

Slutsky Matrix Norms and Revealed Preference Tests of Consumer Behaviour

Victor H. Aguiar* and Roberto Serrano†

This version: January 2015

Abstract

Given any observed finite sequence of prices, wealth and demand choices, we characterize the relation between its underlying Slutsky matrix norm (SMN) and some popular discrete revealed preference (RP) measures of departures from rationality, such as the Afriat index. We show that testing rationality in the SMN approach with finite data is equivalent to testing it under the RP approach. We propose a way to “summarize” the departures from rationality in a systematic fashion in finite datasets. Finally, these ideas are extended to an observed demand with noise due to measurement error; we formulate an appropriate modification of the SMN approach in this case and derive closed-form asymptotic results under standard regularity conditions.

JEL classification numbers: C60, D10.

Keywords: consumer theory; rationality; Slutsky matrix function; revealed preference approach; bounded rationality.

1 Introduction

The generalized axiom of revealed preference (GARP) completely characterizes whether a finite set of prices and demand choices can be rationalized. However, this binary approach is unsatisfying in the sense that we do not know by how much one departs from rationality. This problem has been the motivation of an important body of literature that has focused on how to measure the size of the violations of rationality. The revealed preference approach is the preferred method to test the rationality hypothesis in consumer behavior due to its nonparametric nature. However, in the words of Varian (1983):

“...the revealed preference approach does have some drawbacks. In certain cases, the tests involved may be computationally infeasible for large data sets. Also, the techniques do not typically summarize the data in a useful way. Furthermore it may be rather difficult to incorporate stochastic considerations in a satisfactory manner.”

This cite serves as a strong motivation for the current work. Indeed, here we propose a methodology that may help overcome the last two of these drawbacks (i.e., a useful decomposition of bounded rationality, and a suitable extension to situations with noise). In a previous effort, Aguiar and Serrano (2014) answer the question of how far is a given behavior (demand function)

*Department of Economics, Brown University, victor_aguiar@brown.edu

†Department of Economics, Brown University, roberto_serrano@brown.edu

from rationality using the observed Slutsky matrix function distance to its closest rational Slutsky matrix function. We refer to this as the Slutsky matrix norm (SMN) approach, which provides a way to measure the “size” of the departures from rationality, whatever those might be. Moreover, it yields a closed-form solution when the demand function is observed and provides a useful classification of the types of violations to the classical axioms of revealed demand. This measure can be monetized and provides a way to measure the failures in comparative statics analysis (prediction) when we assume (possibly incorrectly) that the underlying behavior is rational, so it is a “positive” index of bounded rationality. In addition, the smooth nature of the approach is particularly suited to be transported to stochastic environments. Its main limitation, however, is that it requires functional data. In this paper, we attempt to provide a complementary approach to the RP methodology that stems from the tradition of Antonelli (1886), Slutsky (1915) and Hurwicz and Uzawa (1971) by using the Slutsky matrix norm approach to test the empirical implications of consumer behavior in a finite dataset environment.

In contrast, the “revealed preference approach” (RP) born in the seminal contributions of Afriat (1973) and Varian (1983), is specifically designed to deal with finite datasets. Some of the most used and accepted measures include indices such as the Money Pump, the H-M index and the Afriat efficiency index, this is due to its implementation simplicity for small datasets (very common) and its nonparametric nature. However, the Afriat index and the Money Pump indices acquire meaning only if one maintains the assumption that the true behavior is rational but consumer fails to optimize, which makes them “normative” indices. In addition, these measures can be modified in the presence of measurement error, but we are not aware of any result that has a closed form for the asymptotic behavior of their tests. Also, it is not clear how these measures could be transported to an infinite dataset framework.¹

Here we are interested in exploring the connections of these widely used discrete measures of departures from rationality (RP) and the Slutsky matrix norm approach (SMN). Clearly, there are tradeoffs; both measures offer advantages and disadvantages and here we are interested in their complementarities. Efforts like the current paper are meant to unify these approaches in order to improve our ability to understand how to quantify the violations of the classical assumption of rationality and how to classify them. In particular,

- (i) Is testing rationality in the SMN approach with finite data equivalent to testing it under the RP approach?
- (ii) How can we “summarize” the departures of rationality in a systematic fashion?
- (iii) Given that we observe a demand that has noise, what is the appropriate modification of the SMN approach in this case and what is the asymptotic behavior of this measure?

The aim of this paper is to provide an answer to these three questions, and in doing so, to take steps towards answering the fundamental question of what is the best way to measure and classify departures from rationality in consumer decision making settings.

Our main results show that testing rationality in the RP and SMN approaches with finite data is equivalent. Non rational behavior will be accounted for as a positive Slutsky norm for *any* extension of the dataset. Conversely, if a dataset can be rationalized in the sense of Afriat

¹This is a theoretical question of interest since this is needed to understand the connection of the Afriat approach with the Richter environment (and the α and β axioms) as Nishimura et al. (2013) makes clear in their recent work

(1973), there exists an extension thereof that offers a rationalization in the sense of Hurwicz and Uzawa (1971), and also in our Slutsky norm sense. This sheds light on the interesting related results of de Clippel and Rozen (2014) and how their insights can be transported to the standard consumer theory environment.² All this is shown by connecting the properties of revealed demand cycles to properties of classic axioms of consumer theory, such as the Ville axiom and the weak axiom of revealed preference (WARP). Furthermore, the decomposition of the Slutsky norm into the different violations of rationality, as proposed in Aguiar and Serrano (2014), is also shown to apply to finite datasets. We extend our coverage to demand correspondences, which founds our analysis of noisy data. For this case, a smooth test is proposed, whose exact closed-form asymptotic distribution is known, as a function of the estimator of demand (Montecarlo simulations for finite samples support our asymptotic results). Finally, our results shed light on the interesting experimental data of Ahn et al. (2014): the different kinds of violations of rationality appear to correlate with one another, and failures of symmetry (Ville axiom) are more prevalent than violations of the weak axiom.

This is the paper’s outline. Section 2 presents the primitives of our analysis and the model. Section 3 presents the main equivalence result between the RP and the SMN approaches. Section 4 expands the previous analysis and presents a convergence result between the Afriat efficiency index and a function of the Slutsky error matrix. Section 5 extends our previous analysis to nondeterministic environments, in particular, it provides new results regarding the stochastic variant of the SMN approach. Section 6 develops and application using experimental data and the tools developed in the paper. Finally, we present a brief literature review in Section 7 and conclude in Section 8.

2 The Model: Testing Rationality with Limited Datasets

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 that we observed being chosen by the consumer at the price-wealth pair $(p^k, w^k) \in \mathbb{R}_{++}^L \times \mathbb{R}_{++}$. We allow that the consumer may or may not be rational in her choices.

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. There might be, in addition, instances of measurement error, so that $x \in \mathbb{R}^{L \times K}$ is the result of a data generating process as described in section (5).

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. This demand is a deterministic choice rule that represents the behavior of a consumer at a 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 $\mathcal{X}(\mathcal{Z}) \subset \mathcal{C}^1(\mathcal{Z}, \mathbb{R}^L)$. Indeed, we also assume that this is a compact set.

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 $k \in \{1, \dots, K\}$.

²In fact the set of extensions in our work is the same as the set of predictions in theirs.

We say that $x \in \mathcal{X}$ is the “true” DGDF if for any dataset 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 $x \in \mathcal{X}$ 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 at present (it will be allowed in our stochastic extension in the sequel). In addition, we are limiting the first part of our analysis to cases where O^K is such that there is no $p^k x^k = p^m x^m$ for any $k, m \in \{1, \dots, K\}$.

We also assume that Walras’ law holds throughout, for all O^K such that $p^k x^k = w^k$ 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 choice that fulfills 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, that of the extensions of a finite dataset O^K . Every finite dataset 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 dataset. 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 dataset 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 set of extensions is nonempty for all O^K for K finite when the sampling scheme from \mathcal{Z} is done without repetition. The true DGDF of O^K , $x \in \mathcal{X}^K(\mathcal{Z})$ belongs to the set of extensions for all $K > 0$.

2.1 RP and SMN Approaches to Test Rationality

The RP (revealed preference) approach to testing and measuring rationality started with the ground-breaking work of Afriat (1973), which consists of two steps: expressing the utility maximization procedure subject to a budget constraint (a set of linear inequalities), and using quantifier reduction techniques to obtain acyclicity conditions over a dataset O^K . If a dataset O^K is compatible with the system of inequalities, we say that the dataset can be rationalized. Otherwise, the hypothesis of rational decision making is rejected for the given sample of observed choices. This is a binary procedure. However, when attempting to evaluate by how much a dataset O^K departs from being rationalizable, many paths can be followed. In fact, there is a large literature of measuring departures from rationality in finite datasets. We refer to all those that share the common feature that their measures are a function of the revealed demand cycles $C^{D,T}$ for $T \geq 2$, defined next, as RP measures of rationality.

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 $p^t[x^t - x^{t-1}] > 0$ for all $t \in \{1, \dots, T\}$.³

Thus, if we can find a revealed demand cycle $C^{D,T}$ for $T \geq 2$, then we have a violation of the strong axiom of revealed preference (SARP). Namely, we define the relation R as follows: $x^t R x^{t-1}$ whenever $p^t x^t > p^t x^{t-1}$, which is read as x^t being “directly and strictly 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

³We sometimes write $x(t) = x^t$ and $p(t) = p^t$ for all $t \in \{0, \dots, T\}$ and $T \in \mathbb{N}_+$. Recall also that we assume Walras’ law: $p^t x^t = w^t$ for all t .

sequence $x^t R x^1 R x^2 R \dots x^n R x^s$, which we read as x^t being “strictly revealed preferred” to x^s . It is clear that, in $C^{D,T}$, we have $x^T R x^{T-1} R x^{T-2} \dots x^1 R x^0$, 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 (i.e., SARP means that if $x^k \bar{R} x^m$ then it cannot be the case that $x^m R x^k$ for all $(p^k, x^k), (p^m, x^m) \in O^K$).

The most well-known measure in the RP approach is the Afriat index:

Definition 4. (Afriat’s index) The Afriat efficiency index of a dataset 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.

Afriat’s index is such that $\mathcal{G} \in [0, 1]$ and measures the “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.

A more recent measure of rationality is the Money Pump Index (henceforth, MPI). It was formulated by Echenique et al. (2011).

Definition 5. (MPI Index) The MPI is defined as

$$MPI = \max_{C^{D,T}} \{ \sum_{t=1}^T p^t [x^t - x^{t-1}] / \sum_{t=1}^T p^t x^t | (p^t, x^t) \in C^{D,T} \}.$$

The MPI also yields values in $[0, 1]$ and it builds on Afriat’s “inefficiencies” to estimate the money pump required by the consumer in order to avoid instances of irrationality.

The SMN (Slutsky matrix norm) approach, developed in Aguiar and Serrano (2014), consists of measuring 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). This makes the approach suited for studying the structure of bounded rationality models from a theoretical point of view. However, as will be detailed in the sequel, the approach is flexible enough to allow simple extensions to be applied in cases where we have access to a finite dataset.

Definition 6. (SMN approach) The measure of bounded rationality for a given Slutsky matrix function S is

$$d(S) = \min \{ \|E\| : S + E \text{ satisfies the Slutsky regularity conditions} \}.$$

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 7. (JJ-Index) The Jerison and Jerison index is defined as

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

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 $v^0 = v^T$.

This index provides a direct connection between the Slutsky matrix and the revealed preference cycles. It measures violations of the Slutsky properties at a price-wealth reference point and a neighborhood thereof.

2.2 Ville Cycles and Revealed Demand Cycles

One central question that we tackle is how the RP approach measures of departures from rationality fare in comparison with the SMN of a demand function. The key to do this is to draw a connection between the revealed demand cycles from finite data and 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) (Hurwicz and Richter (1979) or HR henceforth). To state this axiom, we need to define a real income path.

Definition 8. (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 $\frac{\partial w}{\partial \tau}(\tau) > \frac{\partial p}{\partial \tau}(\tau)'x^i(p(\tau), w(\tau))$. A Ville cycle is a rising real income path such that $(w(0), p(0)) = (w(b), p(b))$.

We define a Ville Cycle as follows:

Definition 9. (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 $x \in \mathcal{C}^S([0, b]; \mathbb{R}_+^L)$ 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 $p : [0, b] \rightarrow \mathbb{R}_{++}^L$ such that $p(\tau)x(\tau) = w(\tau)$.

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 this 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).*⁴ 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 when we assume that a DGDF exists (i.e., we rule out nondeterministic choice) and with the restrictions of the type we have made on O^K then Weak WARP is simplified. In

⁴The strict version of the weak axiom of revealed preference says that $p^t x^t \geq p^s x^s$ then $p^s x^t > p^s x^s$.

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 $(\bar{p}, \bar{w}) \neq (p, w)$.

The Wald Axiom works on a special kind of dataset where Walras' law is imposed upon the consumer. One can force Walras' law by forcing the experimental subject to choose from the budget hyperplane with the requirement that wealth remains fixed across trials. One example of an O^K to test the Wald axiom is such $p^k x^k = \bar{w}$ for all $k \in \{1, \dots, K\}$. We state the Wald Axiom as follows: $p^k x(p^l, \bar{w}) < \bar{w}$ implies $x(p^l, \bar{w}) > \bar{w}$.

3 Connecting the SMN and RP Approaches

We are ready to make the connection between the finite data case and the infinite data case. The integrability conditions provided by the Frobenius theorem are of a global nature in the sense that they work for any region of the price-wealth space $\mathcal{Z} \subseteq P \times W$. Aguiar and Serrano (2014) prove that with functional data (i.e., if we observe the actual DGDF), we can measure the size of bounded rationality by means of the norm of a perturbation that converts any behavioral Slutsky matrix into a rational Slutsky matrix. It follows that a global connection between the RP and the SMN approaches must rely on how to extend a dataset $O^K = \{p^k, w^k, x^k\}_{k=0}^K$ to the whole region \mathcal{Z} . Again we can allow that an element of the set of possible extensions of the choice function or walrasian demand goes through the data while fulfilling the minimal conditions of continuous differentiability and Walras' law. The set of extensions can be very large, in fact it will generally be infinite. So our first interest is to know if a dataset 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 dataset that violates SARP. The answer in the affirmative will be provided by our main result in this section.

Furthermore, since evidence shows that GARP is routinely violated by a significant fraction of subjects in experimental studies, there is an increasing need to classify experimental datasets about consumer behavior in a useful way in order to understand how rationality is being violated. The SMN approach provides a natural way to do this since it can be decomposed additively in terms of the intensity of violations of the Slutsky regularity conditions (i.e., symmetry, singularity in prices, and negative semidefiniteness). As pointed out in Aguiar and Serrano (2014),

- (i) If WARP holds then $\|E^\pi\| = 0$ and $\|E^\nu\| = 0$.
- (ii) If the Ville Axiom holds then $\|E^\sigma\| = 0$.
- (iii) If no money illusion (homogeneity of degree zero) holds then $\|E^\pi\| = 0$.
- (iv) If the Wald axiom holds then $\|E^\nu\| = 0$.

In this section we also establish that this decomposition remains meaningful in finite datasets. The approach we take is to establish the relationship between the presence of revealed demand cycles and the behavioral axioms on the data generating demand function (DGDF). As a consequence, we can establish the connection with the SMN for the true DGDF (i.e., through the norm $\|E\|$).

