Slutsky Matrix Norms: The Size, Classification, and
Comparative Statics of Bounded Rationality*

Victor H. Aguiar and Roberto Serrano

This version: October 2015

Abstract

Given any observed demand behavior by means of a demand function, we quantify by how much it departs from rationality. The measure of the gap is the smallest Frobenius norm of the correcting matrix function that would yield a Slutsky matrix with its standard rationality properties (symmetry, singularity, and negative semidefiniteness). A useful classification of departures from rationality is suggested as a result. Errors in comparative statics predictions from assuming rationality are decomposed as the sum of a behavioral error (due to the agent) and a mis-specification error (due to the modeller). Comparative statics for the boundedly rational consumer is expressed as path dependence of wealth compensations in price-equivalent paths, or as violations of the compensated law of demand. Illustrations are provided using several bounded rationality models.

JEL classification numbers: C60, D10.

Keywords: consumer theory; rationality; Slutsky matrix function; bounded rationality; comparative statics.

1 Introduction

The rational consumer model has been at the heart of most theoretical and applied work in economics. In the standard theory of the consumer, this model has a unique prediction in the form of a symmetric, singular, and negative semidefinite Slutsky matrix. In fact, any demand system that has a Slutsky matrix with these properties can be viewed as being generated as the result of a process of maximization of some rational preference relation. Nevertheless, empirical evidence often derives demand systems that conflict with the rationality paradigm. In such cases, those hypotheses (e.g., symmetry of the Slutsky matrix) are rejected. These important findings have given rise to a growing literature of behavioral models that attempt to better fit the data.

At this juncture three related questions can be posed in this setting:

*This paper subsumes Aguiar and Serrano (2014). We are especially indebted to Xavier Gabaix, Michael Jerison, and Joel Sobel for suggesting many specific improvements to that earlier version of the paper. We also thank Bob Anderson, Francis Bloch, Mark Dean, Federico Echenique, Drew Fudenberg, Peter Hammond, Susanne Schennach, Larry Selden, Jesse Shapiro and the participants at several conferences and seminars for helpful comments and encouragement.

1Department of Economics, Brown University, victor_aguiar@brown.edu
2Department of Economics, Brown University, roberto_serrano@brown.edu
• (i) How can one measure the distance of an observed demand behavior –demand function– from rationality?

• (ii) How can one compare and classify two behavioral models as departures from a closest rational approximation?

• (iii) Given an observed demand function, what is the best rational approximation model?

The aim of this paper is to provide a tool to answer these three questions in the form of a Slutsky matrix function norm, which allows to measure departures from rationality in either observed Slutsky matrices or demand functions. The answer, provided for the class of demand functions that are continuously differentiable, sheds light on the size and type of bounded rationality that each observed behavior exhibits.

Our primitive is an observed demand function. To measure the gap between that demand function and the set of rational behaviors, one can use the “least” distance and try to identify the closest rational demand function. This approach presents serious difficulties, though. Leaving aside compactness issues, which can be addressed under some regularity assumptions, the solution would require solving a challenging system of partial differential equations. Lacking symmetry of this system, an exact solution may not exist, and one needs to resort to approximation or computational techniques, but those are still quite demanding.

We take an alternative approach, based on the calculation of the Slutsky matrix function of the observed demand. We pose a matrix nearness problem in a convex optimization framework, which permits both a better computational implementability and the derivation of extremum solutions. Indeed, we attempt to find the smallest correcting additive perturbation to the observed Slutsky matrix function that will yield a matrix function with all the rational properties (symmetry, singularity with the price vector on its null space, and negative semidefiniteness). We use the Frobenius norm to measure the size of this additive factor, interpreting it as the “size” of the observed departure from rationality.

We provide a closed-form solution to the matrix nearness problem just described. Interestingly, the solution can be decomposed into three separate terms, whose intuition we provide next. Given an observed Slutsky matrix function,

• (a) the norm of its anti-symmetric or skew symmetric part measures the “size” of the violation of symmetry;

• (b) the norm of the smallest additive matrix that will make the symmetric part of the Slutsky matrix singular measures the “size” of the violation of singularity; and

• (c) the norm of the positive semidefinite part of the resulting corrected matrix measures the “size” of the violation of negative semidefiniteness.

Our main result shows that the “size” of bounded rationality, measured by the Slutsky matrix squared norm, is simply the sum of the squares of these three effects. In particular, following any observed behavior, we can classify the instances of bounded rationality as violations of the Ville axiom of revealed preference –VARP– if only symmetry fails, violations of homogeneity of degree zero or other money illusion phenomena if only singularity fails, violations of the weak axiom of revealed preference –WARP– by a symmetric consumer if only negative semidefiniteness fails,
or combinations thereof in more complex failures, by adding up the nonzero components of the norm.

The size of bounded rationality provided by the Slutsky norm depends on the units in which the consumption goods are expressed. It is therefore desirable to provide unit-independent measures, and we do so following the approach of modifying the Slutsky matrix by a weighting matrix. For example, one can translate the first norm into dollars, hence providing a monetary measure, or one can instead use a budget shares version that is unit-free.

The Slutsky matrix function is the key object in comparative statics analysis in consumer theory. It encodes all the information of local demand changes with respect to small Slutsky compensated price changes. We obtain comparative statics results for a boundedly rational consumer. Importantly, one can decompose the error in comparative statics from assuming a given form of rationality as the sum of two independent terms. The first is the behavioral error (measured by our Slutsky norm and its decomposition), and the second is a mis-specification error given the assumed parameterized rationality model. Further, unlike her rational counterpart, a boundedly rational consumer may exhibit path dependence of wealth compensations in price-equivalent paths or may violate the compensated law of demand, and we relate such phenomena to the different nonzero terms in our decomposition of the norm.

The rest of this paper is organized as follows. Section 2 presents the model. Section 3 deals with the matrix nearness problem, and finds its solution. Section 4 emphasizes its additive decomposition, and provides interpretations of the matrix nearness problem in terms of the axioms behind revealed preference. Section 5 presents weighted Slutsky norms. Section 6 goes over comparative statics. Section 7 presents several examples and applications of the result, including hyperbolic discounting and the sparse consumer model. Section 8 is a brief review of the literature, and Section 9 concludes. Some of the proofs of the more technical results are collected in an appendix.

2 The Model

Consider a demand function $x: Z \mapsto X$, where $Z \equiv P \times W$ is the compact space of price-wealth pairs $(p,w)$, $P \subseteq \mathbb{R}^L_{++}$, $W \subseteq \mathbb{R}^m_{++}$, and $X \equiv \mathbb{R}^L$ is the consumption set. This demand system is a generic function that maps price and wealth to consumption bundles.

Assume that $x(p,w)$ is continuously differentiable and satisfies Walras’ law: $p'x(p,w) = w$ for all $(p,w) \in Z$. Let $\mathcal{X} \subset C^1$ denote a set of functions that satisfy these characteristics. To be more explicit noting the dependence on the domain, define also $\mathcal{X}(Z) \subset C^1(Z)$, with $C^1(Z)$ denoting the complete metric space of vector-valued functions $f : Z \mapsto \mathbb{R}^L$, continuously differentiable, uniformly bounded with compact domain $Z \subset \mathbb{R}^L_{++}$, equipped with a norm.\footnote{For instance, the norm $\|f\|_{C^1} = \max(\|f_l\|_{C,1})_{l=1,...,L}$, with $\|f_l\|_{C,1} = \max(\|f_l\|_{\infty,1})$} where $f(z) = [f_1(z) \cdot \cdot \cdot f_L(z)]' \in \mathbb{R}^L$. We use also the notation for the norm $\|\cdot\|_{\infty,m} = \sup_{z \in Z}|g(z)|$ for $g : Z \mapsto \mathbb{R}^m$, for finite $m \geq 1$ and $|\cdot|$ the absolute value. By the set $C^1(Z)$ we mean the set of functions that have $C^1$ extensions to an open set containing the compact domain $Z$.
Definition 1. Define the distance of \( x \in X \) to the set of rational demands \( R \) by the “least” distance from an element to a set: 
\[
d(x, R) = \inf \{ d(x, x^r) \mid x^r \in R \}.
\]

We shall refer to this problem of trying to find the closest rational demand to a given demand as the “behavioral nearness” problem (Varian, 1990). Observe that the behavioral nearness problem at this level of generality presents several difficulties. First, the constraint set \( R(Z) \), i.e., the set of rational demand functions, is not convex. In addition, the Lagrangian depends not only upon \( x^r \) but also on its partial derivatives. The typical curse of dimensionality of calculus of variations applies here with full force, in the case of a large number of commodities. Indeed, the Euler-Lagrange equations in this case do not offer much information about the problem and give rise to a large partial differential equations system. Finally, to calculate analytically the solution to this program is computationally challenging.

Thus, instead of solving directly the “behavioral nearness” problem, our approach will rely on matrix spaces. Our next goal is to talk about Slutsky norms. Let \( M(Z) \) be the complete metric space of matrix-valued functions, \( F : Z \rightarrow \mathbb{R}^{L \times L} \), equipped with the inner product \( \langle F, G \rangle = \int_{z \in Z} \text{Tr}(F(z)'F(z))dz \). This vector space has a Frobenius norm \( ||F||^2 = \int_{z \in Z} \text{Tr}(F(z)'F(z))dz \).

Let us define the Slutsky substitution matrix function.

Definition 2. Let \( Z \subset P \times W \) be given, and denote by \( z = (p, w) \) an arbitrary price-wealth pair in \( Z \). Then the Slutsky matrix function \( S \in M(Z) \) is defined pointwise: 
\[
S(z) = D_p x(z) + D_w x(z)z^r \in \mathbb{R}^{L \times L},
\]

with entry \( s_{l,k}(p, w) = \frac{\partial x_l(p, w)}{\partial p_k} + \frac{\partial x_l(p, w)}{\partial w_k} x_k(p, w) \).

The Slutsky matrix function is well defined for all \( x \in C^1(Z) \). Restricted to the set of rational behaviors, the Slutsky matrix satisfies a number of regularity conditions. Specifically, when a matrix function \( S \in M(Z) \) is symmetric, negative semidefinite (NSD), and singular with \( p \) in its null space for all \( z \in Z \), we shall say that the matrix satisfies property \( R \), for short.

Definition 3. For any Slutsky matrix function \( S \in M(Z) \), let its Slutsky norm be defined as follows: 
\[
d(S) = \min \{ ||E|| : S - E \in M(Z) \text{ having property } R \}.
\]

The use of the minimum operator is justified. Indeed, it will be proved that the set of Slutsky matrix functions satisfying \( R \) is a closed and convex set. Then, under the metric induced by the Frobenius norm, the minimum will be attained in \( M(Z) \). We shall refer to the minimization problem implied in the Slutsky norm as the “matrix nearness” problem.

3 The Matrix Nearness Problem: Its Solution

In this section we provide the exact solution to the matrix nearness problem, which allows us to quantify the distance from rationality by measuring the size of the violations of the Slutsky matrix conditions.

3.1 Preliminaries on Matrices

We begin by reviewing some definitions.

---

2. We refer the reader to our expanded discussion in Aguiar and Serrano (2014), which includes the use of the “almost implies near” approach.

3. While one could use a different norm, we justify our choice of the Frobenius norm in the sequel.

4. Since \( Z \) is closed, we use the definition of differential of Graves (1956) that is defined not only in the interior but also on the accumulation points of \( Z \).
It will be useful to denote the three regularity conditions of any Slutsky matrix function with shorthands. We shall use \( \sigma \) for symmetry, \( \pi \) for singularity with \( p \) in its null space (\( p \)-singularity), and \( \nu \) for negative semidefiniteness.

Given any square matrix-valued function, \( S \in \mathcal{M}(Z) \), let \( S^{sym} = \frac{1}{2} [S + S'] \) be its symmetric part. With our new notation, if \( S \) is a Slutsky matrix function, \( S^{\sigma} = S^{sym} \). Equivalently, \( S^{\sigma} \) is the projection of the function \( S \) on the closed subspace of symmetric matrix-valued functions, using the inner product defined for \( \mathcal{M}(Z) \).

