(x^t) + Du(x^t)[x^{t-1} - x^t]$. Since the demand $x \in \mathcal{X}_L(\mathcal{Z})$ is the maximizer of the utility function u , we have $Du(x^t) = \lambda^t p^t$ for $\lambda^t > 0$ by local non satiation for all $t \in \mathcal{T}$. This implies $u(x^{t-1}) < u(x^t) + \lambda^t p^t [x^{t-1} - x^t]$. Since we have a revealed demand cycle, we have that $p^t [x^{t-1} - x^t] < 0$ for all t , which implies $u(x^t) > u(x^{t-1})$ for $t \in \{1, \dots, T\}$. However, the presence of the cycle in the data set also implies that $u(x^T) = u(x^0)$, given that $x^T = x^0$, which is a contradiction. Therefore, every element $x \in \mathcal{X}_L(\mathcal{Z})$ must have a Slutsky matrix that violate either σ, π or ν . □

Proof of Lemma 2

Proof First, we establish (i).

Say we observe a finite data set O^K $K \geq 3$, and we find a strict revealed demand cycle $C^{D,T} = \{p^t, x^t\}_{t=0}^T$, for $T \geq 3$. Consider $\mathcal{X}^T \subseteq \mathcal{X}$ to be the subset of demand functions that are consistent with the observed decisions:

$$\mathcal{X}^T = \{x \in \mathcal{X}([0, b]) : x^t = x(p^t, p^t x^t) = x(\tau_t) \quad \forall t \in \{0, \dots, T\} \quad \tau_t \in [0, b]\}.$$

We know that the DGDF must be an element of \mathcal{X}^T by Walras’ law since $p^t x^t = w^t$, so we must prove that all elements of \mathcal{X}^T have at least one Ville Cycle $C^{V(S),b}$ for $S \geq 1$.

Assume, to the contrary, that there is at least one $x \in \mathcal{X}^T$ that satisfies the Ville axiom. In particular, this implies that such a function can generate the data set O^K and at the same time exhibit no Ville cycles.

The absence of Ville cycles $C^{V(S),b}$ for $S \geq 1$ implies by HR that there is an integrating factor $\lambda(x(\tau))$ that is positive for all $\tau \in [0, b]$ and a function $z : x([0, b]) \rightarrow \mathbb{R}$ such that $\frac{\partial z(\tau)}{\partial \tau} = \lambda(\tau) p(\tau)' \frac{\partial x(\tau)}{\partial \tau}$, with $\lambda(\tau) p(\tau) = \nabla_x z(x(\tau))$ and with the property that $x(0) = x^0, x(1) = x^1, \dots, x(T) = x^T$ and $x(0) = x(T)$ and similarly $p(\tau)$ goes through all observed prices. Let the distance between observations in the revealed demand cycle define the length of the intervals of the grid $\tau_{t+1} - \tau_t = \|x^{t+1} - x^t\|_{\mathbb{R}^L}$, fix $\tau_0 = 0$, and define recursively τ_t for $t \geq 1$ (i.e., define $b = \sum_{t=0}^{T-1} \|x^{t+1} - x^t\|_{\mathbb{R}^L}$).

Integrating along the path, we have $\int_0^{\tau_T} \frac{\partial z(\tau)}{\partial \tau} d\tau = \int_0^{\tau_T} \lambda(\tau) [p(\tau) \frac{\partial x(\tau)}{\partial \tau}] d\tau = z(x(T)) - z(x(0)) = 0$ by the observation that $x^T = x^0$. However, by linearity of the integral over an interval, we have

$$\int_0^{\tau_T} \frac{\partial z(\tau)}{\partial \tau} d\tau = \int_0^{\tau_1} \frac{\partial z(\tau)}{\partial \tau} d\tau + \int_{\tau_1}^{\tau_2} \frac{\partial z(\tau)}{\partial \tau} d\tau + \dots + \int_{\tau_{T-1}}^{\tau_T} \frac{\partial z(\tau)}{\partial \tau} d\tau = 0.$$

Recall that by definition, $\int_{\tau_{t-1}}^{\tau_t} \frac{\partial z(\tau)}{\partial \tau} d\tau = z(x^t) - z(x^{t-1})$. Now, since we have a strict cycle $p^t x^t \geq p^t x^{t-1}$ for all $t \in \{0, \dots, T\}$, and $p x^t > p^t x^{t-1}$ for some $t \in \{0, \dots, T\}$ and $T \geq 3$. Also, by the assumption that there are no cycles $C^{D,2}$ ($T = 2$) and no Ville cycles, then there is a quasiconcave utility function z (in fact, strictly quasiconcave since we ruled out revealed indifference) such that $p^t x^t \geq p^t x^{t-1} \iff z(x^t) - z(x^{t-1}) \geq 0$ which implies that $\sum_{t=0}^{T-1} \int_{\tau_t}^{\tau_{t+1}} \frac{\partial z(\tau)}{\partial \tau} d\tau = \sum_{t=0}^{T-1} [z(x^{t+1}) - z(x^t)] > 0$. This contradicts the result that $\int_0^{\tau_T} \frac{\partial z(\tau)}{\partial \tau} d\tau = z(x^T) - z(x^0) = 0$ since both quantities are numerically equivalent. Therefore, each $x \in \mathcal{X}^T$ violates the Ville

axiom so the DGDF must have at least one Ville cycle $C_S^{V,b}$ for $S \geq 1$ and some $b > 0$.