This is the main result of this section:

Theorem 1. *Testing rationality of a limited dataset O^K is equivalent in the SMN and the RP approaches.*

Moreover, in the decomposition of the Slutsky matrix norm for the true DGDF, $\|E\|^2 = \|E^\sigma\|^2 > 0$ if there are $C^{D,T}$ for $T \geq 3$ and no $C^{D,T}$ for $T = 2$; and $\|E\|^2 = \|E^\pi\|^2 + \|E^\nu\|^2 > 0$ if there are $C^{D,T}$ for $T = 2$ and no $C^{D,T}$ for $T \geq 3$.⁵

The structure of the proof is as follows. First, we connect revealed demand cycles with axioms of consumer theory. This is done in Lemma 1 and Proposition 1, both stated and proved below. After that connection is established, Proposition 2 links types of revealed demand cycles to properties of the Slutsky matrix norm. All these results, which are the building blocks for Theorem 1, are presented in the rest of the current section.

Indeed, our first lemma concerns Weak WARP:

Lemma 1. *A given dataset O^K with $K \geq 2$ has at least a revealed demand cycle of order $T = 2$ if and only if all elements $x \in \mathcal{X}^K(\mathcal{Z})$ of the set of extensions of the dataset violate Weak WARP.*

Proof. First, we prove that if we observe a revealed demand cycle of size 2, $C^{D,2}$, then we have a violation of Weak WARP in the data generating demand function. Indeed, suppose we can find a cycle $C^{D,2}$ in a dataset O^K for $K \geq 2$. Then we observe that $p^2x^2 > p^2x^1$ and $p^1x^1 > p^1x^0$. Since $x^0 = x^2$, we have $p^2x^2 > p^2x^1$ and $p^1x^1 > p^1x^2$ at the same time. Although we do not know the true data generating demand function, it must be an element of the set of extensions $\mathcal{X}^T = \{x \in \mathcal{X} | x^t = x(p^t, p^t x^t), \forall t \in \{0, 1, 2\}\}$. Then, it follows from Walras' law that $p^t x^t = p^t x(p^t, w^t) = w^t$ for $t = 0, 1, 2$. It is straightforward to see that we have $p^2x(p^1, w^1) < w^2$ and $p^1x(p^2, x^2) < w^1$, which is a violation of Weak WARP. Thus the DGDF violates weak WARP.

Conversely, if all elements $x \in \mathcal{X}^K(\mathcal{Z})$ of the set of extensions of the dataset violate the Weak WARP, then we can build a revealed demand cycle $C^{D,2}$. To prove this, take the contrapositive, if there is no $C^{D,2}$ in O^K then by Kihlstrom et al. (1976) there is a demand function x that generates the data and satisfies Weak WARP, thus implying that not all $x \in \mathcal{X}^K(\mathcal{Z})$ violate the aforementioned axiom. \square

Corollary 1. *A necessary and sufficient condition to observe a revealed demand cycle of order $T = 2$ in some O^K $K \geq 2$ is that the generating demand function (DGDF) $x \in \mathcal{X}$ violates Weak WARP.*

Proof. If there is $C^{D,2}$ then by the lemma (1) it follows that the true DGDF $x \in \mathcal{X}$ violates Weak WARP because it has to be an element of the set of extensions $\mathcal{X}^K(\mathcal{Z})$.

If the data generating demand function violates weak WARP, we can generate a set of observations $O^K = \{p^k, w^k, x(p^k, w^k)\}_{k=1}^K$ for $K \geq 2$, where $x^k = x(p^k, w^k)$. By the violation of weak WARP, there is a set O^K $K \geq 2$ such that if we have $p^k x^k > p^k x^r$ and $p^r x^k < p^r x^r$ simultaneously for some pair $(p^k, x^k), (p^r, x^r) \in O^K$, then we can build a revealed demand cycle $C^{D,2}$. First, note that we have ruled out the case $p^k x^k = p^r x^r$ in O^K (i.e., we assume that there exists a DGDF). Then we relabel the observation k as $t = 0$ and $t = 2$ and we let r be $t = 1$, hence building the cycle $C^{D,2} = \{p^t, x^t\}_{t=0}^2$ that is a revealed demand cycle of size $T = 2$: $p^r x^k < p^r x^r$ leads to $p^1 x^0 < p^1 x^1$, and $p^k x^k > p^k x^r$ leads to $p^2 x^2 > p^2 x^1$. \square

⁵If we observe also wealth, Walras' law can be relaxed. Then we can separate violations of WARP into violations of Homogeneity of degree zero from violations of the Wald Axiom, resulting in each matrix (E^σ , E^π , and E^ν) in the decomposition having potentially a nonzero norm when the other two have a zero norm.

Next, Proposition 1 asserts that a necessary and sufficient condition to observe a revealed demand cycle of size $T \geq 3$ is that the data generating demand function contains at least one Ville cycle $C_S^{V,b}$ $S \geq 1$.

Proposition 1. *A given dataset O^K with no cycle of order $T = 2$ with $K \geq 3$ has at least a revealed demand cycle $C^{D,T}$ of order $T \geq 3$ only if all elements $x \in \mathcal{X}^K(\mathcal{Z})$ of the set of extensions of the dataset have a Ville Cycle $C^{V(S),b}$ for $S \geq 1$ and $b > 0$. Furthermore, some $x \in \mathcal{X}^K(\mathcal{Z})$ satisfy Weak WARP.*

Let the DGDF satisfy weak WARP. Then, a sufficient condition to observe a revealed demand cycle of order $T \geq 3$ in some O^K of size $K \geq 3$ is that the generating demand function (DGDF) $x \in \mathcal{X}$ has a Ville Cycle $C^{V(S),b}$ for $S \geq 1$ and $b > 0$.

Proof. First, we prove necessity. Say we observe a finite dataset O^K $K \geq 3$, and we find a 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 dataset 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 $\tau_{t+1} - \tau_t = \|x^{t+1} - x^t\|_{\mathbb{R}^L}$ and 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 $p^t x^t > p^t x^{t-1}$ for all $t \in \{0, \dots, T\}$ and $T \geq 3$ and by assumption there are no cycles $C^{D,2}$ ($T = 2$). Then there is a concave utility function z (in fact, strictly concave since we ruled out revealed indifference) such that $p^t x^t > p^t x^{t-1} \iff z(x^t) - z(x^{t-1}) > 0$ 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 $b > 0$. That at least an extension x satisfies Weak WARP follows from Lemma 1, given the absence of demand cycles of size 2.

If all elements $x \in \mathcal{X}^K$ of the set of extensions violate the Ville Axiom and there is a $x^* \in \mathcal{X}^K$ that satisfies Weak WARP, then there is at least a $C^{D,T}$ for $T \geq 3$ and no $C^{D,T}$ with $T = 2$. The latter follows from Lemma 1. To prove the former, we argue by contradiction. That is, suppose

there is no $C^{D,T}$ for $T \geq 2$. Then we conclude by Chiappori & Rochet (1987) that if there is no cycle then there is at least one $x \in \mathcal{X}^T$ that satisfies the Ville Axiom and Weak WARP. \square

A related result follows, linking the previous proposition (Proposition 1) to the DGDF.

Corollary 2. *Let the DGDF satisfy weak WARP. Then, a necessary and sufficient condition to observe a revealed demand cycle of order $T \geq 3$ in some O^K of size $K \geq 3$ is that the generating demand function (DGDF) $x \in \mathcal{X}$ has a Ville Cycle $C^{V(S),b}$ for $S \geq 1$ and $b > 0$.*

Proof. First, we prove sufficiency. If the DGDF has a Ville Cycle $C^{V(S),b}$ for $S \geq 1$, then we have $p(\tau) \frac{\partial x(\tau)}{\partial \tau} > 0$ for $\tau \in [0, b]$ almost everywhere. Partition the interval $[0, b]$ into $T+1$ equally spaced knots such that $0 = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_{T-1} < \tau_T = b$. Denote a sequence of wealth levels by $w(\tau_t) = w^t$ and a corresponding sequence of prices $p(\tau_t) = p^t$, with $\tau_t = \tau_{t-1} + h$ for $h > 0$. By the existence of the Ville cycle, $(p(0), w(0)) = (p^T, w^T)$ and $(p(b), w(b)) = (p^0, w^0)$ with $(p^T, w^T) = (p^0, w^0)$.

Take the first order approximation of the derivative $\frac{\partial x(\tau)}{\partial \tau} \Big|_{\tau=\tau_T} = \frac{[x^T - x^{T-1}]}{h} + \mathbf{1} \cdot o(h)$ for $\mathbf{1} = (1, \dots, 1)' \in \mathbb{R}^L$.

Then we have $p^{T'} \frac{[x^T - x^{T-1}]}{h} + o(h) > 0$, and we can do the same for τ_{T-s} for $s \in \{1, \dots, T\}$ and multiply by $h > 0$ so that

$$p^t x^t + o(h^2) > p^t x^{t-1} \text{ for } t \in \{1, \dots, T\}.$$

Since we fix h , we can make $o(h^2)$ to be small enough so that we have:

$$p^t x^t > p^t x^{t-1} \text{ for } t \in \{1, \dots, T\} \text{ which gives rise to the cycle, by defining}$$

$p^t x^t > p^t x^{t-1} \iff x^t R x^{t-1}$. Furthermore, by noticing that $x^T = x^0$ by assumption, we conclude that there is a revealed demand cycle of size T ($C^{D,T}$):

$x^T R x^{T-1} R x^{T-2} \dots x^2 R x^1 R x^T$, which means that the generated data set $\{p^t, x^t\}_{t=0}^T$ violates SARP. From Lemma 1, if the DGDF satisfies weak WARP then $T \geq 3$.

Sufficiency follows from Proposition 1. \square

The necessity part of this result uses a smooth version of Nishimura et al. (2013) decision making environments when we are focusing on a set of choice functions that are identified with a particular subset of the set of Walrasian demand functions (\mathcal{X}). It is also worth recalling that by HR we know that the absence of Ville cycles does not imply weak WARP nor does weak WARP imply the Ville axiom. But our result shows that when there is a Ville cycle, it can only guarantee cycles of size $T \geq 3$ if WARP holds and we can generate cycles of $T = 2$ when WARP does not hold.

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

Proposition 2. *Any given dataset $O^K = \{p^k, w^k, x^k\}_{k=1}^K$ 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 norm error that is non zero.*

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 dataset 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 & Rochet (1987) and Lee & Wong (2005), 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 SMN that is zero.

For the converse, if we observe a dataset of $K \geq 3$ observations we define the set of walrasian demands passing through each of the observations, a subset of the set $\mathcal{X}(\mathcal{Z})$:

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

This set is nonempty when the dataset contains no repeated points or when two observations do not have simultaneously $p^k x^k = p^m x^m$ when $w^k \neq w^m$, so there is no revealed indifference. We argue by contradiction. Assume that the dataset contains a revealed demand cycle, yet all $x \in \mathcal{X}_L(\mathcal{Z})$ have a Slutsky matrix norm $\|E\| = 0$. Then, pick any of the $x \in \mathcal{X}_L(\mathcal{Z})$ with $\|E\| = 0$. 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 $\{px(p, w) \leq w | (p, w) \in \mathcal{Z} \subseteq P \times W\}$. By Kim and Richter (1986) and Mas-Colell (1974), this implies that there is a strictly concave locally non satiated utility function 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 positive Slutsky matrix norm $\|E\| > 0$.

If we have a dataset with $K = 2$ observations, again we argue by contradiction. We have a revealed demand cycle of size $T = 2$, $C^{D,2}$ and $\|E\| = 0$, then $E^\nu = 0$ which implies that weak WARP must hold by Kihlstrom et al. (1976). This is a contradiction, since we have that $p^2 x^2 > p^2 x^1$ and $p^1 x^1 > p^1 x^0$ with $(p^0, x^0) = (p^2, x^2)$, implying a violation of Weak WARP. \square

Corollary 3. *A necessary and sufficient condition for some dataset $O^K = \{p^k, w^k, x^k\}_{k=1}^K$ for $K \geq 2$ observations to contain at least one revealed demand cycle is that the true DGDF x has an associated Slutsky matrix norm error that is non zero.*

Proof. Suppose a demand function $x \in \mathcal{X}(\mathcal{Z})$ has a Slutsky matrix function with an associated Slutsky matrix norm $\|E\| > 0$. Following Aguiar and Serrano (2014), the norm of E can be decomposed into the sum of three effects, corresponding to the violations of symmetry (E^σ), singularity (E^π), and negative semidefiniteness (E^ν). Then, it follows that at least one of the norms of E^σ, E^π, E^ν is not zero. If $\|E^\sigma\| > 0$, by Darboux theorem, we can build a Ville cycle (details can be found in HR). Then, we can discretize the cycle to obtain a revealed demand cycle by proposition 1. If $\|E^\pi\| > 0$ or $\|E^\nu\| > 0$, the Slutsky matrix is not negative semidefinite. Then, by Kihlstrom et al. (1976) this is equivalent to a violation of Weak WARP, which implies a revealed demand cycle of size $T = 2$.

The converse follows from Proposition 2. \square

Our argument in the previous proof 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 dataset that contains a revealed demand cycle.

This result implies that, when there is a dataset that can be rationalized in the sense of Afriat then we can find at least one demand function completion that can be rationalized in the

sense of Hurwicz and Uzawa (1971).

In fact, for this rational Walrasian demand function completion of the data, we know that there cannot be revealed demand cycles in the region $\mathcal{Z} \subseteq P \times W$ of any size.

Remark 1. This result uses 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 dataset that can be rationalized, there is at least one element $x \in \mathcal{X}^K(\mathcal{Z})$ that has $\|E\| = 0$. Therefore, we can identify such an element $x \in \mathcal{X}^K(\mathcal{Z})$ as the outcome of maximizing a unique concave utility function. Hence, we can think of an injection between the set of rational extensions $\mathcal{R}^K = \{x \in \mathcal{X}^K(\mathcal{Z}) \mid \|E\| = 0\}$ and $\mathcal{U}(\bar{\mathcal{Z}})$ where $\bar{\mathcal{Z}} = \cup_{x \in \mathcal{R}_L} x(\mathcal{Z})$ a set of utility functions defined on the union of ranges of all demands that can be rationalized. This connects the problem of completing a choice function with the problem of finding a multiple-utility representation when there is an observed incomplete order that is transitive.

The global connection between the revealed preference and the Slutsky norm approaches stated in Theorem 1 says that testing SARP and checking that $\|E\| = 0$ boils down to the same thing. In fact, suppose we design a binary test that gives 1 as its output if a dataset satisfies SARP and 0 otherwise, and another binary test that gives 1 as its output when at least one of the elements $x \in \mathcal{X}_L(\mathcal{Z})$ has a Slutsky matrix norm $\|E\| = 0$ and 0 otherwise. Then, both tests are equivalent for the same dataset, and an output of 1 signifies that the dataset O^K can be rationalized.

The actual way of testing rationality in this setting is captured by the following result.