Every square matrix function \( S \in \mathcal{M}(Z) \) can be written as \( S(z) = S^{sym}(z) + S^{skew}(z) \) for \( z \in Z \), also written as \( S = S^{\sigma} + E^{\sigma} \), where \( S^{\sigma} = S^{sym} \) is its symmetric part and \( E^{\sigma} = S^{skew} = \frac{1}{2} [S - S'] \) is its anti-symmetric or skew-symmetric part (the orthogonal complement of the symmetric part).

Every symmetric matrix function \( S^{\sigma} \in \mathcal{M}(Z) \), in particular the symmetric part of a Slutsky matrix function \( S \), can be decomposed into the sum of a singular and not singular parts (with prices in the null space): see Claim 3.5. With our notation, the matrix function part that is singular with \( p \) in its null space will be denoted \( S^{\sigma, \pi} \), that is, \( S^{\sigma, \pi}(z)p = 0 \). Then \( S^{\sigma} = S^{\sigma, \pi} + E^{\sigma} \) where \( S^{\sigma, \pi} = PS^{\sigma}P \) and \( E^{\sigma} = S^{\sigma} - PS^{\sigma}P \) is its orthogonal complement with \( P = I - \frac{pp'}{pp'} \) a projection matrix.6

Any symmetric matrix-valued function \( S^{\sigma} \in \mathcal{M}(Z) \) and in particular any matrix function that is the singular in prices part \( S^{\sigma, \pi} \in \mathcal{M}(Z) \) of a Slutsky matrix can be pointwise decomposed into the sum of its positive semidefinite and negative semidefinite parts. Indeed, we can always write \( S^{\sigma, \pi}(z) = S^{\sigma, \pi}(z)_+ + S^{\sigma, \pi}(z)_- \), with \( S^{\sigma, \pi}(z)_+ S^{\sigma, \pi}(z)_- = 0 \) for all \( z \in Z \). Moreover, for any square matrix-valued function \( S \in \mathcal{M}(Z) \), its projection on the cone of NSD matrix-valued functions under the Frobenius norm is \( S^{\sigma, \pi}_{NSD} \).

In general, a square matrix function may not admit diagonalization. However, we know thanks to Kadison (1984) that every symmetric matrix-valued function in the set \( \mathcal{M}(Z) \) is diagonalizable. In particular, \( S^{\sigma, \pi} \) can be diagonalized: \( S^{\sigma, \pi}(z) = Q(z) \Lambda(z) Q(z)'^{'} \). Here, \( \Lambda(z) = \text{Diag} \{ \lambda(z) \} \), where \( \lambda(z) \in \mathcal{M}(Z) \), with \( \lambda_i : Z \mapsto \mathbb{R} \) a real-valued function with norm \( || \cdot ||_{1} \) (a norm in \( C^1(\mathbb{R}) \)), and \( \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_L \) with the order derived from the metric induced by the \( || \cdot ||_{1} \) norm.7 Let \( Q(z) = [q_1(z) \ldots q_L(z)] \), where \( Q \in \mathcal{M}(Z) \) and its columns \( q_i \in C^1(Z) \) are the eigenvector functions such that for \( l = 1, \ldots, L \): \( S^{\sigma, \pi}(z) q_l(z) = \lambda_l(z) q_l(z) \) for \( z \in Z \).

Any real-valued function can be written as \( \lambda(z) = \lambda(z)_+ + \lambda(z)_- \), with \( \lambda(z)_+ = \max \{ \lambda(z), 0 \} \) and \( \lambda(z)_- = \min \{ \lambda(z), 0 \} \). This decomposition allows us to write \( S^{\sigma, \pi, \nu}(z) = S^{\sigma, \pi}(z)_- = Q(z) \Lambda(z)_- Q(z)'^{'} \) for \( \Lambda(z)_- = \text{Diag} \{ \lambda(z)_- \} \), and \( \lambda(z)_- \) the negative part function for the \( \lambda_l(z)_- \) function. We can write also \( E^{\nu}(z) = S^{\sigma, \pi}(z)_+ = S^{\sigma, \pi}(z) - S^{\sigma, \pi}(z)_- \) for \( z \in Z \), or \( S^{\sigma, \pi}(z)_+ = Q(z) \Lambda(z)_+ Q(z)'^{'} \) with \( \Lambda(z)_+ \) defined analogously to \( \Lambda(z)_- \) as the orthogonal complement of \( S^{\sigma, \pi, \nu} \).

3.2 The Result

We are ready to state the solution to the matrix nearness problem:

---

5This means that we consider the space of symmetric matrix functions \( S^{\sigma} \) that satisfy \( p'S^{\sigma}(z)p = 0 \). This is a consequence of Walras’ law because \( p'S(z) = 0 \), although not necessarily \( S(z)p = 0 \).

6Because of Walras’ law when \( S \) is the Slutsky matrix of some demand function \( x \in \mathcal{L}(Z) \), \( \nu S(z) = 0 \) and \( S^{\sigma, \pi} = S^{\sigma} - E^{\sigma} \) and \( E^{\sigma}(z) = \frac{1}{pp'}[S^{\sigma}(z)pp' + pp'S^{\sigma}(z)] \).

7The order of the eigen-values is inessential to the results but it is convenient for the proofs. Observe that we can have many diagonal decompositions all of which will work for our purposes.
Theorem 1. Given a Slutsky matrix $S$, it has the unique decomposition $S = S^r + E$, where $S^r = S_{\sigma,\pi,\nu}$ is the solution to the matrix nearness problem and $E$ is the sum of the orthogonal complements of the symmetric, singularity in $p$ and NSD parts of $S$: $E = E^s + E^r + E^v$. Furthermore, the Slutsky norm of the solution to the matrix nearness problem can be decomposed as follows:

$$||E||^2 = ||E^s||^2 + ||E^r||^2 + ||E^v||^2.$$ 

We elaborate at length on the different components of this solution in the next section, right after the proof of the theorem.

Proof. We first establish that the matrix nearness problem has a solution, and that it is unique. This is done in Claim 1. Its proof is in the appendix.

Claim 1. The solution to the matrix nearness problem exists, and it is unique.

The rest of the proof of the theorem is done in two parts. Lemma 1 gives the solution imposing only the singularity with $p$ in its null space and symmetry restrictions. After that, Lemma 2 rewrites the problem slightly, and the solution is provided by adding the NSD restriction.

Lemma 1. The solution to $\min_A ||S - A||$ subject to $A(z)p = 0$, $A(z) = A(z)'$ for all $z \in Z$ is $S_{\sigma,r}$.

Proof. The Lagrangian for the subproblem with symmetry and singularity restrictions is:

$$\mathcal{L} = \int_{z \in Z} Tr([S(z) - A(z)]'[S(z) - A(z)])dz + \int_{z \in Z} \lambda(z)'A(z)pdz + \int_{z \in Z} vec(U(z))'vec[A(z) - A(z)'].$$

With $\lambda \in \mathbb{R}^L$ and $U \in \mathbb{R}^{L \times L}$. Using that the singularity restriction term is scalar ($\lambda S_{\sigma,r}(z)p \in \mathbb{R}$), as well as the identity $Tr(A'B) = vec(A)'vec(B)$ for all $A, B \in \mathcal{M}(Z)$, one can rewrite the Lagrangian as:

$$\mathcal{L} = \int_{z \in Z}[Tr([S(z) - A(z)]'[S(z) - A(z)]) + Tr(A(z)p\lambda(z)') + Tr(U(z)'[A(z) - A(z)'])]dz$$

Using the linearity of the trace, and the fact that this calculus of variations problem does not depend on $z$ or on the derivatives of the solution $S_{\sigma,r}$, the pointwise first-order necessary and sufficient conditions in this convex problem (Euler Lagrange Equation) is:

$$S(z) - A(z) + U(z) - U(z)' + \lambda(z)p' = 0.$$ 

Solve for $A(z)$:

$$A(z) = S(z) + U(z) - U(z)' + \lambda(z)p'.$$

By the symmetry restriction,

$$2[U(z) - U(z)'] = S(z)' - S(z) + p\lambda(z)' - \lambda(z)p'.$$

Replace back in the expression for $A(z)$:

$$A(z) = S(z) + \frac{1}{2}[S(z)' - S(z) + p\lambda(z)' - \lambda(z)p'] + \lambda(z)p'.$$

We obtain the expression:

$$A(z) = S^s(z) + \lambda(z)p' + \frac{1}{2}[p\lambda(z)' - \lambda(z)p'].$$

Now we observe that we can write,

$$A(z) = S^s(z) + W^s(z)\text{ with } W(z) = \lambda(z)p' \text{ and } W^s = \frac{1}{2}[p\lambda(z)' + \lambda(z)p'],$$

the symmetric part of $W(z)$.

By the restriction of singularity in prices $A(z)p = 0$ which implies $W^s(z)p = -S^s(z)p$. 

\[^{8}\text{The vec}(A)\text{ symbol stands for the vectorization of a matrix A of dimension L x L in a vector a = vec(A) of dimension } L^2 \text{ where the columns of } A \text{ are stacked to form } a. \text{ Observe that the symmetry restriction can be expressed in a sigma notation (entry-wise) but this matrix algebra notation helps us to make more clear the use of the trace operator in the objective function.}\]
To complete the proof we let \( r(z) = -S\sigma(z)p \) be the feedback error (of non-singularity in prices). Then we can write the matrix function \( W\sigma(z) = \frac{r(z)p' + pr(z)}{p'p} - \frac{[r(z)p'pp']}{(p'p)^2} \) that is the symmetric part of the outer product of the feedback error and the price vector. Replacing \( r(z) \) by its definition, we get that \( W\sigma = -E\sigma \):

\[
E\sigma(z) = \frac{1}{p'p}[S\sigma(z)p'p + pp'S\sigma(z) - \frac{1}{p'p}pp'S\sigma(z)pp'],
\]

then \( E\sigma = S\sigma - PS\sigma P \) with \( P = I - pp' \) and by Walras’ law it simplifies to:

\[
E\sigma(z) = \frac{1}{p'p}[S\sigma(z)p'p + pp'S\sigma(z)],
\]

this expression along with the implied multipliers \( \lambda \) and \( U \) satisfies all the first-order conditions of the problem. Since we can use arguments identical to those in Claim 1 –only not imposing \( \text{NSD}^- \), we know that the solution is unique. Hence, this expression describes the solution of the problem. Since we can use arguments identical to those in Claim 1 –only not

The proof is complete.

If \( S^{\sigma,\pi} \) were \( \text{NSD}^- \), then we would be done, since it would have property \( \mathfrak{R} \) and minimize \( ||E||^2 \). Otherwise, the general solution is provided after the following lemma, which rewrites the problem slightly.

**Lemma 2.** The matrix nearness problem can be rewritten as \( \min_{A \in M(Z)} ||S^{\sigma,\pi} - A||^2 \) subject to \( A \in \mathfrak{R} \).

**Proof.** Recall the matrix nearness problem: \( \min_{A \in M(Z)} ||S - A||^2 \) subject to \( A \) satisfying singularity, symmetry, and \( \text{NSD}^- \). This is equivalent, by manipulating the objective function, to: \( \min_{A \in M(Z)} ||E\sigma + E\pi + S^{\sigma,\pi} - A||^2 \). Writing out the norm as a function of the traces, and using the fact that \( E\sigma \) is skew symmetric, while the rest of the expression is symmetric, we get that this amounts to writing: \( \min_{A \in M(Z)} ||E\sigma||^2 + ||E\pi + S^{\sigma,\pi} - A||^2 \) subject to \( A \) having property \( \mathfrak{R} \). This is in turn equivalent to: \( \min_{A \in M(Z)} ||E\sigma||^2 + ||E\pi||^2 + 2\langle E\sigma, [S^{\sigma,\pi} - A] \rangle + ||S^{\sigma,\pi} - A||^2 \) subject to \( A \) having property \( \mathfrak{R} \). Then, exploiting the fact that \( E\sigma \) and \( S_L = S^{\sigma,\pi} - A \) are orthogonal (as proved in Claim 3), we obtain that the problem is equivalent to \( \min_{A \in M(Z)} ||E\sigma||^2 + ||E\pi||^2 + ||S^{\sigma,\pi} - A||^2 \) solving the program \( \min_{A \in M(Z)} ||E||^2 \) subject to \( A \) having property \( \mathfrak{R} \) is equivalent to solving \( \min_{A \in M(Z)} ||S^{\sigma,\pi} - A||^2 \) subject to the same constraints.