Second, we establish (ii).

If there are $C^{D,T}$ for $T = 2$, then by definition every $x \in \mathcal{X}^T$ fails Weak WARP, because every $x \in \mathcal{X}^T$ has to interpolate the data set that has cycles for $T = 2$. Then by Kihlstrom et al. (1976) we know that ν should fail for the given extension in \mathcal{Z} , since every extension $x \in \mathcal{X}^T$ fails Weak WARP. \square

Proof of Proposition 1

Proof Observe that the set of extensions $\mathcal{X}^K(\mathcal{Z}) = \{x \in \mathcal{X}(\mathcal{Z}) | x^k = x(p^k, w^k) \ \forall k \in \{1, \dots, K\}\}$ is compact and the objective function is continuous. Then, there is at least one solution to the optimization problem. By Proposition 0, if O^K is generated by a rational DGDF, then $E^{x^r} = 0$ for some rational demand $x^r \in \mathcal{X}^K(\mathcal{Z})$ and since $\|E^x\| \geq 0$ for all $x \in \mathcal{X}^K(\mathcal{Z})$, then $x^r = \operatorname{argmin}_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$. When O^K is not generated by a rational DGDF, then all $x \in \mathcal{X}^K(\mathcal{Z})$ have $\|E\| > 0$ and by construction the value $\alpha^{*K} > 0$ is the lower bound of the true value of the Slutsky error norm, still part of the set, by compactness. The “moreover” statement dealing with the decomposition of the norm in the different effects follows from Lemmas 1, and 2. \square

Proof of Proposition 2

Proof The first property follows from Hurwicz and Uzawa (1971) because \mathcal{Z} is assumed to be compact, Walras’ law holds and $E^{x^{K,*}} = 0$ which means that the Slutsky regularity conditions hold. Then it follows that $x^{K,*}$ can be generated by maximizing a twice continuously differentiable utility function subject to a linear constraint. By Kim and Richter (1986) and Mas-Colell (1974), this implies that there is a strictly concave locally non satiated utility function. The second property follows from Quah (2005), where it is proved that a demand that has a singular in prices Slutsky matrix that is NSD can be rationalized by preferences that are complete and nontransitive, but that are regular as defined by the author.¹⁶ The third property follows from the proof of Hurwicz and Uzawa (1971) and Jerison and Jerison (1992) that uses only symmetry of the Slutsky matrix, thus $\|E^{\sigma \cdot x^{K,*}}\|^2 = 0$ to construct a scalar function $\phi : \mathcal{Z} \mapsto \mathbb{R}$ that can generate the data by the differential equation $x(p, w) = -\nabla_p \phi(p, w) / \nabla_w \phi(p, w)$ (Roy’s Identity) such that Walras’ law holds. This means that while the demand system is integrable, in this case ϕ is not necessarily an indirect utility function and in fact its Hessian may not be NSD. \square

Proof of Proposition 3

Proof Following Aguiar and Serrano (2017), we denote by $a(x) = E$ the map that assigns to each demand function the smallest (in the Frobenius norm sense) matrix

¹⁶ The upper contour set of x , $B(x)$ must be closed, convex and fulfill two additional (mainly technical) conditions explained in Definition 2.1 in Quah (2005).

that would make its Slutsky matrix function inherit all the regularity properties. Now we prove that the Gateaux derivative for the map $a(x) = E$ exists.

Let $a : \mathcal{X}(\mathcal{Z}) \mapsto \mathcal{M}(\mathcal{Z})$, be such that $a(x) = E$ and $\mathcal{M}(\mathcal{Z}) = \mathcal{C}(\mathbb{R}^{L+1}, \mathbb{R}^L \times \mathbb{R}^L)$. We are going to prove that a is Gateaux differentiable.

Relying on Aguiar and Serrano (2017), we define $a = s - (\varrho_v \circ (\varrho_{\sigma,\pi} \circ s))$ where $s(x) = S$ assigns to a demand function its Slutsky matrix function, the projection $\varrho_{\sigma,\pi} : \mathcal{M}(\mathcal{Z}) \mapsto \mathcal{M}_{\sigma,\pi}(\mathcal{Z})$ is a linear projection that maps any matrix function into the set of symmetric and singular in p matrix functions, $\varrho_{\sigma,\pi}(S) = S^{\sigma,\pi}$, also $\varrho_v : \mathcal{M}_{\sigma,\pi}(\mathcal{Z}) \mapsto \mathcal{M}_v(\mathcal{Z})$ as the mapping that projects any Slutsky matrix into the space of negative semidefinite matrix functions $\varrho_v(S^{\sigma,\pi,v}) = S_-^{\sigma,\pi,v}$, Aguiar and Serrano (2017) proved that $\varrho_v(S^{\sigma,\pi}) = S^r$.