Corollary 4. *A SMN test for rationality in limited datasets corresponds to the solution to the optimization problem $\gamma = \min_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$, where $\gamma^* = 0$ means that the dataset can be rationalized and otherwise it cannot, with $\gamma^* = \|E^\sigma\|^2 + \|E^\pi\|^2 + \|E^\nu\|^2 > 0$ the lower bound of the true value of $\|E\|^2$ for the underlying DGDF. Moreover, $\gamma^* = \|E^\sigma\|^2 > 0$ when there is a $C^{D,T}$ for $T \geq 3$ and no $C^{D,T}$ with $T = 2$, and $\gamma^* = \|E^\pi\|^2 + \|E^\nu\|^2 > 0$ when there is a $C^{D,T}$ for $T = 2$ and no $C^{D,T} \geq 3$.*

Proof. Observe that the set of extensions $\mathcal{X}^K(\mathcal{Z}) = \{x \in \mathcal{X}(\mathcal{Z}) \mid x^k = x(p^k, w^k) \quad \forall k \in \{1, \dots, K\}\}$ is compact and the objective function is continuous. Then, there is at least one solution to the optimization problem. By Theorem 1, 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 $\gamma^* > 0$ 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 Lemma 1 and Proposition 1. \square

A second conclusion of the proof of Proposition 2 is that there may be infinitely-many ways of extending a finite dataset. Therefore, the analyst needs a set of criteria to choose a reasonable measure of departures from rationality from the many that are available. This problem can be addressed transparently in the framework of the Slutsky matrix norm since all possible extensions of the data are members of the same set \mathcal{X} . Therefore we can state a set of axioms for the analysts that will serve to choose a reasonable measure of bounded rationality with finite data. In the case

of the previous corollary, we have chosen the measure that minimizes the measured violations of rationality. Note that for each infinite dataset (i.e., each demand function), there is only one Slutsky matrix norm associated with the observations.

Finally, with an observed revealed demand cycle, by Proposition 1 we know that there exists at least one $C^{V(S),b}$ Ville Cycle for $S \geq 1$ in the DGDF. Now, we show that we can construct a $C^{V(1),b}$ Ville cycle from a revealed demand cycle by means of an extension of the dataset using a quadratic spline. In fact, this is the shortest smooth Ville Cycle that does the job.

Corollary 5. *There is at least one $C^{V(1),b}$ cycle for a quadratic spline extension $x \in \mathcal{X}_L$ where O^K has a $C^{D,T}$ for $T \geq 3$.*

Proof. See the appendix (8) □

4 Further Specifics on the Uncovered Connections

This section details an approximate method to calculate the Slutsky matrix norm in finite datasets, relates the approach with the Varian supporting sets and with the work of Jerison-Jerison, and establishes a convergence result as the dataset size goes to infinity.

4.1 An Approximate Value of the Slutsky Matrix Norm

The previous results have 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. Clearly, the actual value of the decomposition may be also of interest, especially when this measure is monetized as in Aguiar and Serrano (2014). In this case, we provide an upper bound (possibly not the sharpest one) that helps to compute the error of the Slutsky matrix norm inherited from the error of the extension of a dataset. 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.

Following Aguiar and Serrano (2014), we denote by $a(x) = E$ the map that assigns to each demand function the smallest (in the Frobenius norm sense) matrix that would make its Slutsky matrix function inherit all the regularity properties. Also, the Gateaux derivative for the map $a(x) = E$, is a linear operator, such that $a' = a'(x) = \varrho \circ s'$ where $s'(x, v) = D_p v + D_w[xv']$, along direction $v \in \mathcal{C}^1(\mathcal{Z})$.

Proposition 3. *For every O^K , with x its true DGDF and $x^K \in \mathcal{X}^K(\mathcal{Z})$ with $\|x - x^K\| \leq \epsilon$. Then $\|a(x) - a(x^K)\| \leq \|a'\| \cdot \epsilon^K$.*

Proof. The intermediate value theorem for the Gateaux derivative between Banach spaces says that $\|a(x) - a(\hat{x})\| \leq \|a's\| \cdot \|x - \hat{x}\|$.

Let $\|a(x) - a(\hat{x})\| = \eta$, then $\eta \leq \|a'\| \cdot \epsilon$, with

$$\|a'\| \leq \|\varrho\| \cdot \|s'\|. \quad \square$$

Remark 2. We note that a sharp upper bound for the norm of the map a can be estimated. Notice that ϱ is a projection map. Then by definition $\|\varrho\| \leq 1$ and $\|s'\| = \frac{\|D_p v + D_w[xv']\|}{\|v\|} \leq \frac{\|D_p v\|}{\|v\|} + \frac{\|D_w[xv']\|}{\|v\|}$. With the norm of differential operators under the norm chosen, then $\frac{\|D_p v\|}{\|v\|} =$

1, $\frac{\|D_w[xv']\|}{\|v\|} \leq \kappa$. Where $\kappa = \|x\| + \|D_w[x]\|$, because $\frac{\|D_w[xv']\|}{\|v\|} \leq \frac{\|D_w[v]\| \cdot \|x\|}{\|v\|} + \frac{\|D_w[x]\| \cdot \|v\|}{\|v\|} = \|x\| + \|D_w[x]\|$. Observe that κ can be estimated using the interpolator \hat{x} .

With this result in hand, we can generically know the upper and lower bounds for any extension of a dataset O^K and consequently a workable upper bound for the Slutsky matrix norm. The idea here is to choose the fastest interpolator of a dataset O^K . It follows directly that if the interpolator is convergent when $K \rightarrow \infty$ then $\|a(x) - a(x^K)\| \rightarrow 0$.

4.2 The Jerison-Jerison Modified Slutsky Matrix Index, the Slutsky Matrix Norm, and the Afriat Index

An important question is how the Slutsky matrix norm relates numerically to Afriat's index. The connection is established through the Jerison-Jerison Slutsky index.

Jerison and Jerison (2012) establish a remarkable connection between Afriat's inefficiency cost measure of irrationality and a "Slutsky index" (for us JJ index). Both coincide in a local ball around some price and wealth reference point. This means that in principle one can unify two traditions of testing rationality and measuring the extent of the violations, namely the Afriat approach and the statistical approach of testing the Slutsky conditions.

Here we revisit this result and show that their measure can be seen as a quadratic form associated with the minimal correcting matrix that perturbs any observed Slutsky matrix and makes it rational (as proved in Jerison and Jerison (1992), and in Aguiar and Serrano (2014) for the general case). Our intent in this subsection is to establish the relation of the SMN norm and by extension of the work of Jerison and Jerison (2012) with the Afriat index.

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). This is indeed the case for the JJ index. We next prove a nice feature of the discrete JJ index that has not been previously pointed out.

Lemma 2. *The JJ index is equal to*

$$\iota(\{v_t\}_{t=1}^T, S(\bar{p}, \bar{w})) = \frac{1}{T} \sum_{t=0}^T v_t' S(\bar{p}, \bar{w}) [v_t - v_{t-1}] = \frac{1}{T} \sum_{t=0}^T v_t' E(\bar{p}, \bar{w}) [v_t - v_{t-1}],$$

where E is defined as the minimal perturbation matrix function that yields a rational Slutsky matrix function.

Proof. Aguiar and Serrano (2014) prove that we can decompose any observed Slutsky matrix in two components $S = S^r + E$. Then we can write $\iota(\{v_t\}_{t=1}^T, S) = \iota(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w}) + E(\bar{p}, \bar{w}))$ at point (\bar{p}, \bar{w}) . Then we observe that this index is linearly separable so we can write $\iota(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w}) + E(\bar{p}, \bar{w})) = \iota(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w})) + \iota(\{v_t\}_{t=1}^T, E(\bar{p}, \bar{w}))$. Finally, we apply the JJ index property that a rational Slutsky matrix has a zero index value. We conclude that $\iota(\{v_t\}_{t=1}^T, S(\bar{p}, \bar{w})) = \iota(\{v_t\}_{t=1}^T, E(\bar{p}, \bar{w}))$. \square

That is, this result shows that the JJ index is just a quadratic form associated with the correction matrix E . Then we can see more clearly that the JJ index is directly related to the Slutsky matrix norm approach. More importantly, this simple proof makes the connection between the perturbation matrix E and Afriat's cost inefficiency transparent.

Furthermore, it allows us to write a useful decomposition of the JJ index, exploiting the fact that the rational Slutsky matrix function is equal to $S^r = S + E$, where S is the observed Slutsky matrix and $E = E^\sigma + E^\pi + E^\nu$ is the smallest (under the Frobenius norm) correcting perturbation matrix function.⁶

Remark 3. The JJ index of any matrix $S(\bar{p}, w)$ can be decomposed as: $\iota(\{v_t\}_{t=1}^T, E(\bar{p}, \bar{w})) = \iota(\{v_t\}_{t=1}^T, E^\sigma(\bar{p}, \bar{w})) + \iota(\{v_t\}_{t=1}^T, E^\pi(\bar{p}, \bar{w})) + \iota(\{v_t\}_{t=1}^T, E^\nu(\bar{p}, \bar{w}))$. Moreover, each component is associated with cycles of different type and size. For instance, the $E^\sigma(\bar{p}, \bar{w})$ part is associated with cycles of size $T \geq 3$ while $E^{\pi, \nu}(\bar{p}, \bar{w})$ is associated with cycles of $T = 2$. Finally, if we observe also wealth, then we can separate violations of WARP in violations of Homogeneity of degree zero from the Wald Axiom.

4.3 Convergence Results

We now look at the case of continuous time or infinite datasets. We begin with a local result, and then we 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 $v^t \in V \subset \mathbb{R}^L$, V is a compact set, $v^0 = v^T$ 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 are going to 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 $\mathcal{G}((\bar{p}, \bar{w}) + V_w^{T(\epsilon_p, \eta)}) = \sup_{\{v^t, u^t\}_{t=0}^T \in V^{T(\epsilon_p, \eta)}} \{\min_{t \in \{0, \dots, T\}} \{p^t [x^t - x^{t-1}] / [p^t x^t]\} | (\bar{p} + v^t, \bar{w} + u^t x^t) = C^{D, T}(v^t, u^t) \in \{v^t, u^t\}_{t=0}^T\}$ (i.e., the supremum over the Afriat inefficiency index if we are allowed to sample price and wealth at will).

Proposition 4. *Given a set of cycles $V^{T(\epsilon_p)}$ and $V_w^{T(\epsilon_p, \eta)}$ then for a fixed $T \geq 1$*

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

Proof. It follows trivially from Jerison and Jerison (2012), due to lemma (2). \square

The 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 dataset O^K so the sampling process is given. Hence, this 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 datasets. 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 $|w^t - \bar{w}| \leq \epsilon_p \eta$.

Another popular index of bounded rationality in finite datasets environments is the Money Pump Index (MPI). Although we spare the reader with details here (included in a section of

⁶As detailed in Aguiar and Serrano (2014), this perturbation matrix can be written additively since we can define recursively the corrected Slutsky matrices with symmetry $S^\sigma = S + E^\sigma$, symmetry and singularity in prices $S^{\sigma, \pi} = S^\sigma + E^\pi$ and all the Slutsky regularity conditions $S^r = S^{\sigma, \pi, \nu} = S^{\sigma, \pi} + E^\nu$ where we add to the previous requirements, negative semidefiniteness of the Slutsky matrix.

the appendix – 8), it can be shown that locally both the MPI and the JJ index coincide. Also, by Lemma 2, we immediately get a connection between the MPI index and the Slutsky Matrix norm.

We close the subsection by going back to a global approach. Observe 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. Observe that if the sampling scheme is done over some region of interest \mathcal{Z} with a probabilistic measure with $\mu(E) > 0$ for all $E \in \mathfrak{B}(\mathcal{Z})$ the parts of the sigma-algebra of \mathcal{Z} , then we have that when $K \rightarrow \infty$ then since $x^* \in \mathcal{X}^K(\mathcal{Z})$ then $x^* \in \bigcap_K \mathcal{X}^K$. Finally, if $x \in \bigcap_K \mathcal{X}^K$ then $x \in \mathcal{X}^K$ for all K , and by the sampling scheme it means that $x^*(p, w) = x$ for all $(p, w, x) \in O^K$ for some K . This implies that any two $x, y \in \bigcap_K \mathcal{X}^K$ are essentially equivalent: we say that $x^* \equiv \lim_{K \rightarrow \infty} \mathcal{X}^K$ is the unique limit of the sequence of extensions.

In sum, we have argued that:

Lemma 3. $\lim_{K \rightarrow \infty} \mathcal{X}^K \rightarrow x^*$, the true DGDF.

For each set there is a bracket of SMN values formed by $\gamma_{min}^K = \min_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$ and $\gamma_{max}^K = \max_{x \in \mathcal{X}^K(\mathcal{Z})} \|E^x\|^2$. By the observation that $\{\mathcal{X}^K\}$ is a decreasing sequence of sets, we conclude that $[\gamma_{min}^K, \gamma_{max}^K] \supseteq [\gamma_{min}^{K+1}, \gamma_{max}^{K+1}]$, and by the previous lemma, $\lim_{K \rightarrow \infty} [\gamma_{min}^K, \gamma_{max}^K] = \gamma^*$, the true measure of bounded rationality.

5 The Size of Bounded Rationality with Limited Datasets and Noise

The extension of the RP approach to datasets that allow for indifferences (with data generating demand correspondences –DGDC) instead of functions is straightforward. However, the theory does not extend well when there is measurement error in the observations. Even in the simplest applications, the asymptotic theory is not known and practitioners use simulation techniques under the assumption of asymptotic consistency of tests of rationality. Partial identification is necessary in any case. The main reason is that the RP approach is a system of inequalities that is not easily portable when dealing with stochastic environments.

In this section we propose a smoother approach to this problem. It is an extension of the SMN approach for the deterministic case proposed in Aguiar and Serrano (2014). Let \mathbf{F} be a closed and bounded set of demand correspondences.

Definition 10. (Data Generating Demand Correspondence -DGDC-). A demand correspondence $\mathfrak{X} \in \mathbf{F}$ is a DGDC of $O^K = \{p^k, w^k, x^k\}_{k=1}^K$ if $x^k \in \mathfrak{X}(p^k, w^k)$ for all $k \in \{1, \dots, K\}$.

We say that $\mathfrak{X} \in \mathbf{F}$ is the true DGDC if for any dataset 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 $\mathfrak{X} \in \mathbf{F}$ is the Walrasian demand $\mathfrak{X} : \mathcal{Z} \rightarrow 2^{\mathbb{R}_+^L} / \emptyset$ such that $x^k \in \mathfrak{X}(p^k, w^k)$ for all $k \in \{1, \dots, K\}$ for all O^K . A true DGDC always exists. Of course, we have to impose some restrictions to have a workable method of measuring rationality. We assume henceforth, that the DGDC is upper hemicontinuous, convex-valued, compact-valued, and fulfills Walras’ law. Thanks to Kim and Richter (1986) and Mas-Colell (1974), this implies that $\mathfrak{X} \in \mathbf{X}$ can be generated by maximizing a strict order $\succ \in \mathbb{R}_+^L \times \mathbb{R}_+^L$ that is convex, monotone, and upper hemicontinuous. In practice,

this restriction is very weak in the case of finite data (i.e., to reject the assumption that \mathfrak{X} is compact-valued will require to observe all possible choices).