Now, the best \( \text{NSD}^- \) matrix approximation of the symmetric valued function \( S^{\sigma,\pi} \) is \( S' = S'^{\sigma,\pi} \). Then, the candidate solution to our problem is \( A(z) = S'(z) \) for all \( z \in Z \). Notice that \( S'(z) \) is symmetric and singular with \( p \) in its null space by construction. Indeed, recall that \( S^{\sigma,\pi}(z) = Q(z)\Lambda(z)Q(z)' \) and \( S'(z) = S'^{\sigma,\pi}(z) = Q(z)\Lambda(z)Q(z)' \). Then it follows that \( S'(z) \) is symmetric for \( z \in Z \). Moreover, by definition \( \lambda_l(z)_- = \min(0, \lambda_l(z)) \) for \( l = 1, \ldots, L \). Since \( S^{\sigma,\pi}(z)p = 0 \), it follows that \( \lambda_L(z) = 0 \) is the eigenvalue function associated with the \( q_L(z) = p \) eigenvector. Then we have that \( \lambda_L(z)_- = 0 \) is also associated to the eigenvector \( p \), and we can conclude that \( S'(z)p = 0 \).

As just argued, \( S'(z) \) has property \( \mathfrak{R} \), i.e., \( S'(z) \) is in the constraint set of the matrix nearness problem. We conclude that it is its solution.

The orthogonalities of the relevant subspaces of matrix functions, as made evident in previous steps of the proof, are responsible for the additive decomposition of the norm in the second part

---

9In this step we follow a similar strategy of proof to Dennis & Schnabel (1979); Higham (1989).
of the statement. The Frobenius norm and its weighted variants are essentially the only ones that correspond to an inner product on the space of matrix valued functions. For this reason, they are the only ones that lead to an additive decomposition of the three components of the residual matrix $E$.

Furthermore, we close this section by proving that $S^r$ is a continuous matrix-valued function. Indeed, the following is a property of the solution to our problem:

**Claim 2.** $S^r$ is continuous.

**Proof.** Note that $S^r$ is the result of three projections on closed subspaces applied to the convex set of constraints. Such projections are continuous mappings under the conditions that we have imposed, and then $S^r$ is continuous by construction in all $z \in Z$. 

## 4 Decomposition of the Matrix Nearness Solution

The importance of Theorem (1) is to provide a precise quantification of the size of the departures from rationality by a given behavior, as well as a revealing decomposition thereof.\(^{10}\)

We should think of the three terms in the decomposition of $||E||_2^2$ as the size of the violation of symmetry, the size of the violation of singularity, and the size of the violation of negative semidefiniteness of a given Slutsky matrix, respectively (this idea is illustrated in figure 1). The three terms are the antisymmetric part of the Slutsky matrix function, the correcting matrix function needed to make the symmetric part of the Slutsky matrix function $p$-singular, and the PSD part of the resulting corrected matrix function. Note that if one is considering a rational consumer, the three terms are zero. Indeed, if $S(z)$ satisfies property $\Re$ for all $z \in Z$, $S(z) = S^s(z)$ and $E^s(z) = 0$, $S^{s,\pi}(z) = S^s(z)$ and $E^\pi(z) = 0$, and $S^r(z) = S^{s,\pi,\nu}(z) = S^{s,\pi}(z)$ and $E^\nu(z) = 0$. If exactly two out of the three terms are zero, the nonzero term allows us roughly to quantify violations of the Ville axiom of revealed preference –VARP–, violations of homogeneity of degree 0, or violations of the compensated law of demand (the latter being equivalent to the weak axiom of revealed preference –WARP–), respectively. We elaborate on these connections with the axioms of consumer theory in Subsection 4.1 below.

The violations of the property $\Re$ have traditionally been treated separately. For instance, Russell (1997), using a different approach (outer calculus), deals with violations of the symmetry condition only. In this case, $||E||_2^2 = ||E^s||_2^2$.

Another application of our result connects with Jerison and Jerison (1993), who study the case of violations of symmetry and negative semidefiniteness independently. They prove that the maximum eigenvalue of $S^s(z)$ can be used to bound $||e||_C$ locally when NSD is violated and $E^s(z)$ can be used to bound $||e||_C$ when symmetry is violated. Indeed, this is consistent with our solution. In this case $||E||_2^2 = ||E^s||_2^2 + ||S^s||_2^2$, where $max\{\tilde{\lambda}_l(z)_+\}_{l=1,\ldots,L} \leq ||S^s||_2^2 \leq d \cdot max\{\tilde{\lambda}_l(z)_+\}_{l=1,\ldots,L}$, with $d = \text{Rank}(S^s(z)_+)$ (by the norm equivalence of the maximum eigenvalue and the Frobenius norm).

### 4.1 Connecting with Axioms of Revealed Preference

Since Hurwicz and Uzawa (1971) it is known that for the class of continuous differentiable functions a demand that satisfies Walras’ law can be rationalized if and only if its Slutsky matrix

---

\(^{10}\)The decomposition, a consequence of orthogonalities, can be seen as a generalized Pythagorean theorem.
matrix function is symmetric ($\sigma$) and NSD ($\nu$). A corollary of this result is that its Slutsky matrix function is singular in prices ($\pi$) (John, 1995). These Slutsky regularity conditions ($\sigma, \pi, \nu$) are linked to behavioral demand axioms. In fact, Houthakker (1950) proved that, for the class of continuous demand functions, the strong axiom of revealed preference –SARP– implies that a demand can be rationalized. In this sense, SARP is indeed strong. A weaker axiom implies only the Slutsky matrix function symmetry condition: the Ville axiom of revealed preference –VARP– is equivalent to the symmetry condition and therefore to integrability of the demand system (Hurwicz & Richter, 1979). VARP postulates the nonexistence of a Ville cycle in the income path of a demand function. The weak axiom of revealed preference –WARP– implies that the Slutsky matrix function is NSD and singular in prices. Furthermore, the NSD and singularity in prices are equivalent to a weak version of WARP (Kihlstrom et al., 1976). VARP is a differential axiom and does not imply SARP or WARP (Hurwicz & Richter, 1979). It is also known that WARP does not imply VARP or SARP for dimensions greater than two.

A continuously differentiable demand function is said to be rationalizable if and only if it fulfills SARP. However, we can also combine the other weaker axioms to have the same result while drawing connections to the additive components of our Slutsky norm. For example, VARP and WARP imply that a demand function is rationalizable, and so do the Wald Axiom, homogeneity of degree zero and VARP.

Before presenting the main remark of this subsection, for completeness, it is useful to posit the axioms that we employ.

Assumed throughout, the first basic condition (price is a Slutsky matrix left eigenvector) is given by Walras’ law.

**Axiom 1.** (Walras’ law) $p'x(p, w) = w$.

We have that if $x \in X(Z)$ satisfies Walras’ law then its Slutsky matrix function $S \in \mathcal{M}(Z)$ has the following property: $p'S(z) = 0$ for $z \in Z$. The converse holds when $Z$ is path connected.

The condition that the price $p$ is the Slutsky matrix right eigenvector or singularity in prices is given by “no money illusion”.

Figure 1: Decomposition of the Slutsky matrix norm. The size of the violations of rationality is the sum of the violations of symmetry, singularity in prices and negative semidefiniteness.
Axiom 2. (Homogeneity of degree zero -HD0-) \( x(\alpha z) = x(z) \) for all \( z \in Z \) and \( \alpha > 0 \).

It is easy to prove that if \( x \in X(Z) \) satisfies Walras’ law and HD0 then \( S(z)p = 0 \) for \( z \in Z \). More generally, \( p \) is an eigenvector associated with the null eigenvalue of \( S \) if and only if \( x \) satisfies Walras’ law and HD0.

The symmetry of the Slutsky matrix function is given by VARP.\(^{11}\) To state this axiom we need to define an income path as \( w : [0, b] \to W \) and a price path \( p : [0, b] \to P \). Let \( (w(t), p(t)) \) be a piecewise continuously differentiable path in \( Z \). Jerison and Jerison (1992) define a rising real income situation whenever \( \left( \frac{\partial w}{\partial t}(t), \frac{\partial p}{\partial t}(t) \right) \) exists, leading to \( \frac{\partial w}{\partial t}(t) > \frac{\partial p}{\partial t}(t)'x(p(t), w(t)) \). A Ville cycle is a path such that: (i) \( (w(0), p(0)) = (w(b), p(b)) \); and (ii) \( \frac{\partial w}{\partial t}(t) > \frac{\partial p}{\partial t}(t)'x(p(t), w(t)) \) for \( t \in [0, b] \).

Axiom 3. (Ville axiom of revealed preference -VARP-) There are no Ville cycles.

Hurwicz and Richter (1979) proved that \( x \in X(Z) \) satisfies VARP if and only if \( S \) is symmetric and \( x \) satisfies Walras’ law.

The negative semidefiniteness condition of the Slutsky matrix is given by the Wald axiom. The Wald axiom itself is imposed on the conditional demand for a given level of wealth. Following John (1995) a demand function that satisfies Walras’ law is said to fulfill the Wald axiom. Moreover \( x \in X(Z) \) satisfies Walras’ law and HD0 then \( \partial w \partial t ) = (w(b), p(b)) \); and (ii) \( \frac{\partial w}{\partial t}(t) > \frac{\partial p}{\partial t}(t)'x(p(t), w(t)) \) for \( t \in [0, b] \).

Axiom 4. (Wald axiom) \( x \in X(Z) \) is such that for every \( w \in W \) and for all \( p \) and \( \overline{p}, \; \overline{p}'x^w(p) \leq w \implies p'x^w(\overline{p}) \geq w \).

The Wald axiom implies that \( S \leq 0 \) (John, 1995) for demands that satisfy Walras’ law.

The Slutsky singularity in prices and the NSD conditions are equivalent to the following version of WARP.

Axiom 5. (Weak axiom of revealed preference -WARP-) If for any \( z = (p, w) \) \( \overline{\pi} = (\overline{p}, \overline{w}) \): \( \overline{p}'x(p, w) \leq \overline{w} \implies p'x(\overline{p}, \overline{w}) \geq w \).

This is the weak version of WARP, as in Kihlstrom et al. (1976). We follow John (1995), who proves that for continuously differentiable demands (that satisfy Walras’ law) WARP is equivalent to the Wald Axiom and HD0. Kihlstrom et al. (1976) and John (1995) prove that \( x \in X(Z) \) satisfies WARP if and only if \( S \) is NSD and \( S(z)p = 0 \).

We are ready to summarize the main point of this subsection in the following remarks.

Remark 1. The Slutsky matrix nearness norm \( ||E'||^2 \) will be equal to zero if and only if \( x \) satisfies VARP, homogeneity of degree zero, and the Wald Axiom (in addition to Walras’ law).

Moreover

(i) \( ||E'||^2 = ||E'^w||^2 > 0 \) only if VARP fails. In this case, either HD0 and the Wald Axiom hold, or Warp holds.

(ii) \( ||E'||^2 = ||E'^w||^2 > 0 \) only if Walras’ law fails. In addition, HD0 may also fail; VARP and the Wald axiom hold.

\(^{11}\) We present the axiomatization due to Ville as reinterpreted by Hurwicz and Richter (1979) and Jerison and Jerison (1992). There are alternative discrete axioms due to Jerison and Jerison (1996), that also do the job and are potentially testable.
(iii) $|E|^2 = |E^\nu|^2 > 0$ only if the Wald axiom fails. In this case, VARP holds, and either WARP fails or HD0 holds.