We first prove that s is Gateaux differentiable: $\frac{\partial s(x+tv)}{\partial t}|_{t=0} = s'_x(v) = D_p v + D_w[xv']$, along direction $v \in \bar{\mathcal{C}}^1(\mathcal{Z})$, the limit exists under the compactness of \mathcal{Z} and continuity assumptions on x and its derivatives, where $\bar{\mathcal{C}}^1(\mathcal{Z})$ is a compact set of continuously differentiable functions that contains $\mathcal{X}(\mathcal{Z})$.

The mapping $\varrho_{\sigma,\pi}$ by Aguiar and Serrano (2017) is linear, $\frac{\partial \varrho_{\sigma,\pi}(S+TV)}{\partial t}|_{t=0} = \varrho_{\sigma,\pi}(S)$.

The mapping ϱ_v by Aguiar and Serrano (2017) is a polar-projection, and by Higham (1986) (theorem 2.5) this projection is Gateaux differentiable $\frac{\partial \varrho_v(S+TV)}{\partial t}|_{t=0} = L_v(S, V)$ exists and is a linear operator in the space of matrix functions $V \in \mathcal{M}(\mathcal{Z})$.

By the Gateaux derivative chain rule $\frac{\partial a(x+tv)}{\partial t}|_{t=0} = a'(v)$ exists and is defined as $a'(v) = s'_x(v) - L_v(s(x), \varrho_{\sigma,\pi} \circ s'_x(v))$.

Notice that letting $\varrho = \varrho_v \circ \varrho_{\sigma,\pi}$ and ι the identity function in the space of matrix functions, we say that $\frac{\partial (\iota - \varrho)(S+TV)}{\partial t}|_{t=0} = L_r(S, V)$ it exists by the previous argument.

Then $a'(v) = L_r(s(x), s'_x(v))$ by the fact that $a'(v)$ is linear. We define $\kappa = \|a'\| = \max_{v \in \bar{\mathcal{C}}^1(\mathcal{Z})} \|a'(v)\|$ we know this bound exists because of the compactness of $\bar{\mathcal{C}}^1(\mathcal{Z})$ and continuity of a' (it is linear).

The intermediate value theorem for the Gateaux derivative between Banach spaces says that $\|a(x) - a(x^K)\| \leq \|a'\| \cdot \|x - x^K\|$.

Let $\|a(x) - a(x^K)\| = \eta$, then $\eta \leq \|a'\| \cdot \epsilon$, with

$$\|a'\| \leq \|\iota - \varrho\| \cdot \|s'\|.$$

We can get a sharp upper bound for the norm of the map a' , estimated, as follows. Notice that ϱ , being a projection map, has by definition $\|\varrho\| \leq 1$. Further, $\iota - \varrho$ is also a projection to the orthogonal space of the projection ϱ , and hence, $\|\iota - \varrho\| \leq 1$. Also, $\|s'_x\| = \frac{\|D_p v + D_w[xv']\|}{\|v\|} \leq \frac{\|D_p v\|}{\|v\|} + \frac{\|D_w[xv']\|}{\|v\|}$. With the norm of differential operators under the norm chosen, then $\frac{\|D_p v\|}{\|v\|} = 1$, $\frac{\|D_w[xv']\|}{\|v\|} \leq \kappa'$. Here, $\kappa' = \|x\| + \|D_w[x]\|$, because $\frac{\|D_w[xv']\|}{\|v\|} \leq \frac{\|D_w[v]\| \cdot \|x\|}{\|v\|} + \frac{\|D_w[x]\| \cdot \|v\|}{\|v\|} = \|x\| + \|D_w[x]\|$. □

Proof of Proposition 4

Proof The result follows from the same reasoning as Proposition 3. By the intermediate value theorem for the Gateaux derivative between Banach spaces, we have $\|a(x) - a(x^h)\| \leq \|a'\| \cdot \|x - x^h\|$ with $x^h \in \mathcal{X}_h^K(\mathcal{Z})$ and $x \in \mathcal{X}^K(\mathcal{Z})$. By the definition of

the Sieves approximating spaces, there exists a (projection) mapping $\phi_h : \mathcal{X}^K \rightarrow \mathcal{X}_h^K$ such that $\phi_h \circ x \in \mathcal{X}_h$ and such that $\phi_h \circ x \rightarrow x$, satisfying that $\|x - \phi_h \circ x\| \rightarrow 0$ when $h \rightarrow \infty$. We know also by Wong (1984) (Lemma 2.2) that there is a constant $C > 0$ independent of the data set O^K such that for any $x \in \mathcal{X}^K$ there exists a $x^h \in \mathcal{X}_h^K$ satisfying $\|x^h - x\| \leq C\|\phi_h \circ x - x\|$.¹⁷ This means in turn that $\|a(x) - a(x^h)\| \rightarrow 0$ as $h \rightarrow \infty$, because $\|a(x) - a(x^h)\| \leq \|a'\| \cdot C\|x - \phi_h \circ x\|$.