5.1 A Smooth Test of Rationality for Correspondences

Mas-Colell (1974) and Grodal (1974) prove the density of smooth rational demand functions in the set of rational demand correspondences, and this result is relevant for the class considered here.⁷ Thus, we know that rational demand correspondences admit a smooth rational demand function approximation. Formally, a demand correspondence can be generated by a convex and monotone rational preference relation if and only if for every $\eta > 0$ there exists a smooth demand function that can be rationalized by a strictly convex, strictly monotone and smooth preference relation such that $d_{\mathbf{X}}(x^r, \mathfrak{X}) < \eta$. The norm used in Mas-Colell (1974) is the Hausdorff distance defined on the set of graphs of the correspondences ($d_{\mathbf{X}} : \mathbf{X} \times \mathbf{X} \rightarrow \mathbb{R}_+$). Moreover, when dealing with the subset of demand functions, this metric is equivalent to the sup metric with associated norm $\|\cdot\|_{\infty}$.⁸

It turns out that the density of smooth demands in the set of the demand correspondences that we consider is preserved in the bounded rationality case. In fact, we first point out that we can approximate $\mathfrak{X} \in \mathbf{X}$ with a smooth demand function $x \in \mathcal{X} \cap \mathcal{C}^{\infty}$ arbitrarily closely.

Claim 1. For every $\eta > 0$ and $\mathfrak{X} \in \mathbf{X}$ there exists a smooth demand function $x \in \mathcal{X} \cap \mathcal{C}^{\infty}$ such that $d_{\mathbf{X}}(x, \mathfrak{X}) \leq \eta$.

Proof. First, we observe that:

(i) Thanks to De Blasi and Myjak (1986), we know that for every $v_1 > 0$ and $\mathfrak{X} \in \mathbf{X}$, there exists a continuous demand function x such that $d_{\mathbf{X}}(x, \mathfrak{X}) < v_1$.

(ii) Thanks to Azagra and Cepedello(2004), we know that for every $v_2 > 0$ and $x \in \mathcal{X}(\mathcal{Z})$, there exists a $x^* \in \mathcal{X} \cap \mathcal{C}^{\infty}$ such that $\|x - x^*\|_{\infty} \leq v_2$.

Now, we observe that for $\mathcal{X}(\mathcal{Z}) \subset \mathbf{X}$, the metric $d_{\mathbf{X}}$ and the metric induced by $\|\cdot\|_{\infty}$ in $\mathcal{X}(\mathcal{Z})$ are equivalent. Then, for every $\eta > 0$ and $\mathfrak{X} \in \mathbf{X}$, there exists $x \in \mathcal{X}(\mathcal{Z})$ such that $d_{\mathbf{X}}(\mathfrak{X}, x) \leq \eta/2$ and $x^* \in \mathcal{X} \cap \mathcal{C}^{\infty}$ such that $d_{\mathbf{X}}(x^*, x) \leq \eta/2$. By the triangle inequality, we conclude that $d_{\mathbf{X}}(x^*, \mathfrak{X}) \leq d_{\mathbf{X}}(x^*, x) + d_{\mathbf{X}}(x, \mathfrak{X}) = \eta$. \square

As seen in the previous proof, this result follows directly from known results in the literature, but together with the surprising result of Mas-Colell (1974) and Grodal (1974), it provides the technical basis to formulate a smooth test of rationality when a dataset O^K has a DGDC.

In finite data, we define the empirical metric from the set of observations O^K to the graph of a demand correspondence, so we let X be the DGDC and observe that $Gr(\mathfrak{X}(Z^K)) = O^K$, so we define the distance between two correspondences evaluated at a finite domain $Z^K = \{p^k, w^k\}_{k=1}^K$, so the distance between the dataset O^K generated by X and the dataset generated by $\mathfrak{X} \in \mathbf{X}$ is $d_{\mathbf{X}}(\mathfrak{X}(Z^K), \mathfrak{X}(Z^K)) = d_H(O^K, Gr(\mathfrak{X}(Z^K)))$.

In particular, $d_{\mathbf{X}}(\mathfrak{X}(Z^K), y(Z^K)) = \max_{k \in \{1, \dots, K\}} \|x^k - y(p^k, w^k)\|_{\mathbb{R}} < \eta$ when $O^K = Gr(\mathfrak{X}(Z^K))$ admits a DGDF and $y \in \mathcal{X}$ is a demand function.

⁷Mas-Colell (1974) proves this under the closed convergence topology and using a norm defined on the set of graphs of the correspondences.

⁸In our case, since we are interested in a finite data environment, we note that the (empirical) graph of the DGDC is $O^K = \{z^k, x^k\}_{k=1}^K$, we let $z^k \in \mathcal{Z} = P \times W$ and $x^k \in A = \mathbb{R}_+^L$, with the norm $\rho((z^k, a^k), (z^m, a^m)) = \max\{d_Z(z^k, z^m), d_A(a^k, a^m)\}$ in $\mathcal{Z} \times A$, and the distance defined as the Hausdorff metric $d_{\mathbf{X}}(X, Y) = \max\{\sup_{a \in (X^{-1}(A) \times X(Z))} \inf_{b \in (Y^{-1}(A), Y(Z))} \rho(a, b); \sup_{b \in (Z, Y(Z))} \inf_{a \in (Z \times X(Z))} \rho(a, b)\}$.

Lemma 4. *A dataset O^K has a rational DGDC $\mathfrak{X} \in \mathbf{X}$ if and only if for every $\eta > 0$ there is a smooth demand $x \in \mathcal{X} \cap C^\infty$ such that $d_{\mathbf{X}}(\mathfrak{X}(Z^K), x(Z^K)) < \eta$ and x can be rationalized.*

Proof. If a dataset O^K has a rational DGDC, then as a corollary from the Mas-Colell (1974) and Grodal (1974) theorem stated above, there is at least one smooth demand x whose $d_{\mathbf{X}}(\mathfrak{X}, x) < \eta$. So, by definition, $d_{\mathbf{X}}(\mathfrak{X}(Z^K), x(Z^K)) < \eta$ since $x^k \in \mathfrak{X}(p^k, w^k)$.

Suppose now that for every $\eta > 0$ there is a smooth demand $x \in \mathcal{X} \cap C^\infty$ such that $d_{\mathbf{X}}(\mathfrak{X}(Z^K), x(Z^K)) < \eta$ and that x can be rationalized. Then, for a fixed η , there is a set of smooth functions $\mathcal{X}_\eta = \{x \in \mathcal{X} \cap C^\infty \mid d_{\mathbf{X}}(\mathfrak{X}(Z^K), x(Z^K)) < \eta\}$. Let $\mathcal{R}_\eta \subseteq \mathcal{X}_\eta$ be the subset of demands that can be rationalized (this subset is such that its closure $cl(\mathcal{R}_\eta) \subseteq \mathbf{R}$, by Mas-Colell (1974) and Grodal (1974)). We have obtained a subset of the set of demand correspondences that can be rationalized by a monotone and convex preference relation. By definition, when $\eta \rightarrow 0$, there is an element $\mathfrak{X}_\eta \in cl(\mathcal{R}_\eta)$ such that $d_{\mathbf{X}}(\mathfrak{X}(Z^K), \mathfrak{X}_\eta(Z^K)) \rightarrow 0$. Then, $\mathfrak{X}_\eta \rightarrow \mathfrak{X}_0$ (in the closed convergence topology sense) and \mathfrak{X}_0 is a DGDC of O^K that can be rationalized. \square

Recalling the mapping $a(x) = E$ for $x \in \mathcal{X}(Z)$, we are now ready to state the main theorem of this subsection:

Theorem 2. *O^K can be rationalized if and only if for every $\eta > 0$, $\gamma_\eta^* = 0$, where $\gamma_\eta^* = \min_{x \in \mathcal{X}_\eta} \|a(x)\|^2$.*

Proof. We first establish that for every $\eta > 0$, $\gamma_\eta^* = 0$ if and only if there is an $x \in \mathcal{X}_\eta$ that can be rationalized. In fact, if x can be rationalized, then $a(x) = 0$ and since $\|a(x)\| \geq 0$ for all $x \in \mathcal{X}_\eta$, it follows that $\gamma_\eta^* = 0$.

Conversely, if $\gamma_\eta^* = 0$ then $a(x) = 0$. This implies that the Slutsky matrix of x fulfills the regularity conditions, and by Hurwicz and Uzawa (1971), we conclude that x can be rationalized. Then we use Lemma 4 to conclude that O^K can be rationalized if and only if for every $\eta > 0$, $\gamma_\eta^* = 0$. \square

The practical advantage of the test proposed here is that it is smooth and that it does not require to deal with correspondences. Moreover, there are also theoretical insights stemming from the previous results. In particular, the approximation results for demand correspondences using smooth demand functions provide a characterization of the role of revealed indifference in the RP approach. In fact, in the setting we consider, the DGDC's are of limited interest since they can be approximated arbitrarily closely by smooth DGDF's.

Indeed, for every $\eta > 0$, consider a dataset $O^{\mathfrak{X}, K} = \{p^k, w^k, \mathfrak{X}(p^k, w^k)\}_{k=1}^K$ generated by $\mathfrak{X} \in \mathbf{X}$. Then there is a dataset $O^{x, K} = \{p^k, w^k, x(p^k, w^k)\}_{k=1}^K$, where $x \in \mathcal{X} \cap C^\infty$ and $d_H(O^{\mathfrak{X}, K}, O^{x, K}) \leq \eta$. In this sense, all the datasets O^K that have a DGDC have a η -DGDF, that is, all datasets generated by sampling a DGDC have an arbitrarily close dataset generated by sampling a DGDF.

In sum, $O^{\mathfrak{X}, K} = \{p^k, w^k, \mathfrak{X}(p^k, w^k)\}_{k=1}^K$ almost admits a DGDF, or equivalently, it is almost deterministic. Moreover, the results of Mas-Colell (1974) and Grodal (1974) say that if we can find a rational demand correspondence that generates a dataset, there is an arbitrarily close smooth rational demand function to the rational DGDC. This feature will become critical in extending our theory to the noisy observations setting.

However, a major drawback in the theory just presented is that it does not provide a practical way of testing rationality, since the test is actually a family of tests indexed by $\eta > 0$. If a dataset fails any of them, it cannot be rationalized, but to ensure that the dataset can be rationalized, it must be the case that the test is passed for all $\eta > 0$. We will see in the sequel that this imposes no additional complications in the case where the dataset is measured with error.

5.2 A Smooth Test with Noisy Observations

Consider a sampling procedure of $\mathcal{Z} = P \times W$ and a measurement device of the choices made by the consumer in \mathbb{R}^L , which is subject to error. Notice we relax the positive restriction on consumption: negative consumption means short sales of a good. This is done in order to simplify the asymptotic theory (with some additional work, one can derive the asymptotic results for the case of choice in a truncated space). We assume that there exists a joint probability density function of choices and price-wealth situations: $\mu : \mathbb{R}^L \times \mathcal{Z} \rightarrow [0, 1]$. We also assume that there are two random variables that are square integrable, distributed according to μ , $(X, Z) \in \mathcal{L}^2$.

Definition 11. DGDP. We say that a random variable (X, Z) , or alternatively a probability distribution function μ , is a data generating demand process if $O^{*,K} = \{z^k, x^k\}_{k=1}^K$ is the result of stacking K independent draws from the pdf μ .

We will make some assumptions regarding the DGDP that allows us to link it to the definition of DGDF. The key assumption is that the observed choices that are realizations of the CDF associated with μ have the true DGDF as their mean.

Assumption 1. (A1) Mean-centered measurement error. $\mathbb{E}[X|Z] = x(Z)$ where $x \in \mathcal{X}(Z)$.

In particular, this assumption says that the random variable can be decomposed as $X = x(Z) + U$ with $Z \perp U$ so $x^k = x(z^k) + u^k$. In this sense, we can say that $O^{*,K}$ has an underlying DGDF since the observed choices are the sum of the images of a deterministic DGDF, evaluated at the realized values of prices and wealth $z \in \mathcal{Z}$ plus the realizations of the random variable U that correspond to the noise generated by the measuring device.

Remark 4. We do not consider the case where there is a DGDC $\mathfrak{X} \in \mathbf{X}$ such that $x^k = \bar{x}^k + u^k$ where $\bar{x}^k \in \mathfrak{X}(x^k)$ since in that case, given the results in the previous subsection, for every $\eta > 0$, there is a smooth DGDF such that $x^k \leq x(z^k) + \eta + u^k$. Since $\eta > 0$ can be made arbitrarily small and u^k can be arbitrarily large then $x^k \simeq x(z^k) + u^k$. Moreover, when dealing with finite data, it is impossible to distinguish between the case of a DGDC with noise and the smooth DGDF with noise, since we can as well write $x^k = x(z^k) + \bar{u}^k$ with $\bar{u}^k = \eta^k + u^k$ for some $\eta^k > 0$ for all $k \in \{1, \dots, K\}$ where $\|\bar{u}^k - u^k\|_\infty \leq \eta$. This says that, given a finite dataset $O^{*,K}$, in order to test rationality, in the presence of mean-centered measurement error, we can limit ourselves to test the existence of an underlying smooth DGDF.

The additional complexity of the dataset now considered, $O^{*,K}$, comes at a price. In particular, even if we observe all the realizations of the true DGDP μ we will not be able to say for sure that the dataset can be rationalized. Rather, we will be able to assert, μ almost surely, that the dataset has an underlying DGDF that can be rationalized. This is true for both the RP and the SMN approaches.

In finite datasets, however, we will not be able to make either “sure” or “almost sure” statements about the existence of a rational demand function. In fact, in finite data environments,

we have to make a compromise, that takes the form of a rule that rejects or not rejects the rationality hypothesis at a fixed confidence level. In other words, we are no longer able to decide whether a dataset $O^{*,K}$ can be rationalized or not but we can say that a α confidence level with a certain statistical power $(1 - \beta)$ we can reject the null hypothesis that the underlying DGDF $x \in \mathcal{X} \cap \mathcal{C}^\infty$ can be rationalized. This is the well known statistical hypothesis testing. In what follows, we are explicit about the connection of the statistical methodology to the notions of system decidability used in the RP approach.

An estimator of x given a dataset $O^{*,K}$ is called $x^K \in \mathcal{X}$ to denote its dependence on the number of draws.

Assumption 2. (A2) We assume that there exists a demand function $x^K \in \mathcal{X}$ that is a uniformly consistent estimator of x . Formally, $\|x^K - x\|_\infty \rightarrow 0$.

Let $\gamma^K = \|a(x^K)\|^2$. It is also relevant to describe the shape of the distribution of the hypothesis test of rationality.

Assumption 3. (A3) The regularity conditions in Newey (1994) (Linearization, Stochastic Equicontinuity and Mean-Square continuity) hold for $(\gamma^K)^{1/2}$.

Lemma 5. Under Assumptions A1-A3, and under the null hypothesis that $\gamma = 0$, then $\sqrt{K}\sqrt{\gamma^K} \rightarrow^d |\mathcal{N}(0, \tau^2)|$ a half-normal distribution as $K \rightarrow \infty$.

Proof. See Newey (1994) combined with the non-negativity of $\gamma \geq 0$. □

Equivalently, $\tau^{-2}K\gamma^K \rightarrow^d \chi^2(1)$ where τ^2 is the variance of the half-normal distribution from the previous lemma.