In addition, under additional technical regularity conditions on the domain and on the demand (that are not necessary but sufficient and capture the current state of the literature), we can state even stronger connections. Also note that our results go through with an open price wealth region $Z$, provided it is Lipchitz continuous.

**Remark 2.** Suppose $Z$ is open, convex, path connected and the demand $x \in \mathcal{X}(Z)$ is smooth, bounded and Lipchitz continuous (at least with respect to wealth) and satisfies Walras’ law.

(i) Weak WARP holds if and only if $|E^\pi| = 0$ and $|E^\nu| = 0$ (John, 1995; Kihlstrom et al., 1976).

(ii) The Ville Axiom holds if and only if $|E^\sigma| = 0$ (Hurwicz & Richter, 1979; Hurwicz & Uzawa, 1971).

(iii) No money illusion (homogeneity of degree zero) holds if and only if $|E^\pi| = 0$.\(^{12}\)

(iv) The Wald axiom holds if and only if $|E^\nu| = 0$ (John, 2000, 1995).\(^{13}\)

We do not impose the additional sufficient conditions on $x \in \mathcal{X}(Z)$ described in the previous remark for the general results in the rest of the paper because the necessity part of the above statements hold without them. Such necessity parts of the statements suffice to implement an econometric test of the revealed preference axioms due to the fact that if any of the parts of the Slutsky matrix norm are non zero the corresponding axiom cannot hold for the given demand.

## 5 Weighted Slutsky Norms

The norm of bounded rationality that we have built so far is an absolute measure. Therefore, for a specific consumer, this distance quantifies by how far that individual’s behavior is from being rational. Furthermore, one also can compute how far two or more consumers within a certain class are from rationality, and induce an order of who is closer in behavior to a rational consumer. However, we are limited to the case where the setting of the decision making process is fixed in the sense that the decision problem facing each of the individuals is presented in the same way. This implies that the measure is unit dependent, being stated in the same units (the units in which the consumption goods are expressed). A second undesirable feature, is that the Frobenius norm depends on the area of the arbitrary compact region $Z$ that has positive Lebesgue measure. This section addresses these issues.

Let $W$, $U$ be square and PSD (i.e., with symmetric part also PSD) matrix functions of weights such that it is non-null and $W(z), U(z) \in \mathbb{R}^{L \times L}$. This matrix pair is meant to encode all suitable normalizations and priorities of the modeller. Let $Z \subseteq P \times W$ be sigma-measurable and have positive measure $\mu$ with $\mu$ continuous, strictly positive function, such that $S \in \mathcal{M}(Z)$ is square integrable. Let $||\cdot||_M$ be the Frobenius matrix norm where $M = \mathbb{R}^{L \times L}$. The weighted semi-norm $||S||_{WU,\mu}$ is defined as follows:

$$||S||_{WU,\mu}^2 = \int_{z \in Z} ||W(z)S(z)U(z)||_M^2 \mu(z)dz < \infty.$$\(^{12}\)

\(^{12}\)This result is trivial under path connectedness because $E^\pi = 0$ means that $D_p x(p,w)p + D_w x(p,w)w = 0$.

\(^{13}\)This result is trivial to establish by noticing that the related Slutsky matrix function $S$ is nowhere zero due to smoothness, and the fact that $E^\nu = 0$ guarantees that $D_p x(p,\pi)$ is NSD for any $\pi$. 

11
When $\mu$ is continuous and strictly positive, the spaces $L^2(Z, B(Z); M(Z))$ and the weighted $L^2(Z, B(Z); M(Z); \mu)$ are unitary equivalent or unitary isomorphic. This means that our orthogonality relations under the inner products defined by the Lebesgue measure are preserved by inner product weighted by $\mu$. Hence, suitable variants of Theorem 1 can be obtained for these weighted norms.

**Remark 3.** With a $W = U$ that is positive definite and symmetric, the solution to the matrix nearness problem, $S^{*r} = \arg\min_{A \in M(Z)} ||S - A||_{W, \mu}$ subject to $A$ having the property $R$, is $S^{*r} = W^{-1}(WS^{*r}W)W^{-1}$ and the quadratic norm is given by:

$$||E^*||_{W, \mu}^2 = ||E^*||_{W, \mu}^2 + ||E^{*\pi}||_{W, \mu}^2 + ||E^{*\nu}||_{W, \mu}^2,$$

where $E^{*\pi} = E^{\sigma}$, $E^{*\pi} = E^{\tau}$, and $E^{*\nu} = W^{-1}(WS^{*\pi}W)W^{-1}$.

Observe that only the part concerning the NSD matrix function of violations ($\nu$ property) changes due to the introduction of the weighting matrix function $W$.

**Remark 4.** Using the identity matrix as $W = U$ and the multivariate uniform as $\mu$, one obtains the closest results to Theorem 1. However, the resulting Slutsky norm is independent of the area of the domain $Z$.

**Remark 5.** If $W = \text{Diag}(p)$ and $U = W$, the weighted Slutsky matrix is expressed in (square) dollar amounts. The entries of this normalized matrix express changes in expenditure on each good following a percent change in each price.

**Remark 6.** If $W = \frac{1}{\sqrt{u}}\text{Diag}(p)$ and $U = W$, the weighted Slutsky matrix is unit-free. The entries of this normalized matrix express changes in expenditure shares on each good following a percent change in each price.

**Remark 7.** When we have a degenerate $\mu$ such that $\mu(\zeta) = 1$ (say an equilibrium price and a given exogenous wealth) and $W, U$ are the identity, $||S||_{\mu} = ||S(\zeta)||_M$ and we can apply the local results of Jerison and Jerison (1992).

The standard Frobenius norm we use is chosen when one gives equal weight to each entry of the Slutsky matrix, thus avoiding that a specific price change on a specific commodity receive greater weight, but the modeller may adopt such choices at will and use arbitrary weighting matrices as a function of her interests. Finally, the integral in the norm embodies an expectation operator, which can be justified with Savage-type ideas of independence across consumption data points.

### 6 The Comparative Statics of Bounded Rationality

An approach based on the Slutsky matrix cannot be silent on comparative statics analysis. This section goes over a number of results that shed light on the meaning of the decomposition found in Theorem 1 and that connect with the behavioral nearness problem in the space of demand functions. The first subsection below focuses on mis-specification errors, offering another result based on orthogonalities. The results in the second and third subsections relate errors in comparative statics predictions to the error terms in our Slutsky norm decomposition.
6.1 Connecting with Demand: Mis-Specification Errors

We have solved the matrix nearness problem on the basis of the Slutsky regularity conditions. Now, in order to connect with demand, the exercise is one of integrating from the first-order derivatives of the Slutsky matrix terms. In such an integration step, a constant of integration shows up, which we would like to envision as a “residual” or “mis-specification error.” That is, starting from our observed Slutsky matrix function $S(x)$, we find thanks to Theorem 1 the nearest matrix function $S^*$ satisfying all the regularity properties. Now suppose that we specify $x_1$ and $x_2$ as two demand functions with which we wish to explain the consumer’s behavior (for example, $x_1$ might be a Cobb-Douglas demand properly estimated, whereas $x_2$ might be a different estimate, perhaps allowing the entire class of CES demands).

As will be shown, the error in comparative statics analysis from using each of those specifications is, respectively, $||S(x) - S(x_1)||^2 = ||S^* - S(x_1)||^2 + ||E||^2$ and $||S(x) - S(x_2)||^2 = ||S^* - S(x_2)||^2 + ||E||^2$. This means that $||E||^2 = ||S(x) - S^*||^2$ found in Theorem 1—the behavioral error—does not change with $x_1$ or $x_2$, this term is invariant to the choice of the parametrized rational approximation. The other term is the mis-specification error, as a function of the parametrized class considered. This is a useful separation. We establish it in our next result.

Let $\mathcal{R}(Z) \subset \mathcal{R}(Z)$ be an arbitrary closed and bounded subset of rational demand functions in $\mathcal{X}(Z)$. In particular, we are interested in a symmetric error, where we penalize any mistake symmetrically which corresponds to the Frobenius matrix norm:

$$\min_{x' \in \mathcal{R}(Z)} ||S(x') - S(x)||^2.$$

Our next result follows:

**Proposition 1.** For any solution $x^{r*} \in \arg\min_{x' \in \mathcal{R}(Z)} ||S(x') - S(x)||^2$, it follows that $||S(x^{r*}) - S(x)||^2 = ||S(x^{r*}) - S^*||^2 + ||E||^2$, where $||S(x^{r*}) - S^*||^2$ is a mis-specification error and $||E||^2$ is the Slutsky matrix norm.

**Proof.** The proof has three steps: First, we use the decomposition of a matrix into its symmetric and antisymmetric parts: $||S(x) - S(x')||^2 = ||S - S(x')||^2 = ||S^\sigma + E^\sigma - S(x')||^2 = ||S^\sigma - S(x')||^2 + 2(E^\sigma, S^\sigma - S(x')) + ||E^\sigma||^2,$

where $S^\sigma - S(x')$ is a symmetric matrix function and $E^\sigma$ is a skew-symmetric matrix function. Then $(E^\sigma, S^\sigma - S(x')) = 0$. Then $||S(x) - S(x')||^2 = ||S^\sigma - S(x')||^2 + ||E^\sigma||^2$.

Second, we take the p-singularity projection: $||S^\sigma - S(x')||^2 = ||S^\sigma,\pi + E^\sigma - S(x')||^2 = ||S^\sigma,\pi - S(x')||^2 + 2(E^\sigma, S^\sigma - S(x')) + ||E^\sigma||^2$. Since $E^\sigma$ is orthogonal to the matrix functions that are symmetric and have $p$ as its eigenvector associated with the null eigenvalue (claim 3), by definition $S(x'(p, w))p = 0$ and $S(x')$ is symmetric. Then: $(E^\sigma, S^\sigma,\pi - S(x')) = (E^\sigma, S^\sigma,\pi) - (E^\sigma, S(x')) = 0$. Therefore, $||S(x) - S(x')||^2 = ||S^\sigma,\pi - S(x')||^2 + ||E^\sigma||^2 + ||E^\pi||^2$.

Finally, we deal with the NSD part: $||S^\sigma,\pi,\nu - S(x')||^2 = ||S^\sigma,\pi,\nu + E^\pi - S(x')||^2 = ||S^\sigma,\pi,\nu - S(x')||^2 + 2(E^\nu, S^\sigma,\pi,\nu - S(x')) + ||E^\nu||^2$. Using the vector space structure of the space of demand functions, we can write $e = x - x'$. Then $S(x) = S(x') + G$, where $G = D_p e + D_w [xe'] + D_w [xe']$. It follows that $(E^\nu, S^\sigma,\pi,\nu - S(x')) = (E^\nu, C^\sigma,\pi,\nu) = 0$, where the right hand side follows from the orthogonality of $E^\nu$ and $C^\sigma,\pi,\nu$. This implies: $||S(x) - S(x')||^2 = ||S^\sigma - S(x')||^2 + ||E^\pi||^2$. □

14Lacking compactness, replace the min with the inf operator in the problem below. In fact, $\mathcal{R}(Z)$ is not compact but the result goes through.
In the same spirit, we close this subsection with a local result. Around a given reference price \( \bar{p} \), if one searches for a minimizer of error in comparative statics with respect to \( S^r \) among all twice continuously differentiable functions with bounded Hessians, satisfying Walras’ law and defined on an open ball around \( \bar{p} \), the result is a rational demand function. Note that the space of demands that satisfies Walras’ law is convex.