We have established so far that any $x \in \mathcal{X}^K$ can be approximated by a sequence with elements $x^h \in \mathcal{X}_h^K$. Now, we notice that a is continuous under the C^1 norm on \mathcal{X} (claim (3) in the ‘‘Appendix’’), and that means that $\|E^x\| : \mathcal{X}^K \mapsto \mathbb{R}_+$ is also a continuous mapping under these conditions, by Theorem 3.1 in Chen (2007)¹⁸ we conclude that $\alpha_h^* \rightarrow \alpha^*$ as $h \rightarrow \infty$, and by Wong (1984)¹⁹ (theorem 2.5) we also conclude that for any $x^* \in \operatorname{argmin}_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$ there is a sequence $x^{*,h} \rightarrow x^*$ as $h \rightarrow \infty$ where $x^{*,h} \in \operatorname{argmin}_{x \in \mathcal{X}_h^K(\mathcal{Z})} \|E^x\|^2$. \square

Continuity of the Slutsky error map

Claim 2 The map $s : \mathcal{X}(\mathcal{Z}) \mapsto \mathcal{M}(\mathcal{Z})$ defined as $s(x) = S$ is continuous.

Proof First, we will prove that $D_p : \mathcal{X}(\mathcal{Z}) \mapsto \mathcal{M}(\mathcal{Z})$ and $D_w : \mathcal{X}(\mathcal{Z}) \mapsto \mathcal{C}(\mathcal{Z}, \mathbb{R}^L)$ are not only closed linear operators, but are also continuous maps. In general, differential operators are closed but not continuous. However, in this specific domain, D_p, D_w are defined everywhere by assumption, additionally D_p and D_w are closed operators, and finally $\mathcal{X}(\mathcal{Z}), \mathcal{M}(\mathcal{Z})$ are Banach spaces with the norms $\|\cdot\|_{C1}$ and $\|\cdot\|$ respectively, and so is $\mathcal{C}(\mathcal{Z}, \mathbb{R}^L)$, the space of continuous functions $f : \mathcal{Z} \mapsto \mathbb{R}^L$ with supremum norm $\|\cdot\|_{\infty,L}$. Then, by the closed graph theorem, we can conclude that D_p and D_w are continuous maps.

Second, take a convergent sequence in $\mathcal{X}(\mathcal{Z}), \{x_n\}_{n \in \mathbb{N}} \rightarrow x$. To finish the proof we want to show that $\lim_{n \rightarrow \infty} s(x_n) = s(x)$. By continuity of D_p, D_w and by the properties of the limit of a product of vectors, it follows that $\lim_{n \rightarrow \infty} s(x_n) = \lim_{n \rightarrow \infty} D_p x_n + \lim_{n \rightarrow \infty} D_w x_n [\lim_{n \rightarrow \infty} x]^\prime = S$, where $s(x) = S$, thus proving continuity of s . \square

Claim 3 The map $a : \mathcal{X}(\mathcal{Z}) \mapsto \mathcal{M}(\mathcal{Z})$ defined elementwise as $a(x) = S - S'$ is continuous.

¹⁷ The environment in Wong (1984) is different from ours in that the author requires a convex and closed space of functions and we do not. But Lemma 2.2 does not use these properties. Note also that we consider only the point interpolation case.

¹⁸ Our results are nonstochastic, thus we only require conditions 3.1–3.4 in Chen (2007), namely identification, Sieves space convergence, continuity, and compact sieve space. The first condition holds vacuously because we always have $\|E^x\| < +\infty$. Continuity of $\|E^x\|$ holds also, and the conditions on Sieves space convergence and compactness of the sieves space hold by assumption.

¹⁹ The result in Wong (1984) (theorem 2.5) is proved in an environment with the space of functions that is convex, for example by requiring Lipschitz continuous differentiability in our demand functions \mathcal{X} , but this assumption is not used in the convergence result. Its only role is to establish uniqueness. As we see from the argument in this proof, we can approximate any $x \in \mathcal{X}$, including $x^* \in \operatorname{argmin}_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$ there is an approximation $x^h \in \mathcal{X}_h^K$, the fact that $x^h = x^{*,h} \in \operatorname{argmin}_{x \in \mathcal{X}_h^K(\mathcal{Z})} \|E^x\|^2$ is a consequence of continuity of $\|E^x\|$ and the convergence properties of the sieves space increasing sequence \mathcal{X}_h .