We arrive at the main result of this subsection:

Theorem 3. Testing rationality of a limited dataset $O^{*,K}$ with measurement error is equivalent in the SMN and the RP approaches as $K \rightarrow \infty$. Moreover, its asymptotic distribution is known under assumptions A1-A3. It is a Normal with mean zero and variance τ^2 , known up to a choice of an estimator of the DGDF.

Proof. The first part of the statement follows trivially from Proposition 5 below and Theorem (1). The second part follows from Lemma 5. □

Again, recall that $\|E\| = \|a(x)\|$ and the derivative of $\|a(x)\|^2$ in the direction of v evaluated at x is $L_x(v) = D_p v + D_w[xv']$ (Aguiar & Serrano (2014)).

Proposition 5. Suppose A1 and A2 hold. Then, a dataset $O^{*,K}$ can be rationalized almost surely if and only if, when $K \rightarrow \infty$, $\gamma^K \rightarrow^{\mathbf{p}} 0$, where \mathbf{p} denotes convergence with respect to the true probability measure over DGDP μ , and the estimator x^K converges uniformly to x .

Proof. First, observe that we know from Aguiar and Serrano (2014) that a is a closed map and we have computed its derivative that we called L_v . Second, A1 allows us to ensure the existence of a DGDF such that $\mathbb{E}[X|Z] = x(Z)$. If $O^{*,K}$ can be rationalized, or equivalently when there is an underlying DGDF $x \in \mathcal{X}$ that can be rationalized, by A2 we know that x^K is a uniformly consistent estimator of x and we can write $x^K = x + e^K$, where $e^K \in \mathcal{C}^1$ and $e^K \rightarrow 0$ uniformly. This allows us to bound $\|\gamma^K\| \leq \|L_v\| \cdot \|e^K\|$, where L_v is the Gateaux derivative of a in the

direction of $v \in \mathcal{C}^1$. Then, we conclude that, as $K \rightarrow \infty$, $\|\gamma^K\| \rightarrow^{\mathbf{P}} 0$ (in probability). Since $\|\gamma\| = \gamma$ is just the absolute value and $\gamma^K \geq 0$ for all K then this implies that $\gamma^K \rightarrow^{\mathbf{P}} 0$.

If when $K \rightarrow \infty$ it follows that $\gamma^K \rightarrow^{\mathbf{P}} 0$ then $\|a(x^K)\| \rightarrow^{\mathbf{P}} 0$. By the definition of the Frobenius norm, this implies that $a(x^K) \rightarrow^{\mathbf{P}} 0$. Now, using the closedness of a and Assumption A2, we conclude that since $x^K \rightarrow x$ uniformly then $a(x) = 0$ so x can be rationalized by Hurwicz and Uzawa (1971). Finally by Assumption A1, $O^{*,K}$ can be generated by drawing samples from Z evaluating in x and adding noise such that $\mathbb{E}[X|Z] = x(Z)$. \square

Remark. This way to test rationality has several advantages over the traditional RP approach, namely: (i) first the alternative hypothesis is being not rational, the typical approach in RP is to specify the alternative as the randomly uniform consumer of Becker. (ii) As we will see in a second the asymptotic power is one and in general the rate of convergence is similar or faster to a semiparametric estimator.

To help in the computation of the actual rationality test of the estimator in a finite dataset environment and appreciate the advantage of the Slutsky Matrix norm approach in the presence of measurement error, we establish both first and second order asymptotic analyses of the estimator that serves as a blueprint, using as sufficient conditions different convergence rates of the estimators of demand.

Lemma 6. *If $\|x^K - x\| = O(K^{-1/4})$ and $\mathbb{E}[\Sigma(X)|Z] = \Sigma(Z)$, and if Assumption A2 holds when $K \rightarrow \infty$, then $\sqrt{K}[\gamma^K - \gamma]/\tau \rightarrow^d \mathcal{N}(0, 1)$ as $K \rightarrow \infty$, where τ depends on Σ and on L_v . Moreover, when $\gamma = 0$, $\sqrt{K}[\gamma^K - \gamma] \rightarrow \delta_0$ a degenerate Dirac's delta distribution with all the mass concentrated on 0.*

Proof. First, take the functional Taylor expansion of the estimate x^K at the true x

$$\|a(x^K)\|^2 = \|a(x)\|^2 + 2\langle a(x), \varrho \circ s'(x^K - x) \rangle + O(\|x - x^K\|^2)$$

Then pre-multiply by \sqrt{K} to get

$$\sqrt{K}[\gamma^K - \gamma] = 2\sqrt{K}\langle a(x), \varrho \circ s'(x^K - x) \rangle + \sqrt{K}O(\|x - x^K\|^2).$$

By assumption, we have that $\|x^K - x\| = O(K^{-1/4})$, then $\sqrt{K}O(\|x - x^K\|^2) = O(1)$ and we conclude that: $\sqrt{K}[\gamma^K - \gamma] = 2\sqrt{K}\langle a(x), \varrho \circ s'(x^K - x) \rangle + O(1)$.

By Assumption A2, $\sqrt{K}\|x^K - x\| \rightarrow \mathcal{N}(0, \Lambda)$, by the functional delta method we observe that

$\sqrt{K}[\gamma^K - \gamma] \rightarrow^d \mathcal{N}(0, \tau^2)$, since $\sqrt{K}[\gamma^K - \gamma] = L_{x^K}(\sqrt{K}[x^K - x]) + O(1)$. The constant \sqrt{K} enters inside the argument of the derivative due to the bilinearity of the inner product $\sqrt{K}\langle a(x), \varrho \circ s'(x^K - x) \rangle = \langle a(x), \varrho \circ s'(\sqrt{K}[x^K - x]) \rangle$ and the linearity of ϱ and s' . Then τ depends solely on Σ and on L_v . The derivative L_v is a closed map since it is the composition of a closed map a and other continuous maps. Then, by Proposition 5, we know that $a(x^K) \rightarrow^{\mathbf{P}} a(x)$, and by assumption, $\sqrt{K}\|x^K - x\| \rightarrow^d \mathcal{N}(0, \Sigma)$. We conclude that $\sqrt{K}[\gamma^K - \gamma] \rightarrow^d \mathcal{N}(0, \tau^2)$ with $\tau^2(p, w)$ the variance of the random variable $L_x[\xi]$ where $\xi(p, w) \sim \mathcal{N}(0, \Sigma(p, w))$.

When $\gamma = 0$, then $\sqrt{K}\|a(x^K)\|^2 = \sqrt{K}\|a(x)\|^2 + 2\sqrt{K}\langle a(x), \varrho \circ s'(x^K - x) \rangle + \sqrt{K}O(\|x - x^K\|^2)$ simplifies to $\sqrt{K}\gamma^K \rightarrow^d O(1)$ under the assumption of $\|x^K - x\| = O(K^{-1/4})$. \square

As we can see from the previous result, the SMN approach is a semiparametric estimator and, as such, it shares the nice efficiency properties of estimators of its type. In particular, integrating in a region of prices and wealth averages out noise and allows for a faster convergence.

We see next how, under the null hypothesis of rationality, the linear order terms of the functional expansion of the estimator vanish, thus allowing for an even faster rate of convergence and a chi square like test that has an asymptotic power of one. This is in fact a second order asymptotic result that, as far as we know, is a novel rationality test of consumer behavior. We say that $\mathcal{D} \trianglelefteq \mathcal{F}$ if and only if the distribution of the random variable \mathcal{F} first order stochastically dominates \mathcal{D} .

Lemma 7. *If $\|x^K - x\| = O(K^{-1/3})$ and $\mathbb{E}[\Sigma_X|Z] = \rho(Z)I$, and if the regularity condition in Assumption A2 holds for $\|x^K - x\|^2$ when $K \rightarrow \infty$, then the rationality test statistic $\sigma^{-2}K[\gamma^K] \rightarrow^d \mathcal{D} \trianglelefteq \sigma^{-2}\|\sqrt{K}[x^K - x]\|^2$ with $\sigma^{-2}K\|x^K - x\|^2 \rightarrow \chi^2(L)$, where \mathcal{D} is given implicitly and $\sigma^2 = \|\varrho \circ s'\|^2 \cdot \rho^2$. Moreover, testing rationality with the SMN approach has an asymptotic power of one.*

Proof. First, take the functional Taylor expansion of the true DGDF at the estimate x^K :

$$\|a(x^K)\|^2 = \|a(x)\|^2 + 2\langle a(x), \varrho \circ s'(x^K - x) \rangle + \frac{1}{2}\langle 2\langle \varrho \circ s'(x^K - x), \varrho \circ s'(x^K - x) \rangle + \langle a(x), \varrho(D_w([x^K - x][x^K - x'])) \rangle + O(\|x^K - x\|^3)$$

Under the null hypothesis, $a(x) = 0$:

$$\|a(x^K)\|^2 = \langle \varrho \circ s'(x^K - x), \varrho \circ s'(x^K - x) \rangle + O(\|x^K - x\|^3).$$

There is a scale K and σ^2 , and under the assumption $O(\|x^K - x\|^3) = O(K^{-1})$

$$\sigma^{-2}K\|a(x^K)\|^2 = \sigma^{-2}K\langle \varrho \circ s'(x^K - x), \varrho \circ s'(x^K - x) \rangle + \sigma^{-2}KO(K^{-1});$$

$$\sigma^{-2}K\|a(x^K)\|^2 = \sigma^{-2}K\langle \varrho \circ s'(x^K - x), \varrho \circ s'(x^K - x) \rangle + O(1).$$

Notice that $\sigma^{-2}K\|a(x^K)\|^2 = B(\varrho \circ s'(x^K - x), \varrho \circ s'(x^K - x))$ with B the bilinear operator just defined, or abusing notation $\sigma^{-2}K\|a(x^K)\|^2 = B(\varrho \circ s'(x^K - x), \varrho \circ s'(x^K - x))$.

Continuity of the bilinear operator allows us to say

$$\|B \circ \varrho \circ s'(x^K - x)\|^2 \leq \|B\| \cdot \|\varrho \circ s'\|^2 \cdot \|(x^K - x)\|^2.$$

The norm of the bilinear operator is:

$$\|B\| = \sup_{\|V\| \leq 1} B(V, V) \leq 1 \text{ because } \langle V, V \rangle = \|V\|^2.$$

We conclude:

$$K\langle \varrho \circ s'(x^K - x), \varrho \circ s'(x^K - x) \rangle \leq \|\varrho \circ s'\|^2 \|\sqrt{K}[x^K - x]\|^2.$$

We let $\sigma^2 = \|\varrho \circ s'\|^2 \cdot \rho^2$ with ρ^2 the variance of associated $\sqrt{K}\|x^K - x\| \rightarrow |\mathcal{N}(0, \rho^2)|$ the half-normal distribution, due to Assumption A2, and the positivity constraint of the estimator. Thus, $\sigma^{-2}\|\sqrt{K}[x^K - x]\|^2 \rightarrow \chi^2(L)$.

Finally, the term $K\langle a(x), \varrho \circ s'(x^K - x) \rangle$ when $a(x) \neq 0$ or under the alternative hypothesis goes to infinity as $K \rightarrow \infty$, thus proving consistency of the test or asymptotic power of 1. \square

Since L_v is known analytically, we can compute ρ^2 when we have a specific estimator of x . This is a key advantage in the SMN approach with respect to the RP approach, since we can actually write a closed form for the asymptotic distribution of the size of bounded rationality.

Remark 5. The actual distribution depends on the choice of the estimator, but it can be readily estimated or simulated using the closed form

$K\|a(x^K)\|^2 = \langle \varrho \circ s'(\sqrt{K}[x^K - x]), \varrho \circ s'(\sqrt{K}[x^K - x]) \rangle + O(1)$ and the pointwise distribution of $x^K(z) - x$ for a given estimator. However, rationality can be rejected (conservatively) when the test statistic surpasses a critical value ($c_\alpha(\chi^2(L))$) given by the quantile $1 - \alpha$ of the $\chi^2(L)$ distribution where α is the chosen significance level. This provides a conservative test of rationality that holds for all estimators of demand that fulfill the conditions stated in the previous result. The ρ^2 can be estimated from the statistic γ^K and a polynomial estimator by using

the results in Akerberg et al. (2012), which permits to ignore the nonparametric nature of the first step and still obtain an asymptotically convergent estimator of the quantity. The rejection region is given by $\{\gamma^K : \sigma^{-2}K\gamma^K \geq c_\alpha(\chi^2(L))\}$. Finally, we can use the bootstrap procedure proposed in Haag et al. (2009), since our SMN is a generalization of their work.

We state an immediate, but quite interesting, corollary of the previous results.

Corollary 6. *When $K \rightarrow \infty$, the decomposition $\gamma^K \rightarrow_{\mathbf{P}} \|E^{K,\sigma}\|^2$ if and only if there are $C^{D,T}$ for $T \geq 3$ and $\gamma^K \rightarrow_{\mathbf{P}} \|E^\pi\|^2 + \|E^\nu\|^2$ if and only if there are $C^{D,T}$ for $T = 2$ in the deterministic dataset generated by $x \in \mathcal{X} \cap C^\infty$ with $E[X|Z] = x$.*

Proof. This follows trivially from Proposition 5 and Proposition 1. \square

5.3 Montecarlo Study of Finite Samples

In order to complement our first- and second-order asymptotic results for large enough datasets with mean-centered measurement errors, the current subsection provides some Montecarlo studies of fixed finite samples. The results are encouraging. They exhibit fast convergence, and for a given demand curve fit to a noisy dataset, the associated Slutsky matrix norm can be taken as a “good” asymptotic approximation.

We study a boundedly rational model proposed by Shaffer (1977):

- $x_1(p, w) = \frac{w(p_1 + (1-c)p_2)}{2p_1(p_1 + p_2)}$,
- $x_2(p, w) = \frac{w(p_1 + (1-c)p_2)}{2p_2(p_1 + p_2)}$,
- $x_3(p, w) = \frac{wcp_2}{p_3(p_1 + p_2)}$.

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 symmetric 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 (ν) fail.

We impose that $w = 1$, and consider the budget share version of the SMN, adding a zero-centered random measurement error. We have:

- (i) $b_1^*(p) = \frac{p_1 + (1-c)p_2}{2(p_1 + p_2)} + \epsilon_1$ with $\epsilon_1 \sim N(0, 0.13)$.
- (ii) $b_2^*(p) = \frac{p_1 + (1-c)p_2}{2(p_1 + p_2)} + \epsilon_2$ with $\epsilon_2 \sim N(0, 0.13)$.
- (iii) $b_3^*(p) = 1 - b_1^*(p) - b_2^*(p)$.

The effects of sample size and the size of bounded rationality on power: The number of simulations is fixed at 100, let $s \in \{1, \dots, 100\}$ and the dataset O_s^{*100} is generated with $K \in \{100, 500, 1000, 2000\}$ i.i.d. draws from $\ln(p_l) \sim N(\ln(2), 1)$ for $l \in \{1, 2, 3\}$ and from $\epsilon_l \sim N(0, .13)$ for $l \in \{1, 2\}$, such that we obtain: $O_s^{*K} = \{p^{s,k}, b^{s,*}(p^{s,k})\}_{k=1}^K$ for each value of K .