Define the local problem. Fix a reference price \( \bar{p} \). We can compute the Slutsky matrix of \( x \) at \( \bar{p} \), \( S(\bar{p}, w) \) where we allow \( w \) to vary. We can compute \( S^r(\bar{p}, w) \) using the same techniques developed for the general case. Abusing notation we keep referring to \( S(\bar{p}, \cdot) \) as \( S \) and \( S^r(\cdot, \cdot) \) as \( S^r \). The errors in comparative statics made due to assuming (perhaps incorrectly) rationality, around an open ball centered at \( \bar{p} \) correspond to the problem:

\[
\min_{y \in \mathcal{X}(\bar{p})} \| S(y) - S^r \|^2
\]

where \( \mathcal{X}(\bar{p}) \) is the set of demands that fulfill Walras’ law that are defined in an open ball around \( \bar{p} \), and that are twice continuously differentiable with a bounded Hessian. Let \( \mathcal{R}(\bar{p}) \subseteq \mathcal{X}(\bar{p}) \) be the subset of rational demands with the same properties. This is a local errors problem.

**Proposition 2.** If \( y^* \in \arg\min_{y \in \mathcal{X}(\bar{p})} \| S(y) - S^r \|^2 \), then \( y^* \in \mathcal{R}(\bar{p}) \).

**Proof.** The proof is organized in the following series of lemmata. We start with a definition due to Jerison and Jerison (1992; 1993).

Let \( (\hat{p}, \hat{w}) \in \mathbb{R}_{++}^{k+1} \), and let a price path be defined \( p(t) = t\bar{p} + (1-t)\hat{p} \), where \( \bar{p} \in \mathbb{R}_{++}^{k+1} \) is fixed. Then we study the Slutsky wealth compensation problem: \( \frac{\partial w(p,\hat{w},\bar{p},t)}{\partial t} = \frac{\partial p(t)}{\partial t} \) with initial condition \( w(\hat{p}, \hat{w}, \bar{p}, 0) = \hat{w} \). The unique solution to this problem is \( w^*(\hat{p}, \hat{w}, \bar{p}, \cdot) \).

The following definition of approximate local indirect utility is borrowed from Jerison and Jerison (1992; 1993). The indirect local utility at \((\hat{p}, \hat{w})\) is defined by \( v^*(\hat{p}, \hat{w}, \bar{p}, 1) \). Jerison and Jerison (1993) prove that \( w^* \) is continuously differentiable when \( x \) is continuously differentiable, and it is known that continuous functions can be approximated arbitrarily closely (in the sup norm) by smooth functions (Azagra & Boiso, 2004), thus we assume that \( w^* \) is smooth.

**Lemma 3.** [Jerison and Jerison (1993)] For every demand \( x \in \mathcal{X}(\bar{p}) \) that is twice continuously differentiable there exists \( x^* \in \mathcal{X}(\bar{p}) \) such that \( S(x^*)(\bar{p}, w) = S^r(\bar{p}, w) \). Moreover, \( x^*(p, w) = -\sum_{\sigma \in E} \frac{\partial w^*(p, w)}{\partial \sigma} \).

**Proof.** Remark 1 of [Jerison and Jerison (1993)].

**Lemma 4.** If \( y \in \arg\min_{y \in \mathcal{X}(\bar{p})} \| S(y) - S^r \|^2 \) then \( S(y) \) is symmetric. Moreover, if Walras’ law holds and \( S^r: \mathcal{R} \) there is a \( y^* \in \arg\min_{y \in \mathcal{X}(\bar{p})} \| S(y) - S^r \|^2 \) such that \( y^* = x^* \).

**Proof.** Take \( y^* \in \arg\min_{y \in \mathcal{X}(\bar{p})} \| S(y) - S^r \|^2 \) and assume that \( S(y^*) \) is not symmetric. Then let \( S(y^*) = S^u \). Now we can compute \( \| S(y^*) - S^r \|^2 = \| S^u - S^r \|^2 \). Observe that:

\[
\| S^u - S^r \|^2 = \| S^u - S^r \|^2 + 2\langle E^u, S^u - S^r \rangle + \| E^u \|^2.
\]

By the fact that \( S^u - S^r \) is symmetric, we have:

\[
\| S^u - S^r \|^2 = 2\| S^u - S^r \|^2 + \| E^u \|^2.
\]

By the lemma (3) there exists \( y^* \in \mathcal{X}(\bar{p}) \) for all \( y^* \) such that \( \| S(y^*) - S^r \|^2 = \| S^u - S^r \|^2 \). This means that \( \| S(y^*) - S^r \|^2 > \| S(y^*) - S^r \|^2 \) which is a contradiction to \( y^* \in \arg\min_{y \in \mathcal{X}(\bar{p})} \| S(y) - S^r \|^2 \). We conclude that \( S(y^*) \) is symmetric.
The moreover part of the statement follows from the existence of \( x^\sigma \) such that \( S(x^\sigma) = S^x,\sigma \) locally. By Walras’ law and symmetry, \( S^x,\sigma(p,w)p = 0 \) thus \( x^\sigma \) is HD0 around \( (p,w) \), this means that locally \( S(x^\sigma) = S^x,\sigma \). Thus \( ||S(x^\sigma) - S^r||^2 = 0 \) when \( S^x,\sigma = S^r \) around \( (p,w) \), thus making \( x^\sigma \) a solution.

\[ \square \]

Lemma 5. \( {\text{CCD}} \) Convex concave decomposition. Any scalar twice continuously differentiable function with a bounded Hessian can be decomposed in the sum of a concave and convex function parts.

\[ \text{Proof. Yuille \& Rangarajan (2003)} \]

Our task now is to define a local expenditure function. For a fixed \( w \), let \( e^\sigma(p,w) \) be defined implicitly as \( v^\sigma(p,e^\sigma(p,w)) = u \). The \( e^\sigma \) is not necessarily concave in \( p \). The function \( e^\sigma \) is unique and continuous when the indirect utility is monotone in wealth.

Now notice that, by construction, the integrable demand \( x^\sigma \) can be written as \( x^\sigma(p,e^\sigma(p,w)) = \nabla_p e^\sigma(p,u) \) because by differentiating \( v^\sigma(p,e^\sigma(p,w)) = u \) with respect to prices we obtain:

\[
\nabla_p e^\sigma(p,u) = -\frac{\nabla_x v^\sigma(p,e^\sigma(p,w))}{\nabla_v v^\sigma(p,e^\sigma(p,w))}.
\]

In turn, the Slutsky matrix of a “integrable” demand \( x^\sigma \) can be written as:

\[
S(x^\sigma)(p,e^\sigma(p,w)) = H e^\sigma(p,u),
\]

where \( H \) stands for the Hessian operator.

Lemma 6. If \( y \in \arg\min_{y \in X(p)} ||S(y) - S^r||^2 \), then \( S(y) \) is NSD.

\[ \text{Proof. By lemma (3) we know that } y^* = y^\sigma \text{ or equivalently } S(y^*) \text{ is symmetric. Also we know that there is a local quasi-expenditure function } e^\sigma(p,v(p,w)) = w \text{ for a fixed } w \text{ such that } S(y^*)(p,e^\sigma(p,w)) = H e(p^\sigma, u). \text{ Now take the matrix function } S^r \text{ and evaluate it at the same local quasi-expenditure function and obtain the matrix function: } H^{x,r}(p,u) = S^r(p,v^\sigma(p,w)). \]

Next, decompose the quasi-expenditure function in a concave part \( e^\sigma_{\text{cave}} = e^\sigma - r \) with \( r \) a convex function and a convex part \( e^\sigma_{\text{ex}} = r \). We focus on the local \( \overline{p} \) on quadratic expenditure functions (this is without loss of generality locally due to the Taylor expansion under twice continuous differentiability). Then we can ensure that we can find the minimal decomposition in the sense of Frobenius such that \( H e^\sigma = H e^\sigma_{\text{cave}} + H e^\sigma_{\text{ex}} \) such that \( \langle H e^\sigma_{\text{cave}}, H e^\sigma_{\text{cave}} \rangle = 0 \) and \( H e^\sigma_{\text{cave}} \) is NSD and \( H e^\sigma_{\text{ex}} \) is PSD (because locally we can think of the allowable functions as approximately quadratic forms). We study the equivalent problem \( ||S(y^*) - S^r||^2 = ||H e^\sigma - H^{x,r}||^2 \).

Assume by contradiction that \( y^* \) (a solution to the problem) is such that \( S(y^*) \) is not NSD. Observe that \( e^\sigma \) associated with \( y^* \) provides the value \( ||H e^\sigma - H^{x,r}||^2 = ||H e^\sigma_{\text{cave}} + H e^\sigma_{\text{ex}} - H^{x,r}||^2 \). This in turn can be expanded \( ||H e^\sigma_{\text{cave}} + H e^\sigma_{\text{ex}} - H^{x,r}||^2 = ||H e^\sigma_{\text{cave}} - H^{x,r}||^2 + 2\langle H e^\sigma_{\text{cave}}, H e^\sigma_{\text{ex}} - H^{x,r} \rangle + ||H e^\sigma_{\text{ex}}||^2 \). By construction, \( 2\langle H e^\sigma_{\text{cave}}, H e^\sigma_{\text{ex}} \rangle = 0 \). Then \( ||H e^\sigma_{\text{cave}} - H^{x,r}||^2 + 2\langle H e^\sigma_{\text{ex}} - H^{x,r} \rangle + ||H e^\sigma_{\text{ex}}||^2 \geq ||H e^\sigma_{\text{cave}} - H^{x,r}||^2 \). Because \( ||H e^\sigma_{\text{ex}} - H^{x,r}|| \geq 0 \), because both matrices are PSD. This contradicts that \( y^* \) is a solution. Therefore, \( y^* \) is such that \( S(y) \) is symmetric and NSD.

\[ \square \]

Lemma 7. If Walras’ law holds and if \( y^* \in \arg\min_{y \in X(p)} ||S(y) - S^r||^2 \), then \( S(y^*)(p,w)p = 0 \) (singular in \( p \)). Moreover, \( y^* \) is HD0.

\[ ^{15}\text{If it were not monotone, in particular if existence or uniqueness are not guaranteed (i.e., the problem is ill-posed) then we could use a regularization procedure to obtain a unique continuous expenditure (Tikhonov regularization). However, the result goes through also if there exists at least one continuous and differentiable solution \( e^\sigma \) to the equation above so we do not need uniqueness. Thus without loss of generality we assume that } e^\sigma(p,w) \text{ is monotone in wealth.} \]
Proof. Under Walras’ law and by the results of lemma (6) we apply John (1995) result to conclude that 
\( S(y^*)(p, w)p = 0 \) so it is singular in \( p \). By Walras’ law again, this leads us to conclude that 
\( D_p y^*(p, w)p + D_w y^*(p, w)w = 0 \), this implies that \( y^* \) is HDO.

This last result also yields an explicit solution locally. Observe that by definition \( e^\sigma(p, v(p, w)) = w. \)

Definition 4. [Concave part local indirect utility] We define the “concave part” of the indirect utility \( v_{cave} \) implicitly \( e^\sigma_{cave}(p, v_{cave}(p, w)) = w. \)

We also define a notion of Slutsky matrix function integrability:

Definition 5. [Slutsky integrability] We say a matrix function \( S \in M(Z) \) is Slutsky integrable if there exists some \( x \in \mathcal{X}(Z) \) such that \( S(x) = S \).

This definition can be modified appropriately for the local setup:

Corollary 1. The closest rational demand to \( x \in \mathcal{X}(\mathbb{P}) \) in the sense of minimizing its comparative statics error is given by \( x^* \in \mathcal{R}(\mathbb{P}) \) defined locally as \( x^*(p, w) = -\sum_{v \in v_{cave}(p, w)} xy_{cave}(p, w) \). Moreover its value function can be decomposed in four components \( ||S(x^*) - S(x)||^2 = ||E^x||^2 + ||E^\sigma||^2 + ||E^\sigma||^2 + \) where \( ||E^\sigma||^2 = \min_{y \in \mathcal{X}(\mathbb{P})} ||S(y) - S^*||^2 \) is the mis-specification error (distance of \( S^* \) from being the Slutsky matrix of some demand in \( \mathcal{X}(\mathbb{P}) \), i.e., Slutsky integrability).

Proof. This follows from Proposition 2 and Theorem 1.

6.2 Path Independence of Wealth Compensations

In what follows we study the Slutsky Wealth compensation problem and its relation with the Slutsky Error \( E^\sigma \). In this and the next subsection we assume that the price-wealth region \( Z \) is path connected. We need some preliminary definitions.