Proof The continuity of the map a follows directly from the continuity of the Slutsky map s and the continuity of the projections maps that generates S' . Formally, $a = \varrho \circ s$ where $\varrho = (\iota - \rho_\nu \circ \rho_\pi \circ \rho_\sigma)$ is a projection mapping, ι is the identity map in $\mathcal{M}(\mathcal{Z})$ and ρ_σ is the projection on the space of symmetric matrices, ρ_π is the projection in the space of symmetric matrices that are singular with prices in its eigen-space associated with the null eigen-value, and ρ_ν is the projection on the space of symmetric, singular in prices and negative semidefinite matrices. By Claim 2, we know that $s : \mathcal{X}(\mathcal{Z}) \mapsto \mathcal{M}(\mathcal{Z})$ is continuous. It remains to show that the projection maps are indeed continuous. For this we need that the range of the projection map is a closed subspace under the metric induced by the norm of $\mathcal{M}(\mathcal{Z})$. In Aguiar and Serrano (2017) this is proven to be the case for each projection, in fact each range is convex and closed. It follows that a is continuous. \square

Simulations: methodological details

We use corollary (1) we approximate the lower bound of $\|E^{sf}\|_W$ for the fixed sample O^K . The lower bound is $\alpha^K = \min_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|_W$ where $\mathcal{X}^K(\mathcal{Z})$ is the set of extensions of O^K that belongs to $\mathcal{X}(\mathcal{Z})$. In practice, we cannot optimize over the whole $\mathcal{X}^K(\mathcal{Z})$ numerically.

1. First, we recast the problem in terms of budget shares and we refer to the functional space of budget shares corresponding to $\mathcal{X}^K(\mathcal{Z})$ as $\mathcal{X}^{K,b}(\mathcal{Z})$.
2. We replace this space by the approximate sieve space $\mathcal{X}_h^{K,b}(\mathcal{Z})$ that corresponds to the set of extensions that are B-splines or piecewise polynomials of degree h . A multivariate B-spline is obtain by the tensor product of $L + 1$ univariate B-splines of degree h , with the same degree for all dimensions, and they are evaluated at the data set. We obtain the matrix B_l of dimension $K \times H$ given fixed knots and the given degree h for each good l .
 - (a) For this simulation we choose $h \in \{5, 7, 11\}$.
 - (b) The support of the univariate spline is computed using the empirical support of the logarithm of prices.
 - (c) The number of experiments is $K \in \{20, 50, 100\}$.
 - (d) We use the function “spline.des()” in the Software “R”, in the package “splines” by Douglas Bates and Bill Venables.
 - (e) The tensor product is computed using “tensor.prod.model.matrix” in the Software “R”, in the package “mgcv” by Simon N. Wood.
3. We can obtain the vector $b_l = \{b_l^{sf}(p^k, w^k)\}_{k=1}^K$, for any element of the set of extensions of the data set $\mathcal{X}_h^{K,b}(\mathcal{Z})$, by finding $a \in \mathbb{R}^H$ such that $Ba = b$ in the data points.
4. In addition, we need to compute the derivatives of the interpolated functions which can be automatized since there are recursive formulas to obtain the derivatives of an spline (Chambers 1992), which we exploit in the numerical implementation of our exercise. In the same spirit as the previous step, we can obtain the partial derivatives of any extension in $\mathcal{X}^{K,h}(\mathcal{Z})$ by the following automatic procedure. We choose a

variable with respect to which we want to differentiate, say $\ln(p_1)$, derive the univariate B-spline corresponding to $\ln(p_1)$ and obtain the tensor product with respect to the remaining B-splines (for $\ln(p_2), \ln(p_3), \ln(w)$) to obtain a matrix $B_{l, \ln(p_1)}$ of dimension $K \times H$,²⁰ the partial derivative of the budget share b_l with respect to $\ln(p_1)$ at the data points, or the vector $b_{l, p_1} = \left\{ \frac{\partial b_l(\ln(p), \ln(w))}{\partial \ln(p_1)} \Big|_{p=p^k, w=w^k} \right\}_{k=1}^K$ can be obtained by multiplying $B_{l, \ln(p_1)} a = b_{l, p_1}$ for the same vector of weights as before. Using this procedure for all partial derivatives, we can obtain the Slutsky matrix evaluated at the data points. We call it $S^b(p^k, w^k)$ to denote its dependence on a given extension in the restricted space.