The key parameter c is fixed at three different values $c \in \{0, \frac{3}{2}, 3\}$, to generate different intensities of both types of violations (i.e., of σ and ν). We estimate the Budget Share Slutsky matrix, $\hat{S}^u(p, \bar{w}) = [\bar{D}_{\ln(p)} \hat{b}(p, \bar{w})] - [\bar{D}_{\ln(p)} \hat{b}(p, \bar{w}) \mathbf{1} \hat{b}(p, \bar{w})']$ where $\bar{D}_{\ln(p)} \hat{b}(p, \bar{w}) = D_{\ln(p)} \hat{b}(p, \bar{w}) -$

Null hypothesis	Sample	Proportion null is rejected.		
		$c = 0$ ($\ E\ _u^2 = 0$)	$c = 3/2$ ($\ E\ _u^2 = 0.15$)	$c = 3$ ($\ E\ _u^2 = 11.09$)
$H_0 : \ \hat{E}\ _u^2 = 0$	$K = 100$	0.01	0.11	0.88
	$K = 500$	0.00	0.53	1.00
	$K = 1000$	0.00	0.83	1.00
	$K = 2000$	0.00	0.97	1.00

Table 1: Montecarlo simulation for $n_s = 100$, proportion of cases where the null is rejected.

$Diag(\hat{b}(p, \bar{w}))$. Notice that $\hat{S}^u(p, \bar{w}) = \frac{1}{w} Diag(p)S(p, \bar{w})Diag(p)$ is the unit-free Slutsky matrix budget share from (Aguiar & Serrano, 2014). We use a non parametric kernel estimator of the conditional expectation of $b^*(p)$, $\hat{b}(p) = \mathbb{E}[b^*|P = p]$. The estimator is the Nadaraya-Watson estimator (Li & Racine (2011)) that has a pointwise convergence upper bound rate of $O(K^{-2/7}) \simeq O(K^{-1/3})$.⁹ The bandwidth is chosen by cross-validation at the first iteration and kept the same for all others. The Slutsky matrix norm is computed for each s finite set of observations: $\|E\|_{u,s}^2 = \frac{1}{K} \sum_{k=1}^K \|E^{s,u}(p^{s,k}, \bar{w})\|_{\mathbb{R}^{3 \times 3}}^2$ for $K \in \{100, 500, 1000, 2000\}$. The estimator is denoted $\|\hat{E}\|_{u,s}^2$ and the mean of the $n_s = 100$ simulations is denoted $\|\hat{E}\|_u^2 = \frac{1}{100} \sum_{s=1}^{100} \|\hat{E}\|_{u,s}^2$. The hypothesis test is done by computing the statistic $\frac{K \cdot \|\hat{E}\|_u^2}{se^2} \geq cr_{\alpha=0.05}$ where cr_{α} is given as the 95% percentile of a $\chi^2(K)$ and se^2 is the standard error of $\|\hat{E}\|_u^2$ (a naive estimator of the variance). The results of the simulation are provided in Table 1 and show that even for small $K = 100$ the size of the test is appropriate, but the power increases strongly with the sample size, with a power near 1 at $K = 2000$ for the case of $c = \frac{3}{2}$. In general, for all considered K , the size is appropriate (less than 5% error of type I cases) and the power depends strongly on the sample size and on the intensity of the rationality violations with acceptable power for the moderate case at $K = 1000$ with a power of 0.83.

We must underline that the Naradaya-Watson estimator is biased and the absolute error contains that effect too. Our theoretical discussion was done under the assumption of an unbiased estimator, but this bias is typically small.

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 in its components (σ, ν) , by construction the same hypothesis test can be applied to each type of violation since the violations are additive and orthogonal. The power and the size of the hypothesis test specialized for violations of symmetry (σ) and NSD (ν) will depend on its intensity and sample size in the same way that the overall measure of bounded rationality did. The Montecarlo analysis for the analysis for the decomposition of the SMN are given in the appendix (8).

The precision of the estimate in finite sample: Let $s \in \{1, \dots, 1000\}$ dataset O_s^{*100} is generated with $K = 100$ i.i.d. draws from $\ln(p_l) \sim N(\ln(2), 1)$ for $l \in \{1, 2, 3\}$ and from $\epsilon_l \sim N(0, \frac{1}{10 \cdot l})$ for $l \in \{1, 2\}$, such that we obtain: $O_s^{*100} = \{p^{s,k}, b^{s,*}(p^{s,k})\}_{k=1}^{100}$. The key parameter c is fixed at $c = \frac{3}{2}$, to generate both types of violations (i.e., of σ and ν). The results of the simulation are provided in table (2) and show that even for small $K = 100$ 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.

⁹Its uniform convergence is slower $O(\frac{\ln(K)}{K})^{\frac{2}{7}}$ for 3 continuous regressors -denominator $q + 4$ with $q = 3$.

Variables	Estimated	Actual	Abs.Err.
$\ \hat{E}\ _u^2$ vs. $\ E\ _u^2$	0.109	0.145	0.035
$\ \hat{E}^\sigma\ _u^2$ vs. $\ E^\sigma\ _u^2$	0.029	0.031	0.002
$\ \hat{E}^\nu\ _u^2$ vs. $\ E^\nu\ _u^2$	0.080	0.114	0.033

Table 2: Montecarlo simulation for $n_s = 1000$ experiments for $K = 100$.

6 Application

We use in this section an experimental dataset gathered by Ahn et al. (2014), to which we refer the reader for the experimental design and details of the dataset. Here, we provide only the necessary information for this section to be self-contained. The dataset is the outcome of 154 subject solving 50 independent portfolio choice problems. Ahn et al. (2014), present the subjects with three states of the world $l \in \{1, 2, 3\}$. For each state, the subjects can buy an Arrow security that pays 1 token in state l and nothing in other states. The probability of state 2 occurs with $\pi_2 = \frac{1}{3}$ and the other states occur with unknown probabilities that add up to $\pi_1 + \pi_3 = \frac{2}{3}$. Let x_l be the demand for each state contingent Arrow security and p_l its price. The budget constraint can be written as $p \cdot x = 1$. The subject must choose any $x \geq 0$ that satisfies Walras' law. The subject's task is to pick a point on the budget hyperplane, which is represented graphically, using a mouse. The subjects receive a payment according to the probability of each state at each round. The effective choice set $X = \mathbb{R}_+^3$ and the budget hyperplanes are drawn with uniform probability. Any given individual produces a sequence of observations: $O^{50} = \{p^k, x^k\}_{k=1}^{50}$ (where we omit the index i for each individual). The wealth is fixed at $p^k x^k = 1$, in that case we have to consider the conditional Slutsky matrix as in John (1995), but all our results go through. Due to the experimental set-up we can only observe violations of the symmetry and Negative semidefiniteness conditions. Alternatively, rationality here amounts to the Ville Axiom and the Wald Axiom (which is equivalent to WARP in this setup). The main objective of this section is to measure the Slutsky matrix norm (in its unit-free version, as described in Aguiar and Serrano (2014)) for each individual that does not satisfy GARP, in order to measure the average error of making comparative statics analysis under the assumption of rationality given the data.

Using Theorem 1, all extensions of the data O^K (\mathcal{X}^K) when rationality is rejected have a positive SMN, and its decomposition $\|E\|^2 = \|E^\sigma\|^2 + \|E^\nu\|^2$ is related to revealed demand cycles of different lengths and to the underlying Ville axiom and WARP. For this reason, we investigate only the subsample of subjects that fail the GARP test and are boundedly rational. We know by Proposition 3 that any extension of the data will have these properties and will approximate the true value (up to some error) with accuracy increasing in K . We show that the distribution of such errors associated with symmetry first-order stochastically dominates the distribution of errors associated with the NSD property. If we aim to improve our predictions, this provides evidence in favor of models that maintain only the NSD restriction, allowing violations of symmetry, which seem uniformly more likely. We also show that both types of violations are dependent in a monotonic way (the rank they induce is highly correlated), suggesting that violations of the two SMN conditions have a common cause. In addition, the SMN measure is moderately related (rank correlation) with the Afriat index, even when using an arbitrary interpolator. We compute the Budget Share Slutsky matrix, let $b_l(p, \bar{w}) = p_l x_l(p, \bar{w}) / \bar{w}$ for $l = \{1, 2, 3\}$ such that $\sum_{l=1}^3 b_l(p, \bar{w}) = 1$, $S^u(p, \bar{w}) = [\bar{D}_{ln(p)} b(p, \bar{w})] - [\bar{D}_{ln(p)} b(p, \bar{w}) \mathbf{1} b(p, \bar{w})']$ where

Variables	Desc.	Spearman -order- corr.	Pearson corr.
$\ E\ _u^2$ vs. \mathcal{G}	SMN vs. Afriat's	-0.602	-0.499
$\ E^\sigma\ _u^2$ vs. \mathcal{G}	SMN symmetry vs. Afriat's	-0.532	-0.455
$\ E^\nu\ _u^2$ vs. \mathcal{G}	SMN NSD vs. Afriat's	-0.594	-0.411
$\ E^\sigma\ _u^2$ vs. $\ E^\nu\ _u^2$	SMN symmetry vs. NSD	0.587	0.460

Table 3: Order -Spearman- correlation and Pearson Correlation of the Slutsky Matrix Norm estimates.

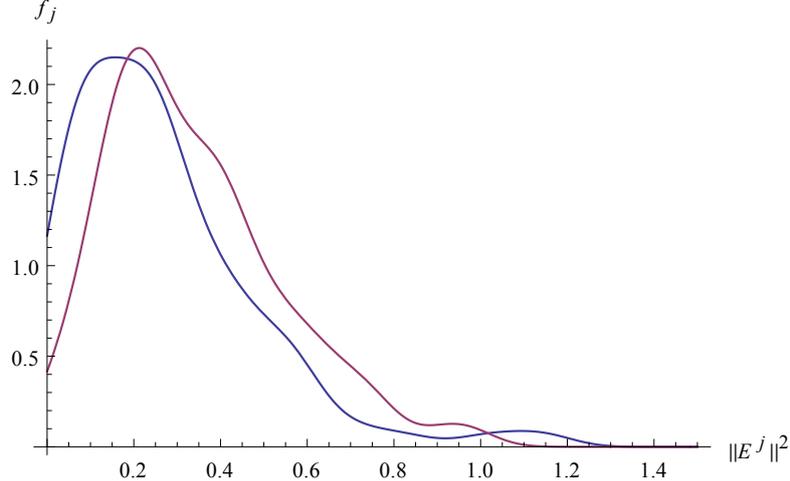
$\bar{D}_{ln(p)}b(p, \bar{w}) = D_{ln(p)}b(p, \bar{w}) - \text{Diag}(b(p, \bar{w}))$. Notice that $S^u(p, \bar{w}) = \frac{1}{w} \text{Diag}(p)S(p, \bar{w})\text{Diag}(p)$ is the unit-free Slutsky matrix budget share from (Aguiar & Serrano, 2014). The Slutsky matrix norm is computed only for the finite set of observations so $\|E\|_u^2 = \frac{1}{50} \sum_{k=1}^{50} \|E^u(p^k, \bar{w})\|_{\mathbb{R}^{3 \times 3}}^2$. The estimator of $\|E^u(p^k, \bar{w})\|_{\mathbb{R}^{3 \times 3}}^2$ is computed using a Hermite polynomial of order one, that uses polygonal hypersurfaces to connect the data. This is done in order to obtain an approximate lower bound of the Slutsky matrix norm.¹⁰ We use commercial software algorithms of automatic differentiation to construct a fast algorithm to approximate the quantity of interest. The main difference of this approach with the traditional work of Haag et al. (2009) is that in the Ahn et al. (2014) dataset, there is no measurement error and so we use interpolators of the dataset instead of approximating functions. Any departure from rationality in this setup is the combination of an error of approximation of the derivative and the behavior of the subjects. Due to the moderately large sample, the SMN is taken as a good estimate of the true deviations from rationality (error estimates are provided in the appendix using a leave one out scheme).

The Empirical Relation of Afriat's Index with SMN Measures: Using this framework, we provide initial evidence of the relationship between violations of symmetry and violations of NSD. The distribution of both types of errors has a high ordinal correlation (Spearman ρ correlation) and a moderate Pearson correlation (traditional) suggesting that the relation between this two random variables is strong, monotone but not linear (Table 3). This result has interest in its own right, and relates tangentially to the literature concerned with learning the joint distribution of behavioral biases (Dean & Ortoleva, 2014). The second stylized fact is that the distribution of violations of symmetry first-order stochastically dominates the distribution of violations of NSD. This evidence supports models that relax the first condition and maintain the second, as was already expected to happen in the original work of Kihlstrom et al. (1976). This finding gives empirical grounds to the theoretical idea that symmetry is a fragile property (based on equalities), while the NSD is a more robust behavioral feature (based on an inequality) – (Aguiar & Serrano, 2014). Finally, the ordinal correlation of the SMN and Afriat's efficiency index surpasses 60%, again coexisting with a comparable Pearson correlation. This monotone dependence corroborates the results delineated by Jerison & Jerison (2012). The numerical value of the SMN itself has interest as a measure of error in prediction of comparative statics analysis. As we can see, this application mainly illustrates the usefulness of the SMN approach to summarize the violations of rationality in a way that is complementary to the RP approach.

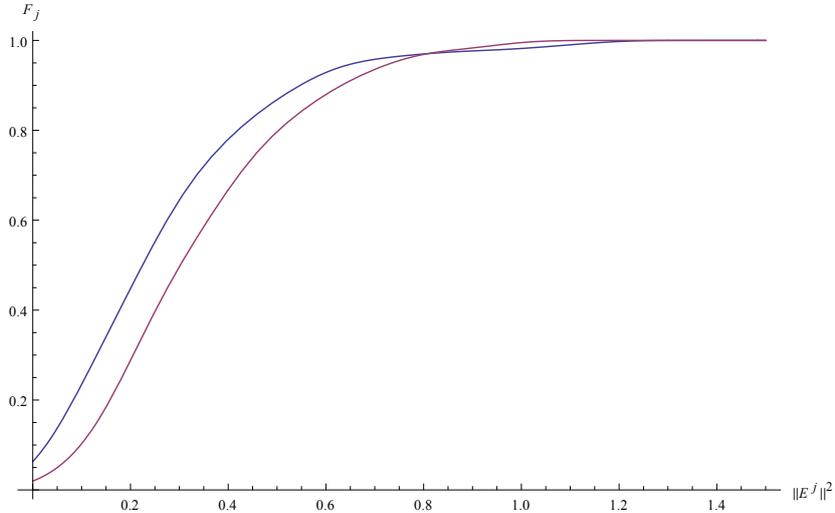
The Stochastic Dominance of Violations of Symmetry w.r.t. NSD: Next, we consider only the population that violates SARP, and we compute the distribution of the violations of

¹⁰The exact solution to the interpolator problem formulated in Corollary 4 is not solved analytically nor it is implemented, and it is left as an open problem. However, the work of Pinkus (1978) on perfect splines that have a similar set-up for a unidimensional case suggest the use of low order interpolators to obtain the lower bound of the SMN while the highest orders approximate the upper bounds (the simulations support this intuition).

the symmetry and NSD conditions. The stochastic dominance of the violations of the symmetry condition can be visualized in Figure 1. The average Afriat index of this dataset is $\mathcal{G} = 0.947$.¹¹ The mean of $\|E^\sigma\|$ is 0.33% that is interpreted as a measure of dispersion of the prediction errors done when assuming symmetry given the data (e.g., when given 1 token, the average prediction amounts to 0.33 tokens). The mean of $\|E^\nu\|_u$ is 0.27%. This evidence supports the theoretical insights of Kihlstrom et al. (1976) that symmetry is more unstable than NSD.