Definition 6. A price path is \( p : [0, 1] \mapsto \mathbb{R}^L_{++} \) with \( p(0) = p_0 \) and \( p(1) = p_1 \) is a continuous function that connects two points in a convex subset of \( \mathbb{R}^L_{++} \).

The Slutsky wealth compensation problem is that of compensating a given price change through a change in wealth in order to keep the initial demand quantities affordable after the change in prices. We ask what is the local wealth compensation needed to offset a small price change at any point in the price path. We want to compute the total Slutsky wealth compensation that is required to move from price situation \( p_0 \) to \( p_1 \). This means that along the path the consumer is always compensated in the sense of Slutsky for small price changes.

The local Slutsky compensation quantity change is related to the actual amount or level of wealth compensation by the following relation: \( \frac{\partial w(t)}{\partial t} = \frac{\partial p(t)'}{\partial t} x(p(t), w(t)) \) with Walras’ law holding along the path \( w(t) = p(t)'x(t) \), in particular we fix \( w(0) = w_0 \) at some level. Hence, the Slutsky compensation level at some \( t \) is given by \( w(t) = w(0) + \int_0^t \frac{\partial p(t)'}{\partial t} x(p(\tau), w(\tau))d\tau \) along a price path \( p \). Of course, the sum total wealth compensation change from \( (p_0, w_0) \) to \( p_1 \) is:
\[ w(1) - w(0) = \int_0^1 \frac{\partial p(t)'}{\partial t} x(p(\tau), w(\tau))d\tau. \]

We next show that \( E^\sigma \) is related to a notion of path independence in this wealth compensation. To introduce the path independence definition, we define a composite price path.
Definition 7. (Composite price path) Let \( p \) and \( p^* \) be price paths such that \( p(0) = p^*(0) \) and \( p(1) = p^*(1) \). A composite price path is defined as \( p(s, t) = p_s(t) = s p(t) + (1 - s) p^*(t) \) for all \( s \in [0, 1] \) and all \( t \in [0, 1] \).

To define path independence, we first need to consider the Slutsky wealth compensation problem for a fixed \( s \):
\[
\frac{\partial x}{\partial s} w_s(0) + \int_0^t \frac{\partial x}{\partial \tau} \frac{\partial p_s(\tau)}{\partial \tau} x(p_s(\tau), w_s(\tau)) d\tau.
\]

Next, we need to allow \( s \) to vary in order to capture the sensitivity to the paths. This leads to the following definition:

Definition 8. (Slutsky Wealth Compensation within-Path gap) The Slutsky Wealth Compensation within-Path gap is \( w_0(1) - w_1(1) \).

In this exercise, we have allowed \( \tau \) first to trace the interval \([0, 1]\), followed by a similar move in the parameter \( s \).

Alternatively, we can allow first \( s \) to move, followed by a move in \( \tau \) in determining the wealth compensations. To study first the relation of the dependence of the wealth compensation on the path (parameter \( s \)), we propose the following quantity (whose exact expression stems from the integral calculations in the ensuing proof). This is well defined for any composite price and wealth paths \( p_s \) and \( w_s \).

Definition 9. (Uncompensated across-path demand correction term) The uncompensated across-path demand correction term \( c = \frac{1}{2} \int_0^1 \int_0^1 \frac{\partial p_s(\tau)}{\partial \tau} D_w x(p_s(\tau), w_s(\tau)) \frac{\partial x}{\partial \sigma} A_\tau d\sigma d\tau \), (i.e., the weighted inner product \( \frac{\partial p_s}{\partial \tau}, \frac{\partial x}{\partial \sigma} \)). We refer to \( c \) as uncompensated across-path demand correction term because in general changes in wealth are not required to be compensated with respect to the parameter \( s \), which is first allowed to move, and \( c \) captures this divergence.

When the \( p_s \) and \( w_s \) are not Slutsky compensated paths with respect to \( s \), the \( c \) correction term allows us to obtain an equivalence between a quadratic form of \( E^s \) and a monetary measure of the possible departures of rationality. We propose the following definition, in which the order of moving \( s \) or \( \tau \) is irrelevant in affecting the wealth compensations:

Definition 10. (Slutsky wealth compensation path independence) Slutsky wealth compensation path independence holds whenever
\[
[w_0(1) - w_1(1)] - c = 0.
\]

Our next result follows:

Proposition 3. The Slutsky Wealth Compensation is path independent if and only if \( E^s = 0 \).

Moreover, for any price \( p_s \) and wealth \( w_s \) paths, the corrected path dependence wealth gap
\[
\frac{1}{2} [w_1(1) - w_0(1)] - c = \int_0^1 \int_0^1 \frac{\partial p_s(\tau)}{\partial \tau} E^s(p_s(\tau), w_s(\tau)) \frac{\partial p_s(\tau)}{\partial \tau} d\tau d\sigma.
\]

Proof. The first part of the statement follows from the Frobenius-Stokes theorem of integrability (Guggenheimer, 1962).

For the “moreover” part of the statement, given \( s \), we consider the function \( w_s(t) \), the solution for that \( s \) of the Slutsky Wealth Compensation problem:
\[
w_s(t) = w_s(0) + \int_0^t \frac{\partial p_s(\tau)}{\partial \tau} x(p_s(\tau), w_s(\tau)) d\tau.
\]
We compute the derivative with respect to $s$, $\frac{\partial w_s(t)}{\partial s}$:

$$\frac{\partial w_s(t)}{\partial s} = \int_0^t \frac{\partial p_s(\tau)}{\partial \tau} D_{\tau} x(p_s(\tau), w_s(\tau)) \frac{\partial p_s(\tau)}{\partial s} d\tau + \int_0^t \frac{\partial p_s(\tau)}{\partial s} D_{\tau} w x(p_s(\tau), w_s(\tau)) \frac{\partial w_s(\tau)}{\partial s} d\tau.$$

Focusing on $w_s(1)$ and using integration by parts on the third term:

$$\int_0^1 \frac{\partial^2 p_s(\tau)}{\partial \tau \partial s} x(p_s(\tau), w_s(\tau)) d\tau = \frac{\partial p_s(\tau)}{\partial s} x(\tau)|_0^1 - \left[ \int_0^1 \frac{\partial p_s(\tau)}{\partial s} \left[ D_{\tau} x(p_s(\tau), w_s(\tau)) \right] \frac{\partial p_s(\tau)}{\partial \tau} d\tau + \int_0^1 \frac{\partial p_s(\tau)}{\partial s} \left[ D_{\tau} w x(p_s(\tau), w_s(\tau)) \right] \frac{\partial w_s(\tau)}{\partial \tau} d\tau \right].$$

Notice that $\frac{\partial p_s(\tau)}{\partial s} x(\tau)|_0^1 = p^*(\tau) x(\tau)|_0^1 - p(\tau) x(\tau)|_0^1$. By Walras’ law this term vanishes.

Thus:

$$\frac{\partial w_s(1)}{\partial s} = \int_0^1 \frac{\partial p_s(\tau)}{\partial \tau} \left[ D_{\tau} x(p_s(\tau), w_s(\tau)) - D_{\tau} w x(p_s(\tau), w_s(\tau)) \right] \frac{\partial p_s(\tau)}{\partial s} d\tau + \int_0^1 \frac{\partial p_s(\tau)}{\partial s} \left[ D_{\tau} w x(p_s(\tau), w_s(\tau)) \right] \frac{\partial w_s(\tau)}{\partial \tau} d\tau.$$

By definition of the price path $p_s$ and by Walras’ law, we have $w_s(\tau) = p_s(\tau) x_s(\tau)$. This implies:

$$\frac{\partial w_s(\tau)}{\partial s} = x_s(\tau) \frac{\partial p_s(\tau)}{\partial s} + \frac{\partial w_s(\tau)}{\partial \tau} p_s(\tau)$$

and by the Slutsky compensation equation $\frac{\partial w_s(\tau)}{\partial \tau} = x(\tau) \frac{\partial w_s(\tau)}{\partial s}$.

Now, plugging this inside the integral equation above:

$$\int_0^1 \frac{\partial p_s(\tau)}{\partial \tau} D_{\tau} x(p_s(\tau), w_s(\tau)) \frac{\partial w_s(\tau)}{\partial s} d\tau - \int_0^1 \frac{\partial p_s(\tau)}{\partial s} D_{\tau} w x(p_s(\tau), w_s(\tau)) \frac{\partial w_s(\tau)}{\partial \tau} d\tau =$$

$$\int_0^1 \frac{\partial p_s(\tau)}{\partial \tau} \left[ D_{\tau} w x(p_s(\tau), w_s(\tau)) x(\tau) - x(\tau) D_{\tau} w x(p_s(\tau), w_s(\tau)) \right] \frac{\partial p_s(\tau)}{\partial s} d\tau +$$

$$- \int_0^1 \frac{\partial p_s(\tau)}{\partial s} \left[ D_{\tau} w x(p_s(\tau), w_s(\tau)) \right] \frac{\partial x_s(\tau)}{\partial s} p_s(\tau) d\tau.$$

Therefore,

$$\frac{\partial w_s(1)}{\partial s} = 2 \int_0^1 \frac{\partial p_s(\tau)}{\partial s} \left[ E^s(p_s(\tau), w_s(\tau)) \right] \frac{\partial p_s(\tau)}{\partial s} d\tau + c(s)$$

with $c(s) = - \int_0^1 \frac{\partial p_s(\tau)}{\partial s} \left[ D_{\tau} w x(p_s(\tau), w_s(\tau)) \right] \frac{\partial x_s(\tau)}{\partial s} p_s(\tau) d\tau$, and $E^s(p_s(\tau), w_s(\tau)) = \frac{S(p_s(\tau), w_s(\tau)) - S(p_s(\tau), w_s(\tau))'}{2}$.

Finally, by integration:

$$w_1(1) - w_0(1) = \int_0^1 \frac{\partial w_s(1)}{\partial s} ds + \int_0^1 c(s) ds$$

$$\frac{w_1(1) - w_0(1)}{2} = c = \int_0^1 \int_0^1 \frac{\partial p_s(\tau)}{\partial s} E^s(p_s(\tau), w_s(\tau)) \frac{\partial p_s(\tau)}{\partial s} d\tau ds.$$

16This function is differentiable (by the results given by Guggenheimer (1962)).
Next, we establish the connection with the Slutsky matrix norm. The local change of wealth compensation with respect to the change of path is measured by \( \partial^2 w_{s}(t) / \partial s \partial t \) –c\( s(t) \), and this term is controlled globally by a weighted norm of \( E^\pi \) –suitable marginal quantity changes correcting for possible path dependence – properly weighted by the changes in prices. Specifically:

**Proposition 4.** Given a composite price path, for any Slutsky wealth compensated path such that \( w_s(t) = p_s(\tau)' x(\tau) \), it is the case that \( \left[ \frac{\partial^2 w_{s}(t)}{\partial s \partial t} - c_s(t) \right]^2 = || E^\pi(s,t) ||_{WU}^2 \) with \( || \cdot ||_{WU} \) the weighted Frobenius norm with weighting \( W_s(t) = \sqrt{2} \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} \) when \( W \) is PSD pointwise and \( U \) is the identity. Moreover, the path dependence wealth gap from any two paths is bounded above by the average weighted Frobenius norm: \( \frac{1}{2} [w_1(1) - w_0(1)] - c^2 \leq || E^\pi ||_{WU}^2 \).