5. We obtain the Slutsky error matrix $E^b(p^k, w^k)$ (applying the formula in Aguiar and Serrano (2017)) for each data point and we approximate the integral $\|E^{sf}\|_W$, by the numerical approximation $\|E^{bh}\|_W = \frac{1}{K} \sum_{k=1}^K \|E^b(p^k, w^k)\|_{\mathbb{M}, W}$ for any given $b^h \in \mathcal{X}_h^{K, b}(\mathcal{Z})$.²¹
6. We use a Spectral Projection Gradient (SPG) large scale optimization algorithm to solve this optimization problem. The reader can notice that the minimization is with respect to the weights with a fixed $\mathcal{X}_h^{K, b}(\mathcal{Z})$.
 - (a) We initialize the algorithm by finding a feasible point, that is we first find any extension $x^0 \in \mathcal{X}_h^{K, b}(\mathcal{Z})$, and then we use it as an starting point for finding the extension with the minimal Slutsky matrix norm.
 - (b) For this last point we provide more details. The problem of finding the Minimal SMN for a finite data set is possibly non-convex:

$$\alpha^{*K} = \min_{x \in X^K(\mathcal{Z})} \|E^x\|^2 \quad (SMN1).$$

The mapping E^x is not concave on the demand $x \in X^K(\mathcal{Z})$.

We note that the restriction that $x \in \mathcal{X}(\mathcal{Z})$ implies that $x \in \mathcal{X}(\mathcal{Z})$ and $h_k(x) = x(p^k, w^k) - x^k$, such that $h_k(x) = 0$ for all $k = 1, \dots, K$. Let $h(x) = (h_k(x))_{k=1}^K \in \mathbb{R}^K$. The euclidean norm in \mathbb{R}^K is denoted by $\|\cdot\|_{\mathbb{R}^K}$.

We can follow Bertsekas (1979) to pose the following penalized problem, for a given constant $c > 0$:

$$\alpha_c^{*K} = \min_{x \in X^K(\mathcal{Z})} \|E^x\|^2 + \frac{1}{2}c \|h(x)\|_{\mathbb{R}^K}^2 \quad (SMN2).$$

From Bertsekas (1979) we know that: (i) For any $c > 0$, if $x^* \in X(\mathcal{Z})$ solves SMN2 then it solves SMN1. (ii) For any $c > 0$, the Minimal SMN associated to SMN1 is equivalent to the minimal SMN associated with SMN2, i.e., $\alpha_c^{*K} = \alpha^{*K}$. Finally, (iii) for $c > 0$ sufficiently large, there exists a unique solution to SMN2. We have transformed the non-convex SMN1 problem into a strictly convex problem SMN2. For practical purposes $c > 0$ can be chosen as a very high positive number. For any $c > 0$, the decomposition of $\alpha_c^{*K} = \alpha_c^{\sigma *K} + \alpha_c^{\pi *K} + \alpha_c^{v *K}$ is going to be close to the

²⁰ H is determined by the degree of the polynomial h and the step length of the univariate sieve, we use a step of $3/4$.

²¹ Alternatively, we could integrate in the region using automatic integration taking advantage of the B-spline structure.

true decomposition $\|E^{x^r}\|^2 = \|E^{x^r,\sigma}\|^2 + \|E^{x^r,\pi}\|^2 + \|E^{x^r,\nu}\|^2$, for the DGDF x^r , when K is large as in Proposition 3. This is true, because the distance between the DGDF and the interpolator that solves SMN2 (and SMN1) converges to 0. In practice, the practitioner can pick a sufficiently large $c > 0$ and solve the problem above in terms of its b-Spline approximation.

Proof of lemma (3)

Proof Aguiar and Serrano (2017) prove that we can decompose any observed Slutsky matrix in two components $S = S^r + E$ locally. Then we can write $\gamma(\{v_t\}_{t=1}^T, S) = \gamma(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w}) + E(\bar{p}, \bar{w}))$ at the point (\bar{p}, \bar{w}) . Then we observe that the JJ index is linearly separable $\frac{1}{T} \sum_{t=0}^T v'_t S(\bar{p}, \bar{w}) [v_t - v_{t-1}] = \frac{1}{T} \sum_{t=0}^T v'_t [S^r(\bar{p}, \bar{w})][v_t - v_{t-1}] + \frac{1}{T} \sum_{t=0}^T v'_t [E(\bar{p}, \bar{w})][v_t - v_{t-1}]$ so we can write $\gamma(\{v_t\}_{t=1}^T, S(\bar{p}, \bar{w})) = \gamma(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w})) + \gamma(\{v_t\}_{t=1}^T, E(\bar{p}, \bar{w}))$. \square

Proof of lemma (4)