(a) Density (f_j), where $j \in \{\sigma, \nu\}$ and σ is red and ν is blue.



(b) Cumulative distribution (F_j), where $j \in \{\sigma, \nu\}$ and σ is red and ν is blue.

Figure 1: First order stochastic dominance of $\|E^\sigma\|_u$ to $\|E^\nu\|_u$ for individuals that fail SARP.

Testing Rationality: There exists a sharp test based on the SMN that corresponds to the solution and implementation of the class of estimators delineated in Corollary 4. However, the main point here is to show that even for small samples (consisting of as few as 50 observations) the use of an out-of-the-box interpolator does a good job in detecting rationality violations in terms of power and significance. More importantly, by our results in Proposition 3, we know

¹¹The minimum Afriat efficiency is 0.74681 and the maximum is 1 – 12% of subjects satisfy SARP ($\mathcal{G} = 1$). The standard deviation is 0.06.

Passes SARP	Satisfies		
	Symmetry σ	NSD ν	σ, ν .
Yes (12%)	95%	95%	90%
No (88%)	57.5%	20%	17%
Total	62%	29%	25%

Table 4: Testing rationality using the Slutsky Matrix Norm estimates.

that, as the number of observations increases, the SMN estimate will converge to its true value. Thus, we know its performance will improve with richer datasets. Recall that this exercise is deterministic but contains error due to the use of a fast hermite polynomial interpolator that has gains in terms of speed and interpretation, but its use has a cost in terms of accuracy. In order to formulate a test that takes into account this fuzzyness, we compute a test quantity (analogous to a test statistic) for the estimated SMN. In particular, we compute the error of the estimation (η_j for $j \in \{\sigma, \nu\}$) and then compute the test quantity $\hat{\sigma}^i = \frac{\|E^{i,\sigma}\|_{\mathbf{u}}^2}{|\eta_{\sigma}^i|}$, and $\hat{\nu}^i = \frac{\|E^{i,\nu}\|_{\mathbf{u}}^2}{|\eta_{\nu}^i|}$ for all $i \in \{1, \dots, 154\}$ individuals.¹² The critical value of each quantity is computed to guarantee a 5% significance level (proportion of type I error).

In the case of testing the symmetry property the power to test rationality on its own is 42% while the power of the NSD property to test rationality is 80% alone. It is encouraging that the NSD test alone has a high power to test rationality, maintaining a standard size (significance) of 5%. We can reject that the individual Slutsky matrix has the property (σ, ν) when $\hat{\sigma}^i > \tau_{\sigma}$ and similarly when $\hat{\nu}^i > \tau_{\nu}$.

7 Literature Review

The current research can be considered as an extension and generalization of the work of Chiappori and Rochet (1987) and Lee & Wong (2005) that prove the equivalence of SARP and rationalizability 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 result generalizes these findings, taking them to the realm of quantifying departures from rationality. Jerison and Jerison (2012) innovative work regarding the local equivalence between Afriat's inefficiency index and a quadratic form of the Slutsky matrix is also used as a stepping stone in establishing the relationship between the SMN and the RP approaches. 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 & Martin (2013) and others. Our work can be seen as a specialization of the work of Nishimura et al. (2013), by establishing a unified approach to the Richter and the Afriat traditions using multiple choice correspondence extensions of a finite dataset for the results of the first section. Furthermore, it also generalizes the approach by establishing its extension to a noisy or measurement error environment. The SMN approach has desirable properties such as power against all alternatives and closed-form asymptotic behavior under regularity conditions. The results in the nondeterministic part of the paper generalize those in Haag et al. (2009) for testing symmetry

¹²The rationale for doing this is that the numerator corresponds to the value of absolute error under the null that the SMN is zero while the denominator is the estimate of the true absolute error. When the null is false the numerator will increase, thus we conclude that the true SMN is positive when the test quantity is large enough. This is in line with δ -decidability (Gao et al., 2012) or approximate decidability (Hemmerling, 2003).

to all Slutsky regularity conditions. The bootstrapping techniques proposed by these authors can be combined with our approach fruitfully. Our approach is an alternative to the methodologies proposed in Blundell et al. (2008) or Epstein & Yatchew (1985), which depend on the demand function instead of the Slutsky matrix function. It improves upon them by obtaining closed-form solutions for the asymptotics and by having a simpler implementation, as well as a clear interpretation as a positive index of departures from rationality that has an axiomatic foundation provided in Aguiar & Serrano (2014). Finally, we also provide a new way to analyze the power of consumer rationality tests with respect to all possible alternatives, a very significant improvement with respect to the current practice born with Bronars (1987) that uses as the non-rational benchmark the Becker (uniformly) random consumer. We show that our SMN approach in the presence of noise has an asymptotic power of 1 with respect to all possible alternatives. This result is similar to Lewbel (1995), but the latter is only done for the symmetry condition.

8 Conclusion

We present a unification of the RP and SMN approaches to measure departures from rationality. In particular, we conclude that: (i) testing rationality in both settings is equivalent, (ii) Locally a quadratic form of the Slutsky error norm converges to zero proportionally to the Afriat's inefficiency index. (iii) In the presence of noise, the smooth approach that characterizes the SMN allows us to obtain a closed-form asymptotic theory for the measures of bounded rationality. Our unified approach is smooth and easy to implement, and it avoids the issues of partial identification related to the use of inequalities in the RP approach. Finally, the SMN measure of departures from rationality is a positive measure, and as such, it allows interpersonal comparisons and it can be monetized. It also provides a new criterion to judge whether maintaining the rationality hypothesis is acceptable, that of failures of making comparative statics analysis. We have applied it to experimental data, to conclude that insisting on the Slutsky matrix symmetry hypothesis leads to making a higher error in this exercise than maintaining the NSD of the Slutsky matrix. As far as we know, this is the first evidence brought forward on the claims of the higher robustness of WARP with respect to the Ville axiom as already suggested in Kihlstrom et al. (1976).

References

- Ackerberg, D., Chen, X., & Hahn, J. (2012). A practical asymptotic variance estimator for two-step semiparametric estimators. *Review of Economics and Statistics*, 94(2), 481–498.
- Afriat, S. N. (1973). On a system of inequalities in demand analysis: an extension of the classical method. *International economic review*, 14(2), 460–472.
- Aguiar, V. H. & Serrano, R. (2014). Slutsky matrix norms and the size of bounded rationality. *Available at SSRN 2378585*.
- Ahn, D., Choi, S., Gale, D., & Kariv, S. (2014). Estimating ambiguity aversion in a portfolio choice experiment. *Quantitative Economics*, 5(2), 195–223.
- Antonelli, G. B. (1886). *Sulla teoria matematica della economia politica*. Malfasi.

- Azagra, D. & Boiso, M. C. (2004). Uniform approximation of continuous mappings by smooth mappings with no critical points on hilbert manifolds. *Duke Mathematical Journal*, 124(1), 47–66.
- Blundell, R., Browning, M., & Crawford, I. (2008). Best nonparametric bounds on demand responses. *Econometrica*, 76(6), 1227–1262.
- Bronars, S. G. (1987). The power of nonparametric tests of preference maximization. *Econometrica*, 55(3), 693–698.
- Chiappori, P.-A. & Rochet, J.-C. (1987). Revealed preferences and differentiable demand. *Econometrica: Journal of the Econometric Society*, (pp. 687–691).
- De Blasi, F. S. & Myjak, J. (1986). On continuous approximations for multifunctions. *Pacific Journal of Mathematics*, 123(1), 9–31.
- de Clippel, G. & Rozen, K. (2014). Bounded rationality and limited datasets: Testable implications, identifiability, and out-of-sample prediction. Working Paper.
- Dean, M. & Martin, D. (2013). Measuring rationality with the minimum cost of revealed preference violations.
- Dean, M. & Ortoleva, P. (2014). Is it all connected? a testing ground for unified theories of behavioral economics phenomena. *Unpublished manuscript*.
- Echenique, F., Lee, & Shum, M. (2011). The money pump as a measure of revealed preference violations. *Journal of Political Economy*, 119(6), 1201–1223.
- Epstein, L. G. & Yatchew, A. J. (1985). Non-parametric hypothesis testing procedures and applications to demand analysis. *Journal of Econometrics*, 30(1), 149–169.
- Gao, S., Avigad, J., & Clarke, E. (2012). Delta-decidability over the reals. In *2012 27th Annual IEEE Symposium on Logic in Computer Science (LICS)* (pp. 305–314).
- Grodal, B. (1974). A note on the space of preference relations. *Journal of Mathematical Economics*, 1(3), 279–294.
- Haag, B. R., Hoderlein, S., & Pendakur, K. (2009). Testing and imposing slutsky symmetry in nonparametric demand systems. *Journal of Econometrics*, 153(1), 33–50.
- Hemmerling, A. (2003). Approximate decidability in euclidean spaces. *Mathematical Logic Quarterly*, 49(1), 34–56.
- Hurwicz, L. & Richter, M. K. (1979). Ville axioms and consumer theory. *Econometrica*, 47(3), 603–619. ArticleType: research-article / Full publication date: May, 1979 / Copyright © 1979 The Econometric Society.
- Hurwicz, L. & Uzawa (1971). On the integrability of demand functions. In *Cap 6. Preferences, Utility and Demand* (pp. 114). New York: Harcourt.
- Jerison, D. & Jerison, M. (1992). Approximately rational consumer demand and ville cycles. *Journal of Economic 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. *Economics Letters*, 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, M. K. (1986). Nontransitive-nontotal consumer theory. *Journal of Economic Theory*, 38(2), 324–363.
- Lee, P. M. & Wong, K.-C. (2005). Revealed preference and differentiable demand. *Economic Theory*, 25(4), 855–870.
- Lewbel, A. (1995). Consistent nonparametric hypothesis tests with an application to Slutsky symmetry. *Journal of Econometrics*, 67(2), 379–401.
- Li, Q. & Racine, J. S. (2011). *Nonparametric econometrics: Theory and practice*. Princeton University Press.
- Mas-Colell, A. (1974). Continuous and smooth consumers: Approximation theorems. *Journal of Economic Theory*, 8(3), 305–336.
- Newey, W. K. (1994). The asymptotic variance of semiparametric estimators. *Econometrica: Journal of the Econometric Society*, (pp. 1349–1382).
- Nishimura, H., Ok, E. A., & Quah, J. (2013). A unified approach to revealed preference theory: The case of rational choice.
- Pinkus, A. (1978). Some extremal properties of perfect splines and the pointwise Landau problem on the finite interval. *Journal of Approximation Theory*, 23(1), 37–64.
- Shafer, W. J. (1977). Revealed preference cycles and the Slutsky matrix. *Journal of Economic Theory*, 16(2), 293–309.
- Slutsky, E. (1915). Sulla teoria del bilancio del consumatore. *Giornale degli economisti e rivista di statistica*, (pp. 1–26).
- Varian, H. R. (1983). Non-parametric tests of consumer behaviour. *The Review of Economic Studies*, 50(1), 99–110.

Appendix

This appendix presents several results that have been referred to in the main text. In particular, we have the following technical remark that tightens the connection of the discrete and continuous approaches.

Proof of Corollary (5)

Proof. To violate the Ville axiom we just have to build a Ville cycle C^1 . We use a quadratic spline,

Let $x \in \mathcal{X}_L$, be defined pointwise by a quadratic spline $x(\tau) = \frac{z^{t+1}-z^t}{2[\tau_{t+1}-\tau_t]}[\tau-\tau_t]^2 + z^t[\tau-\tau_t] + x^t$ for $\tau \in [\tau_t, \tau_{t+1}]$.

It's derivative is: $\frac{\partial x(\tau)}{\partial \tau} = \frac{z^{t+1}-z^t}{[\tau_{t+1}-\tau_t]}[\tau-\tau_t] + z^t$

with $z^{t+1} = -z^t + 2\left(\frac{x^{t+1}-x^t}{\tau_{t+1}-\tau_t}\right)$ with $z^0 = \frac{\partial x(\tau)}{\partial \tau}|_{\tau_0=0}$.

Notice that letting $z^{t+1} - z^t = B^{t+1}$ we have that

$$\frac{\partial x(\tau)}{\partial \tau} = \frac{B^t}{[\tau_{t+1}-\tau_t]}[\tau-\tau_t] + z^t$$

$$B^t = 2\left(\frac{x^{t+1}-x^t}{\tau_{t+1}-\tau_t}\right)$$

Then multiply by prices

$$p^{t+1}B^t = 2\left(\frac{p^{t+1}[x^{t+1}-x^t]}{\tau_{t+1}-\tau_t}\right) > 0$$

and

$$\frac{\partial x(\tau)}{\partial \tau} = \frac{B^t}{[\tau_{t+1}-\tau_t]}[\tau-\tau_t] + z^t$$

$$\frac{\partial x(\tau)}{\partial \tau} = \frac{B^t}{[\tau_{t+1}-\tau_t]}[\tau-\tau_t] + z^t$$

$$p^{t+1}\frac{\partial x(\tau)}{\partial \tau} = \frac{p^{t+1}B^t}{[\tau_{t+1}-\tau_t]}[\tau-\tau_t] + p^{t+1}z^t$$

$$p^{t+1}\frac{\partial x(\tau)}{\partial \tau} = p^1B^0[\tau-\tau_0] + p^1z^0$$

Since I have a free parameter fix $p^1z^0 > 0$ then

$$p^{t+1}\frac{\partial x(\tau)}{\partial \tau} = p^1B^0[\tau-\tau_0] + p^1z^0 > 0 \text{ for } \tau \in [0, \tau_1]$$

Also,

$$p^1B^0 = p^1z^1 - p^1z^0 > 0 \text{ implies } p^1z^1 > p^1z^0 > 0$$

$$p^2B^1 = p^2z^2 - p^2z^1 > 0 \text{ implies } p^2z^2 > p^2z^1 > 0.$$

This can be extended for all t

$$p^tB^t = p^tz^t - p^tz^{t-1} > 0 \text{ such that } p^tz^t > p^tz^{t-1} > 0.$$

And this in turns implies that

$$p^{t+1}\frac{\partial x(\tau)}{\partial \tau} > 0 \text{ for all } \tau \in [\tau_t, \tau_{t+1}]$$

Let $p(\tau) = p^{t+1}$ if $\tau \in (\tau_t, \tau_{t+1}]$ such that $p^0 = p(0) = p(T) = p^T$ and let $w(\tau)$ be defined implicitly as $w(\tau) = p(\tau)x(\tau)$.

Then $x : [0, T] \rightarrow X$ forms a C^1 Ville Cycle thus breaking the Ville Axiom.

We can mollify the step function $p(\tau)$ for it to be continuous and smooth without loosing the main result stated above (i.e. convolute it with an infinitely differentiable function such that $\bar{p} : [0, T] \rightarrow P$ is close enough to $p : [0, T] \rightarrow P$ in order not to alter the inequalities. \square)

The Connection of the Slutsky Matrix Norm and the Money Pump Index

We start with some preliminary definitions.