**Proof.** Observe that

\[
\partial^2 w_{s}(t) - c_s(t) = 2 \frac{\partial p_s(\tau)' E^\sigma(p_s(\tau), w_s(\tau)) \partial p_s(\tau)'}{\partial t}.
\]

Note that we can write \( \left[ \frac{\partial^2 w_{s}(t)}{\partial s \partial t} - c_s(t) \right]^2 = Tr \left( \left| \frac{\partial^2 w_{s}(t)}{\partial s \partial t} - c_s(t) \right|^2 \right) \), where the left hand side is a scalar and thus its trace is itself.

\[
Tr \left( \left| \frac{\partial^2 w_{s}(t)}{\partial s \partial t} - c_s(t) \right|^2 \right) = 2Tr \left( \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} E^\pi(p_s(t), w_s(t)) \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} E^\pi(p_s(t), w_s(t)) \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} \right).
\]

By the properties of the trace,

\[
Tr \left( \left| \frac{\partial^2 w_{s}(t)}{\partial s \partial t} - c_s(t) \right|^2 \right) = 2Tr \left( 2E^\pi(p_s(t), w_s(t)) \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} E^\pi(p_s(t), w_s(t)) \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} \right).
\]

Then we can establish the first part of the result:

\[
Tr \left( \left| \frac{\partial^2 w_{s}(t)}{\partial s \partial t} - c_s(t) \right|^2 \right) = || E^\pi(p_s(t), w_s(t)) \sqrt{2} \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} ||_M^2 \text{ and}
\]

\[
|| E^\pi(p_s(t), w_s(t)) \sqrt{2} \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} ||_M^2 = || W_s(t) E^\pi(p_s(t), w_s(t)) ||_M^2.
\]

Finally, for the second part of the statement we apply Holder’s inequality to conclude that

\[
\frac{1}{2} [w_1(1) - w_0(1)] - c^2 \leq \int_0^1 \int_0^1 \left| \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} E^\pi(p_s(\tau), w_s(\tau)) \right|^2 \partial \tau \partial ds.
\]

By the first part of the statement, we conclude

\[
\int_0^1 \int_0^1 \left| \frac{\partial p_s(t) \partial p_s(t)'}{\partial t} E^\pi(p_s(\tau), w_s(\tau)) \right|^2 \partial \tau \partial ds = \int_0^1 \int_0^1 || W_s(t) E^\pi(p_s(t), w_s(t)) ||_M^2 \partial \tau \partial ds = || E^\pi ||_{WU}^2.
\]

6.3 The Compensated Law of Demand

So far we have related our error Slutsky matrix \( E^\pi \) to a path independence notion in wealth compensated paths. Now we move on to relate \( E^\pi \) and \( E^\nu \) to WARP and the Compensated law of demand.

The compensated law of demand states that prices and quantities “move in opposite direction,” once the consumer’s wealth is compensated so that she can afford previous bundles. The compensated law of demand can be written in terms of price paths as \( \frac{\partial p(t) \partial x(t)}{\partial t} \leq 0 \) for all \( t \in [0,1] \), for every Slutsky Wealth Compensated path. A Slutsky Wealth Compensation path \((p(t), w(t))_{t \in [0,1]} \) is such that \( \frac{\partial w(t)}{\partial t} = x(p(t), w(t))' \frac{\partial x(t)}{\partial t} \) for all \( t \in [0,1] \).

Here we are interested in the sign and magnitude of the term \( \frac{\partial p(t) \partial x(t)}{\partial t} \) for Slutsky Wealth Compensated paths \( \frac{\partial w(t)}{\partial t} = x(p(t), w(t))' \frac{\partial x(t)}{\partial t} \), which we call locally compensated demand paths. The consumer that does not satisfy the compensated law of demand will have \( \frac{\partial p(t) \partial x(t)}{\partial t} > 0 \) for some paths.

It is well known that the locally compensated demand is related to a quadratic form of the Slutsky matrix function:

\[
\frac{\partial p(t) \partial x(t)}{\partial t} = \frac{\partial x(t)'}{\partial t} S(p(t), w(t)) \frac{\partial p(t)}{\partial t}.
\]

We note that the locally compensated demand path is measured in money.

Our next result relates the locally compensated law of demand to \( E^\pi \) and \( E^\nu \).
**Proposition 5.** Consider any locally compensated path. The locally compensated demand path satisfies the compensated law of demand if and only if $E^\pi = 0$ and $E^\nu = 0$. Moreover, 
\[
\frac{\partial x(t)}{\partial t} = \frac{\partial x(t)}{\partial t} [S^r + E^\pi + E^\nu] \frac{\partial p(t)}{\partial t}.
\]

Proof. We start by stating the definition \(\frac{\partial p(t)}{\partial t} (S^r(t), w(t)) \frac{\partial p(t)}{\partial t} \) for \((p(t), w(t))\) such that \(\frac{\partial p(t)}{\partial t} = x(p(t), w(t)) \frac{\partial p(t)}{\partial t} \). By virtue of the decomposition in Theorem 1, we can write 
\[S = S^r + E^\pi + E^\nu.\]
Also, we notice that \(\frac{\partial p(t)}{\partial t} E^\pi(t) \frac{\partial p(t)}{\partial t} = 0\) for all \(t \in [0,1]\) and any price path because \(E^\pi\) is a skew-symmetric matrix function and it vanishes in a quadratic form.

This establishes the second statement in the proposition. That is, this implies that: \(\frac{\partial p(t)}{\partial t} \frac{\partial x(t)}{\partial t} = \frac{\partial p(t)}{\partial t} [S(t) + E^\pi(t) + E^\nu(t)] \frac{\partial p(t)}{\partial t} \).

Now, we turn to the first statement. If the compensated law of demand holds for any price path and Slutsky wealth compensated path, it follows that: \(\frac{\partial p(t)}{\partial t} [S^r(t)] \frac{\partial p(t)}{\partial t} \leq 0\) which implies that \(S^r(t)\) is NSD so \(S^\pi,\pi\) is NSD and \(E^\nu = 0\) must be zero by definition. This also implies that \(S\) is NSD because so is its symmetric part, and by John (1995) (theorem 3) this implies that \(x\) is HD0 which in turn implies that \(E^\pi = 0\).

Now, if \(E^\pi = 0\) and \(E^\nu = 0\) then we conclude that \(S = S^r\). This implies in turn that \(\frac{\partial x(t)}{\partial t} = \frac{\partial x(t)}{\partial t} S^r(p(t), w(t)) \frac{\partial p(t)}{\partial t}\) because \(E^\pi\) is skew-symmetric. By construction, \(\frac{\partial x(t)}{\partial t} S^r(p(t), w(t)) \frac{\partial p(t)}{\partial t} \leq 0\) so the compensated law of demand holds for any locally compensated demand path.

The dot product of price and quantity changes defined above is important. Its discretized version \((p^t - p^{t-1})(x^t - x^{t-1}) \leq 0\) corresponds to the Compensated law of demand for compensated price changes \(p^t(x^t - x^{t-1}) = w^t\). The result above quantitatively connects violations of this necessary condition for rationality with the error Slutsky matrix functions \(E^\pi\) and \(E^\nu\). Next, we connect this to the Slutsky matrix norm.

**Proposition 6.** Consider any locally compensated demand path. For two demand functions \(x\) and \(y\) such that their nearest rational Slutsky matrix function is the same \(S^\pi(x, y) = S^\nu(x, y)\), it holds that for the weighting matrix \(W(t) = \frac{\partial p(t)}{\partial t} \frac{\partial x(t)}{\partial t}\) and \(U\) the identity, one has that 
\[||E^\pi||_U + ||E^\nu||_U \geq ||E^\nu||_U + ||E^\nu||_U\] if and only if \(\frac{\partial p(t)}{\partial t} \frac{\partial x(t)}{\partial t} \geq \frac{\partial p(t)}{\partial t} \frac{\partial x(t)}{\partial t}\).

Proof. If \(\frac{\partial p(t)}{\partial t} \frac{\partial x(t)}{\partial t} \geq \frac{\partial p(t)}{\partial t} \frac{\partial x(t)}{\partial t}\) under the assumption \(S(x, y) = S(y, x)\) it holds that \(\frac{\partial p(t)}{\partial t} E^\pi(x, y) + E^\nu(y, x) \leq \frac{\partial p(t)}{\partial t} [E^\nu(x, y) + E^\nu(y, x)] \frac{\partial p(t)}{\partial t} \).

Observe that 
\[Tr[\frac{\partial p(t)}{\partial t} (E^\pi(x, y) + E^\nu(y, x)) \frac{\partial p(t)}{\partial t} \frac{\partial p(t)}{\partial t} [E^\pi(x, y) + E^\nu(y, x)] \frac{\partial p(t)}{\partial t} \frac{\partial p(t)}{\partial t}] = ||\frac{\partial p(t)}{\partial t} ||_U^2\] because the expression is a quadratic form, a scalar.

By the cyclical properties of the trace:
\[Tr[\frac{\partial p(t)}{\partial t} [E^\pi(x, y) + E^\nu(y, x)] \frac{\partial p(t)}{\partial t} [E^\pi(x, y) + E^\nu(y, x)] \frac{\partial p(t)}{\partial t} ] =
\[Tr[\frac{\partial p(t)}{\partial t} [E^\pi(x, y) + E^\nu(y, x)] \frac{\partial p(t)}{\partial t} [E^\pi(x, y) + E^\nu(y, x)] \frac{\partial p(t)}{\partial t} ] =
\[||W(t) ||^2_2\]
where \(W(t) = \frac{\partial p(t)}{\partial t} \frac{\partial p(t)}{\partial t}\).

By an analogous argument to that in Theorem 1 above, 
\[||W(t) ||^2_2\]

This implies that 
\[||E^\pi||_U + ||E^\nu||_U \geq ||E^\nu||_U + ||E^\nu||_U\].
Conversely, if \( \|Ex,\pi\|_W + \|Ex,\nu\|_W \geq \|Ey,\pi\|_W + \|Ey,\nu\|_W \) and \( S_{x,r}(p,w) = S_{y,r}(p,w) \), it follows that \( \frac{\partial p(t)}{\partial t} \frac{\partial x(t)}{\partial t} \geq \frac{\partial p(t)}{\partial t} \frac{\partial y(t)}{\partial t} \) and this inequality is preserved under monotone transformations so that \( \frac{\partial p(t)}{\partial t} \frac{\partial x(t)}{\partial t} \geq \frac{\partial p(t)}{\partial t} \frac{\partial y(t)}{\partial t} \).

\[ \square \]

7 Examples and Applications

The rationality assumption has long been seen as an approximation of actual consumer behavior. Nonetheless, to judge whether this approximation is reasonable, one should be able to compare any alternative behavior with its best rational approximation. Our results may be helpful in this regard, as the next examples illustrate.

7.1 The Sparse-Max Consumer Model of Gabaix (2014)

This model generates analytically tractable behavioral demand functions and Slutsky matrices. In this example, we compare the matrix nearness distance to the “underlying rational” Slutsky matrix function proposed by Gabaix and compare it to the one proposed here. This example shows that there exists a rational demand function that is behaviorally closer to the sparse max consumer demand proposed by Gabaix than the “underlying rational” model of his framework.

Consider a Cobb-Douglas model \( x^{CD}(p,w) \) such that:

\[
\begin{align*}
x_{i}^{CD} &= \frac{\alpha_{i}w}{p_{i}}, \quad i = 1,2, \\
x_{i}^{CD} &= -\frac{\alpha_{i}w}{p_{i}}, \\
x_{i}^{CD} &= \frac{\alpha_{i}}{p_{i}}, \\
x_{i}^{CD} &= \frac{\alpha_{i}w}{p_{i}} + \frac{\alpha_{i}w}{p_{1}}, \quad i = 1,2, \\
s_{i,j}^{CD} &= \frac{\alpha_{i}w}{p_{i}} \frac{\alpha_{j}w}{p_{j}}. \\
S^{CD}(p,w) &= \begin{bmatrix}
-\frac{\alpha_{1}\alpha_{2}w}{p_{1}^{2}} & \frac{\alpha_{1}\alpha_{2}w}{p_{1}p_{2}} & \frac{\alpha_{1}w}{p_{1}} \\
\frac{\alpha_{1}\alpha_{2}w}{p_{1}p_{2}} & -\frac{\alpha_{1}w}{p_{2}} & \alpha_{1}w \\
\frac{\alpha_{1}w}{p_{1}} & \frac{\alpha_{1}w}{p_{2}} & \frac{\alpha_{1}w}{p_{2}}
\end{bmatrix}.
\end{align*}
\]

Let us denote Gabaix’s theory of behavior of the Sparse-max consumer by \( G \). Then the demand system under \( G \) is:

\[
x_{i}^{G} = \frac{\alpha_{i}w}{p_{i}} \sum_{j} \frac{\alpha_{j}w}{p_{j}} \quad \text{for } i = 1,2.
\]

This demand system fulfills Walras’ law. This function has an additional parameter with respect to \( x^{CD}(p,w) \), the perceived price \( \bar{p}_{i}(m) = m_{i}p_{i} + (1 - m_{i})p_{i}^{d} \). The vector of attention to price changes \( m \), weighs the actual price \( p_{i} \) and the default price \( p_{i}^{d} \).