Proof Notice that by subadditivity of the supremum $sup_{\mathbf{v}} \gamma(\mathbf{v}, S(\bar{p}, \bar{w}) - S^r(\bar{p}, \bar{w})) \leq sup_{\mathbf{v}} \gamma(\mathbf{v}, S(\bar{p}, \bar{w})) - sup_{\mathbf{v}} \gamma(\mathbf{v}, S^r(\bar{p}, \bar{w}))$ and by the properties of the JJ index (Remark 2 in Jerison and Jerison 2012) we have $sup_{\mathbf{v}} \gamma(\mathbf{v}, S^r(\bar{p}, \bar{w})) = 0$ because S^{x^r} is symmetric and NSD. Then $sup_{\mathbf{v}} \gamma(\mathbf{v}, S(\bar{p}, \bar{w}) - S^r(\bar{p}, \bar{w})) \leq sup_{\mathbf{v}} \gamma(\mathbf{v}, S(\bar{p}, \bar{w}))$. Now, recall that $E(\bar{p}, \bar{w}) = S(\bar{p}, \bar{w}) - S^r(\bar{p}, \bar{w})$, and we want to prove that $sup_{\mathbf{v}} \gamma(\mathbf{v}, E(\bar{p}, \bar{w})) \geq sup_{\mathbf{v}} \gamma(\mathbf{v}, S(\bar{p}, \bar{w}))$. In fact, this inequality follows from the properties of the JJ index (Remark 2 in Jerison and Jerison 2012), because the following two conditions hold: (i) $E(\bar{p}, \bar{w}) - S(\bar{p}, \bar{w}) = -S^r(\bar{p}, \bar{w})$ is both symmetric and PSD. (ii) $E(\bar{p}, \bar{w})' \bar{p} = S(\bar{p}, \bar{w})' \bar{p} - S^r(\bar{p}, \bar{w})' \bar{p} = 0$ because $\bar{p}' S(\bar{p}, \bar{w}) = 0$ and $\bar{p}' S^r(\bar{p}, \bar{w}) = 0$ by Walras' law, and by construction respectively. Thus we conclude that $sup_{\mathbf{v}} \gamma(\mathbf{v}, S(\bar{p}, \bar{w}) - S^r(\bar{p}, \bar{w})) \geq sup_{\mathbf{v}} \gamma(\mathbf{v}, S(\bar{p}, \bar{w}))$. Because of the previous two inequalities we establish the result. \square

Proof of proposition (5)

Proof By Jerison and Jerison (2012) we know that $lim_{\epsilon_p \rightarrow 0} \epsilon_p^{-2} \mathcal{G}_T((\bar{p}, \bar{w}) + V_w^{T(\epsilon_p, \eta)}) = sup\{\gamma(\mathbf{v}, S(\bar{p}, \bar{w}))/\bar{w} : \{v^t\}_{t=0}^T \in V^{T(\epsilon_p)}\}$. By Lemma 3 we know that $\gamma(\mathbf{v}, S(\bar{p}, \bar{w})) - \gamma(\mathbf{v}, S^r(\bar{p}, \bar{w})) = \gamma(\mathbf{v}, E(\bar{p}, \bar{w}))$. Finally by Lemma 4 we know that $sup\{\gamma(\mathbf{v}, E(\bar{p}, \bar{w}))/\bar{w}, \mathbf{v} \in V^{T(\epsilon_p)}\} = sup\{\gamma(\mathbf{v}, S(\bar{p}, \bar{w}))/\bar{w} : \mathbf{v} \in V^{T(\epsilon_p)}\}$. Therefore, $lim_{\epsilon_p \rightarrow 0} \epsilon_p^{-2} \mathcal{G}_T((\bar{p}, \bar{w}) + V_w^{T(\epsilon_p, \eta)}) = sup\{\gamma(\mathbf{v}, E(\bar{p}, \bar{w}))/\bar{w} : \mathbf{v} \in V^{T(\epsilon_p)}\}$. \square

Proof of proposition (6)

Proof The proof uses the fact that the solution to the matrix nearness problem is $E = S^r - S$, as shown in Aguiar and Serrano (2017). We want to show that $a(x) = S^r - S$

is AN at $0 \in \mathcal{M}(Z)$. By Claim 3, found in the “Appendix”, $a : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ is continuous when $\mathcal{X}(Z)$ is endowed with the $\|\cdot\|_{C_1}$ norm. Additionally, recall that the set $\mathcal{X}(Z)$ is compact under our assumption. Then we conclude [applying Proposition 3.1 in Boualem and Brouzet (2012)] that a is (AN) everywhere, i.e., $a(x) = S^r - S = C$ has the AN property for all $C \in \mathcal{M}(Z)$. In particular, a is AN at $0 \in \mathcal{M}(Z)$. Moreover, it follows that $\lim_{\delta \rightarrow 0} (\delta) = 0$ (applying Proposition 2.6 in Boualem and Brouzet (2012)). \square