Definition 12. (Slutsky demand) A Slutsky demand is a function $q : P \times \mathbb{R}_+^L \mapsto X$, defined pointwise as $q(p, r) = x(p, p'r)$, given a price p and a reference bundle r .

The main feature of the Slutsky demand that we are interested in is that its jacobian with respect to prices is $D_p q(p, r) = D_p x(p, p'r) + D_w x(p, p'r)r'$. In particular, when $r = x(p, w)$ we get the Slutsky matrix $D_p q(p, r)|_{r=x(p, w)} = S(p, w)$ under Walras' law. Additionally, we have the simple property that $q(p, x(p, w)) = x(p, w)$. Finally, the jacobian with respect to the second argument is $D_r q(p, r) = D_w x(p, p'r)p'$ that will come in handy later.

It is interesting to notice that a ‘‘rational’’ Slutsky demand has a symmetric, NSD and singular in prices jacobian only evaluated at $r = x(p, w)$, elsewhere it may not be the case.

The first step is to characterize what we mean by the Walrasian and Slutsky demands to be rational. The first discrete characterization will be given by analogy to the GARP.

Definition 13. The following statements are equivalent.

- (i) A walrasian demand is said to be rationalizable if it is generated by maximizing a locally nonsatiated continuous and concave utility function subject to a linear budget constraint.
- (ii) A walrasian demand is said to be rational if it generates a dataset that satisfies SARP evaluated at any sequence of prices and wealth.
- (iii) A walrasian demand is said to be rational if it satisfies the Ville axiom and the weak WARP.

In particular, we say that a dataset generated by a demand function or DGDF $x \in \mathcal{X}$ satisfies SARP if there are not revealed preference cycles.

If we observe a revealed demand cycle $C^{D, T} = \{p^t, w^t, x^t\}_{t=0}^T$ we have: $p^t x(p^t, w^t) > p^t x(p^{t-1}, w^{t-1})$, where $x \in \mathcal{X}$ is the true DGDF, we can also write the same cycle condition using Slutsky demands:

$$p^t q(p^t, x^t) > p^t q(p^{t-1}, x^{t-1}) = p^t x(p^{t-1}, w^{t-1})$$

Finally, we know that $p^t x^t > p^t x \iff x^t R x$ for all $x \in \mathbb{R}_+^L$, where R denotes the revealed preference relation.

Then we can define the more general revealed Slutsky demand cycle $C^{S, T} = \{p^t, r^t, q^t\}_{t=0}^T$, where q is a Slutsky demand and $q^t = q(p^t, r^t)$ and for all $t \in \{0, \dots, T\}$, we have $p^t q(p^t, r^t) > p^t q(p^{t-1}, r^{t-1})$ with $r^0 = r^T$, $p^0 = p^T$, and $r^t \in \mathbb{R}_+^L$.

In particular, this means that for a constant sequence $r^t = r$ (a trivial cycle), there is a violation of rationality if we observe the following chain $x(p^T, p^T r) R x(p^{T-1}, p^{T-1} r) \dots x(p^1, p^1 r) R x(p^0, p^0 r)$ which is a violation of SARP. These generalized cycles are of interest because we can make them equivalent to the standard cycles around a reference point (\bar{p}, \bar{w}) (i.e., locally) and they are strongly related to the Slutsky matrix norms, thus constituting the bridge across the two approaches. In a similar spirit, a generalized Ville cycle or Slutsky cycle $C^{G(S), b} = (p, r)$ ($p, r : [0, b] \rightarrow \mathbb{R}_+^L$ and $p, r \in \mathcal{C}^S([0, b])$) can be defined as a situation where $p(\tau)' \frac{\partial q(p(\tau), r(\tau))}{\partial \tau} \geq 0$ for $\tau \in [0, b]$ and $p(0) = p(b)$ and $r(0) = r(b)$. In particular notice that $p(\tau)' \frac{\partial q(p(\tau), r(\tau))}{\partial \tau} = p(\tau)' \frac{\partial x(\tau)}{\partial \tau}$ when we let $r(\tau) = x(\tau)$ for all $\tau \in [0, b]$. Hence, the Ville cycle is a specific kind of Slutsky cycle. It is also useful to define a continuous version of the JJ index.

Definition 14. Continuous JJ index $\iota(v, S) = \int_0^b v(\tau) S(\bar{p}, \bar{w}) \frac{\partial v(\tau)}{\partial \tau}$ where $v : [0, b] \rightarrow \mathbb{R}^L$, $v(0) = v(b)$, $p(\tau) = \bar{p} + v(\tau)$ and $p(\tau)x(\tau) = w(\tau)$ with $S(\bar{p}, \bar{w})$ the Slutsky matrix of x at the fixed price and wealth pair.

The first main result of this subsection is a connection between the existence of a Slutsky cycle with the JJ index.

Lemma 8. *The aggregated real income growth $\int_0^b p(\tau)' \frac{\partial x(\tau)}{\partial \tau} d\tau$ for a cycle around a reference point (\bar{p}, \bar{w}) can be linearly approximated by the continuous JJ index for an associated cycle around the same reference point.*

Proof. Let $r(\tau) = x(\bar{p}, \bar{w})$ for $\tau \in [0, b]$ and obtain the first order approximation of the Slutsky demand $q(p, x(\bar{p}, \bar{w})) = x(\bar{p}, \bar{w}) + D_p q(\bar{p}, x(\bar{p}, \bar{w}))[p - \bar{p}] + o(\mathbf{1} \cdot |p - \bar{p}|^2)$.

Then we get the approximation of the time derivative $\frac{\partial q(p(\tau))}{\partial \tau} = D_p q(\bar{p}, \bar{w}) \frac{\partial p(\tau)}{\partial \tau} + o(|p - \bar{p}|) = S(\bar{p}, \bar{w}) \frac{\partial p(\tau)}{\partial \tau} + o(\mathbf{1} \cdot |p(\tau) - \bar{p}|)$, where $|\cdot|$ is the supremum norm defined for the path functions $p(\cdot)$.

This implies that $p(\tau)' \frac{\partial q(p(\tau), r(\tau))}{\partial \tau} = p(\tau)' S(\bar{p}, \bar{w}) \frac{\partial p(\tau)}{\partial \tau} + o(|p - \bar{p}|)$, since it is the constant approximation of its derivative.

Finally, we can write without loss of generality $p(\tau) = \bar{p} + v(\tau)$, under Walras' law this implies $p(\tau)' S(\bar{p}, \bar{w}) \frac{\partial p(\tau)}{\partial \tau} \simeq v(\tau)' S(\bar{p}, \bar{w}) \frac{\partial v(\tau)}{\partial \tau}$ since $\bar{p}' S(\bar{p}, \bar{w}) = 0$.

Then we obtain the following result $\epsilon_p = |p - \bar{p}|$, then let $\gamma^S = \int_0^b p(\tau)' \frac{\partial q(p(\tau), r(\tau))}{\partial \tau} d\tau$ and $\gamma^{JJ} = \int_0^b v(\tau)' S(\bar{p}, \bar{w}) \frac{\partial v(\tau)}{\partial \tau}$, so $\gamma^S - \gamma^{JJ} = o(\epsilon_p)$ and $\lim_{\epsilon_p \rightarrow 0} \frac{\gamma^S - \gamma^{JJ}}{\epsilon_p} \rightarrow 0$.

The γ^{JJ} is nothing but the continuous time version of the JJ index for a cycle $v : [0, b] \rightarrow \mathbb{R}^L$ with $v(0) = v(b)$.

Notice that $|p(\tau)' \frac{\partial q(p(\tau), x(\bar{p}, \bar{w}))}{\partial \tau} - p(\tau)' \frac{\partial q(p(\tau), x(\tau))}{\partial \tau}| \rightarrow 0$ as $\epsilon = |(\bar{p}, \bar{w}) - (p, w)| \rightarrow 0$, more simply $|x(\tau) - x(\bar{p}, \bar{w})| \rightarrow 0$ when we let $\epsilon \rightarrow 0$ then we let $\gamma^V = \int_0^b p(\tau)' \frac{\partial x(\tau)}{\partial \tau} d\tau$ and we conclude $\lim_{\epsilon \rightarrow 0} \frac{\gamma^V - \gamma^{JJ}}{\epsilon} = 0$. \square

The main point of this result is that for a small enough $\epsilon \rightarrow 0$ we have $x(\tau) \simeq x(\bar{p}, \bar{w})$ thus making $\lim_{\epsilon \rightarrow 0} \epsilon^{-2} [\int_0^b p(\tau)' \frac{\partial x(\tau)}{\partial \tau} d\tau - \int_0^b v(\tau)' S(\bar{p}, \bar{w}) \frac{\partial v(\tau)}{\partial \tau} d\tau] = 0$. In other words, locally the Slutsky cycles coincide with the Ville cycles thus yielding the same real income growth in the interval $[0, b]$. This reasoning provides a “continuous” version of the conclusion reached by Jerison and Jerison (2012) for the “discrete” case. More precisely, the Slutsky cycle for $r = x(\bar{p}, \bar{w})$ converges at a rate of ϵ^{-2} to the Ville cycle in a small open ball around the reference point (\bar{p}, \bar{w}) . This local convergence result will be preserved in the discrete or finite data case thus making the connection with several measures based on revealed preference cycles.

To see this, note that $p(\tau)' \frac{\partial q(p(\tau), r(\tau))}{\partial \tau} = p(\tau)' D_p q(p(\tau), r(\tau)) \frac{\partial p(\tau)}{\partial \tau} + p(\tau)' D_r q(p(\tau), r(\tau)) \frac{\partial r(\tau)}{\partial \tau}$ when $r(\tau) = x(\tau)$ then $D_r q(p(\tau), x(\tau)) = D_w x(\tau) p(\tau)'$, also $p(\tau)' D_w x(\tau) = 1$ clearly for $p(\tau)' D_p q(p(\tau), x(\tau)) = p(\tau)' S(p(\tau), w(\tau)) = 0$ such that $p(\tau)' \frac{\partial q(p(\tau), r(\tau))}{\partial \tau} = p(\tau)' \frac{\partial x(\tau)}{\partial \tau}$.

Remark 6. The discrete JJ index can be generated by discretizing the path $v(\tau)$ and letting the interval $[0, b]$ be replaced by T evenly spaced knots $0 \leq \tau_2 \leq \tau_3 \leq \dots \tau_{T-1} \leq b$ ($\tau_t = \tau_{t-1} + h$). Then $\int_0^b v(\tau)' S(\bar{p}, \bar{w}) \frac{\partial v(\tau)}{\partial \tau} d\tau \simeq \frac{1}{T} \sum_{t=0}^T v_t' S(\bar{p}, \bar{w}) [v_t - v_{t-1}]$, for $v_t = v(\tau_t)$ and $v_1 = v_T$ a cycle.

This discrete JJ index will converge to the continuous JJ index when $h \rightarrow 0$ and the continuous JJ index will converge to aggregated real income growth around a reference point (\bar{p}, \bar{w}) at the rate $\epsilon = |(\bar{p}, \bar{w}) - (p, w)|$. By definition $h < \epsilon$ then the discrete JJ index will approximate the discrete aggregated real income.

Our Slutsky cycle results allow us to make another connection of the JJ index with the well known Money pump index, however this requires the convergence of $\epsilon \rightarrow 0$, that is both $\epsilon_p \rightarrow 0$ and $\epsilon_w \rightarrow 0$. In that case, the denominator of the Money pump index converges to the JJ index.

First, note that the numerator of the MPI can be defined along a cycle as

$$MPC = \sum_{t=0}^K p^t(x^t - x^{t-1}) = \sum_{t=1}^K (w^t - p^t x^{t-1}) \text{ we let } t = \tau_t \text{ and } \tau_t = \tau_{t-1} + h \text{ and } h \rightarrow 0.$$

$$MPC = \sum_{t=0}^K p^t(x^t - x^{t-1}) \rightarrow \int_0^b p(\tau)' \frac{\partial x(\tau)}{\partial \tau} d\tau.$$

Now, we can state the second result of this subsection, which makes use of Lemma 8.

Corollary 7. *The MPI index of a Ville cycle converges towards a normalized JJ index locally.*

Proof. For a Ville cycle, let $MPI = \int_0^b p(\tau)' \frac{\partial x(\tau)}{\partial \tau} d\tau / \int_0^b p(\tau)' x(\tau) d\tau$, then notice that the numerator of the MPI converges to the JJ index at a rate of ϵ^{-2} by proposition 2.

Then we conclude that:

$$\lim_{\epsilon \rightarrow 0} \epsilon^{-2} \int_0^b p(\tau)' \frac{\partial x(\tau)}{\partial \tau} d\tau / \int_0^b p(\tau)' x(\tau) d\tau = \int_0^b v(\tau)' S(\bar{p}, \bar{w}) \frac{\partial v(\tau)}{\partial \tau} d\tau / \int_0^b p(\tau)' x(\tau) d\tau. \quad \square$$

Finally, we can discretize the normalized JJ index or the JJ money pump $JJ - MPI$:

Remark 7. We can discretize the JJ-MPI for a given ville cycle as:

$$\int_0^b v(\tau)' S(\bar{p}, \bar{w}) \frac{\partial v(\tau)}{\partial \tau} d\tau / \int_0^b p(\tau)' x(\tau) d\tau \simeq \frac{\sum_{t=0}^T v^t S(\bar{p}, \bar{w}) [v^t - v^{t-1}]}{\sum_{t=0}^T p^t x^t} \text{ with } \{v^t\}_{t=0}^T \text{ a cycle and } p^t = \bar{p} + v^t \text{ and } p(0) = p^1, p^T = p(b).$$

The Montecarlo Analysis of the Decomposition of the SMN with Noise

The results of the σ, ν cases of bounded rationality are presented in table (5) and table (6):

Null hypothesis (σ)	Sample	Proportion null is rejected.		
		$c = 0$ ($\ E^\sigma\ _u^2 = 0$)	$c = 3/2$ ($\ E^\sigma\ _u^2 = 0.03$)	$c = 3$ ($\ E^\sigma\ _u^2 = 0.13$)
$H_0 : \ \hat{E}^\sigma\ _u^2 = 0$	$K = 100$	0.12	0.32	0.55
	$K = 500$	0.01	0.12	0.50
	$K = 1000$	0.01	0.26	0.63
	$K = 2000$	0.01	0.10	0.41

Table 5: Montecarlo simulation for $n_s = 100$, proportion of cases where the null (symmetry) is rejected.

Null hypothesis (ν)	Sample	Proportion null is rejected.		
		$c = 0$ ($\ E^\sigma\ _u^2 = 0$)	$c = 3/2$ ($\ E^\sigma\ _u^2 = 0.11$)	$c = 3$ ($\ E^\sigma\ _u^2 = 10.97$)
$H_0 : \ \hat{E}^\nu\ _u^2 = 0$	$K = 100$	0.00	0.15	0.88
	$K = 500$	0.00	0.75	1.00
	$K = 1000$	0.00	0.90	1.00
	$K = 2000$	0.00	1.00	1.00

Table 6: Montecarlo simulation for $n_s = 100$, proportion of cases where the null (NSD) is rejected.