References

- Afriat SN (1973) On a system of inequalities in demand analysis: an extension of the classical method. *Int Econ Rev* 14(2):460–472
- Aguiar V, Serrano R (2016) Classifying bounded rationality in limited data sets: a Slutsky matrix approach. *Mimeo*
- Aguiar VH, Serrano R (2017) Slutsky matrix norms: the size, classification, and comparative statics of bounded rationality. *J Econ Theory* 172:163–201
- Ahn D, Choi S, Gale D, Kariv S (2014) Estimating ambiguity aversion in a portfolio choice experiment. *Quant Econ* 5(2):195–223
- Anderson RM (1986) Almost implies near. *Trans Am Math Soc* 296(1):229–237
- Andreoni J, Harbaugh W et al (2005) Power indices for revealed preference tests. Social Systems Research Institute, University of Wisconsin
- Antonelli GB (1951) Sulla teoria matematica della economia politica. *Giornale degli Economisti e Annali di Economia*
- Apesteguia J, Ballester MA (2015) A measure of rationality and welfare. *J Polit Econ* 65:135–150
- Balasko Y, Tvede M (2010) General equilibrium without utility functions: how far to go? *Econ Theory* 45(1–2):201–225
- Bertsekas D (1979) Convexification procedures and decomposition methods for nonconvex optimization problems. *J Optim Theory Appl* 29(2):169–197
- Boualem H, Brouzet R (2012) On what is the almost-near principle. *Am Math Mon* 119(5):381–397
- Chambers J (1992) Linear models. In: Chambers M, Hastie TJ (eds) *Wadsworth and Brooks, Cole*
- Chen X (2007) Chapter 76 large sample sieve estimation of semi-nonparametric models. In: Heckman JJ, Leamer EE (eds) *Handbook of econometrics, volume 6, Part B*. Elsevier, Amsterdam, pp 5549–5632
- Chiappori P-A, Rochet J-C (1987) Revealed preferences and differentiable demand. *Econometrica*, pp 687–691
- Dean M, Martin D (2015) Measuring rationality with the minimum cost of revealed preference violations. *Rev Econ Stat* 98:524–534
- Echenique F, Lee, Shum M (2011) The money pump as a measure of revealed preference violations. *J Polit Econ* 119(6):1201–1223
- Epstein LG, Yatchew AJ (1985) Non-parametric hypothesis testing procedures and applications to demand analysis. *J Econom* 30(1):149–169
- Gabaix X (2014) A sparsity-based model of bounded rationality. *Q J Econ* 129(4):1661–1710
- Graves LM (1956) *The theory of functions of real variables, volume 121*. New York
- Halevy Y, Persitz D, Zrill L (2014) Parametric recoverability of preferences. *Mimeo*, New York
- Higham NJ (1986) Computing the polar decomposition-with applications. *SIAM J Sci Stat Comput* 7(4):1160–1174
- Hosoya Y (2017) The relationship between revealed preference and the Slutsky matrix. *J Math Econ* 70:127–146
- Hurwicz L, Richter MK (1979) Ville axioms and consumer theory. *Econometrica* 47(3):603–619
- Hurwicz L, Uzawa (1971) On the integrability of demand functions. In: *Cap 6. Preferences, utility and demand*. Harcourt, New York
- Jerison D, Jerison M (1992) Approximately rational consumer demand and ville cycles. *J Econ Theory* 56(1):100–120
- Jerison D, Jerison M (2012) Real income growth and revealed preference inconsistency. *Manuscript*

- John R (1995) The weak axiom of revealed preference and homogeneity of demand functions. *Econ Lett* 47(1):11–16
- Kihlstrom R, Mas-Colell A, Sonnenschein H (1976) The demand theory of the weak axiom of revealed preference. *Econometrica* 44(5):971–978
- Kim T, Richter MK (1986) Nontransitive-nontotal consumer theory. *J Econ Theory* 38(2):324–363
- Lee PM, Wong K-C (2005) Revealed preference and differentiable demand. *Econ Theory* 25(4):855–870
- Mas-Colell A (1974) Continuous and smooth consumers: approximation theorems. *J Econ Theory* 8(3):305–336
- Matzkin RL, Richter MK (1991) Testing strictly concave rationality. *J Econ Theory* 53(2):287–303
- Nishimura H, Ok EA, Quah J (2013) A unified approach to revealed preference theory: the case of rational choice. Mimeo, New York
- Pan VY, Chen ZQ (1999) The complexity of the matrix eigenproblem. In: Proceedings of the thirty-first annual ACM symposium on Theory of computing. ACM, pp 507–516
- Quah JK-H (2005) Weak axiomatic demand theory. *Econ Theory* 29(3):677–699
- Richter MK (1996) Revealed preference theory. *Econometrica*, pp 635–645
- Shafer WJ (1977) Revealed preference cycles and the slutsky matrix. *J Econ Theory* 16(2):293–309
- Shiozawa K, et al (2015) Note on the goodness-of-fit measure for garp; np-hardness of minimum cost index. Graduate School of Economics and Osaka School of International Public Policy (OSIPP) Osaka University discussion papers in economics and business 15:1–6
- Slutsky E (1915) Sulla teoria del bilancio del consumatore. *Giornale degli economisti e rivista di statistica*, 1–26
- Toraichi K, Katagishi K, Sekita I, Mori R (1987) Computational complexity of spline interpolation. *Int J Syst Sci* 18(5):945–954
- Varian HR (1983) Non-parametric tests of consumer behaviour. *Rev Econ Stud* 50(1):99–110
- Varian HR (1990) Goodness-of-fit in optimizing models. *J Econom* 46(1–2):125–140
- Wong WH (1984) On constrained multivariate splines and their approximations. *Numer Math* 43(1):141–152

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.