Consider the following matrix of attention for the sparse-max consumer:

\[
M = \begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix}
\]

That is, the consumer does not pay any attention to price changes in \( p_{2} \), but perceives price changes perfectly for \( p_{1} \). One of Gabaix’s elegant results relates the Slutsky matrix function of \( x^{G} \), to the Cobb-Douglas benchmark. The behavioral Slutsky matrix evaluated at default prices (in all this example \( p = p^{d} \)) is:

\[
S^{G}(p,w) = S^{CD}(p,w)M
\]

\[
S^{G}(p,w) = \begin{bmatrix}
-\frac{\alpha_{1}\alpha_{2}w}{p_{1}^{2}} & 0 \\
\frac{\alpha_{1}w}{p_{1}} & \frac{\alpha_{1}w}{p_{2}} & 0
\end{bmatrix}.
\]

This matrix is not symmetric, not NSD, nor singular with \( p \) in its null space. Applying Theorem 1, the nearest Slutsky matrix when \( p = p^{d} \) is:

\[ \square \]
\[ S'(p, w) = \frac{p^2}{p_1 + p_2} \begin{bmatrix} -\frac{\alpha_1 \alpha_2 w}{p_1} & \frac{\alpha_1 \alpha_2 w}{p_2} \\ \frac{\alpha_1 \alpha_2 w}{p_1} & -\frac{\alpha_1 \alpha_2 w}{p_2} \end{bmatrix} \]

Also, one has
\[
E(p, w) = S(p, w) - S'(p, w)
\]
\[
E(p, w) = - \begin{bmatrix} \left[ 1 - b(p) \right] \left[ \frac{\alpha_1 \alpha_2 w}{p_1} \right] & b(p) \left( \frac{\alpha_1 \alpha_2 w}{p_2} \right) \\ \left[ 1 - b(p) \right] \left[ \frac{\alpha_1 \alpha_2 w}{p_1} \right] & b(p) \left( -\frac{\alpha_1 \alpha_2 w}{p_2} \right) \end{bmatrix}
\]
with
\[ b(p) = \frac{p^2}{p_1 + p_2}. \]

Now, we compute a useful quantity:
\[
Tr(E'E) = \frac{w^2 \alpha_1 \alpha_2}{p_1 + p_2}.
\]

It is convenient to compute the contributions of the violations of symmetry and singularity in \( p \) separately.
\[
Tr(E'E) = Tr(E\sigma' E\sigma) + Tr(E\nu' E\nu) = \frac{1}{2} \frac{w^2 \alpha_1 \alpha_2}{p_1 + p_2} + \frac{1}{2} \frac{w^2 \alpha_1 \alpha_2}{p_1 + p_2}.
\]
In this case, regardless of the values that \( w \) takes, the contribution of each kind of violation is equal and amounts to exactly half of the total distance. In fact, we have: \[ ||E||^2 \geq \frac{1}{2} \int \! \! \! \!
\]
\[
\int \! \! \! \! dE \left( \frac{E\sigma'(E\sigma)}{p_1 + p_2} \right) \alpha_1 \alpha_2 w^2 \] \[ \frac{1}{p_1 + p_2} \]
with \[ p = p_d \].

Note, however, that in this example the third component of the violations, the one stemming from NSD, is zero when the prices are evaluated at the default. Since the behavioral model proposed by Gabriel does not satisfy WARP and its Slutsky matrix function violates NSD, we conclude that the violation of the WARP is not massive, in fact, it affects the size of the Slutsky norm only through its interactions with homogeneity of degree zero or “money illusion”. Our approach can also be used in the general case. We compute the previous quantities at any \( p \), and any \( p_d \) with \( m = [1, 0]' \).
\[
Tr(E'E) = \frac{w^2 \alpha_1 \alpha_2}{p_1 + p_2} \frac{[p_d]^2}{[p_2 + p_2 - p_2 \alpha_1]^2}, \text{ with } Tr(E\sigma' E\sigma) = Tr(E\nu' E\nu) \text{ and } E\nu = 0.
\]

The expression above has a positive derivative with respect to \( p_d \) for \( \alpha_1 + \alpha_2 = 1 \), this indicates that \[ \frac{\partial}{\partial p_d} \delta(\cdot) > 0 \] for any \( p \). A sparse consumer that does not pay attention to \( p_2 \) will be further from rationality when the default price \( p_d \) is high. Furthermore, the power of our approach lies in the decomposition of \[ ||E||^2 = ||E\sigma||^2 + ||E\nu||^2 \]. In this case, the decomposition suggests that the violation of WARP can be seen as a byproduct of the violations of symmetry and singularity stemming from the “lack of attention” to price changes of good 2 and the “nominal illusion” or lack of homogeneity of degree zero in prices and wealth in such a demand system.\(^{17}\)

To show the tractability of our approach in this example, we will study a very simple region \( Z \), with the aim of illustrating how one can learn from the effect of a behavioral parameter such as \( \alpha_1 \) and \( p_d \). Let \( Z = \{ w, p_1 = 1, p_2 \in [1, 2] \} \), then \[ \delta(\alpha_1, p_d) = \frac{1}{3} (\alpha_1 - 1) \alpha_1^2 [p_d^2] \]
\[
\left( \frac{1}{p_2^2 + 1} - \frac{1}{p_2^2 + 1} \right) \]
One can now visualize this \( \delta \) in the \( \alpha_1, p_d \) space, that is at \( p_d \in [1, 2] \) and \( \alpha_1 \in [0, 1] \) (figure 2). We can observe that \( \alpha_1 \) has a non-linear effect on \( \delta \), and the distance toward the rational matrix goes to zero when either \( \alpha_1 \to 0 \) or \( \alpha_1 \to 1 \) for all \( p_d \in [1, 2] \).

To finish this example we study the errors in comparative statics analysis. This is done locally.\(^{18}\) Fix \((p, w) = (p_d, \pi) \) as the reference point. First, we compare the distance between \( S^G \) and \( S^{CD} \), in terms of errors in comparative statics analysis. If we use \( x^{CD} \) as an approximation

---
\(^{17}\)This observation is robust to other specifications (e.g., CES family).

\(^{18}\)In general, our approach does not require finding the closest \( z' \in R(Z) \) to \( x^G \), nevertheless doing so may help to complement the understanding of how much this particular bounded rationality model differs from the standard rational one in practical terms.
of \( x^G \) at the default prices, we obtain the following local comparative statics error 
\[ ||S^C(p^d, \overline{w}) - S^{CD}(p^d, \overline{w})||_M^2 = \frac{(p_1^d)^2 + (p_2^d)^2}{(p_1^d)^2 + (p_2^d)^2} \cdot \frac{w^2}{(p_1^d)^2 + (p_2^d)^2}. \]
By our results in Subsection 6.1, we can decompose this quantity as the sum of a mis-specification error and \( ||E(p^d, \overline{w})||_M^2 \). In fact, \( ||S^{CD}(p^d, \overline{w}) - S^G(p^d, \overline{w})||_M^2 = ||S^{CD}(p^d, \overline{w}) - S^R(p^d, \overline{w})||_M^2 + ||E(p^d, \overline{w})||_M^2 \), with \( ||S^{CD}(p^d, \overline{w}) - S^R(p^d, \overline{w})||_M^2 = \frac{w^2}{(p_1^d)^2 + (p_2^d)^2} \) and \( ||E(p^d, \overline{w})||_M^2 = \frac{w^2}{(p_1^d)^2 + (p_2^d)^2} \).

We know also, by the results in Subsection 6.1, that even if we could improve the mis-specification error, the Slutsky matrix error norm \( ||E(p^d, \overline{w})||_M^2 \) will not change at all. To illustrate this, we find the Cobb-Douglas demand that minimizes the error in comparative statics analysis at the reference price-wealth pair \((p^d, \overline{w})\). We can write this problem parametrically, through characterizing the Cobb-Douglas family by a parameter \( \beta \in [0, 1] \): \( x^{CD}(p, w, \beta) = (\frac{w}{p_1^d})^{-(1-\beta)} w^\beta \). We solve the problem \( \hat{\beta} = \arg\min_{\beta \in [0, 1]} ||S^{CD2}(p^d, \overline{w}, \beta) - S^G(p^d, \overline{w})||_M^2 \). There are two solutions \( \hat{\beta}^* = \frac{1}{2} \pm \frac{\sqrt{\beta^2 + (\frac{2\nu}{w})^2(1 - 2\alpha)^2}}{2\nu p^d} \), which depend on the given parameters and the fixed default price. The mis-specification error is zero, that is \( ||S^{CD2}(p^d, \overline{w}, \beta^*) - S^R(p^d, \overline{w})||_M^2 = 0 \). Because of this, in this case, the error in comparative statics reduces to the behavioral error captured by the Slutsky norm: \( ||S^{CD2}(p^d, \overline{w}, \beta) - S^G(p^d, \overline{w})||_M^2 = ||E(p^d, \overline{w})||_M^2 \), where \( ||E(p^d, \overline{w})||_M^2 = \frac{w^2}{(p_1^d)^2 + (p_2^d)^2} \).

Observe that it does not depend on \( \beta^* \) and remains unchanged with respect to the case of \( S^{CD} \). Of course, even if we used the “best” rational model to minimize errors in comparative statics analysis locally, we would not reduce at all the Slutsky matrix error norm that stems from lack of rationality as captured by properties \( \sigma, \pi, \nu \).

Figure 2: Level curves sparse-max consumer example. The figure plots \( \alpha \in [0, 1] \) on its vertical axis and \( p_2^d \in [1, 2] \) on its horizontal axis.
7.2 Hyperbolic Discounting

The literature on self-control and hyperbolic discounting has flourished in macroeconomics and development economics. In this example, we study a three-period model that allows us to illustrate the use of our methodology. Our aim is to measure the violations of property $\mathcal{R}$ by naive and sophisticated quasi-hyperbolic discounters.

The optimization problem for a consumer that can pre-commit is: $\max_{\{x_i^h\}_{i=1,2,3}} u(x_i^h) + \beta \theta u(x_i^2) + \beta^2 u(x_i^3)$ subject to the budget constraint $\sum_{i=1}^{3} p_i x_i^h = w$.

Which gives the demand system: (i) $x_1^h = [p_1 + p_2 [\beta \theta p_3]^{\frac{1}{2}} + p_3 [\beta^2 \theta p_3]^{\frac{1}{2}}]^{-1} w$. (ii) $x_2^h = [\beta \theta p_3]^{\frac{1}{2}} x_1^h$. (iii) $x_3^h = [\beta^2 \theta p_3]^{\frac{1}{2}} x_1^h$.

The naive quasi-hyperbolic discounter will have the following demand system:

In the first period, the consumer assumes he will stick to his commitment in the second period and consumes the same amount as in the pre-commitment case (i.e. $x_1^h = x_1^f$). However, when period two arrives, he re-optimizes taking as given the remaining wealth $w - p_1 x_1^h : x_2^h = \frac{w - p_1 x_1^h}{p_2 + p_3 [\beta \theta p_3]^{\frac{1}{2}}}$ and $x_3^h = [\beta \theta p_3]^{\frac{1}{2}} x_2^h$.

The analytical result for the matrix nearness norm has a nice expression:

$$Tr(E^E) = \frac{(\sigma - 1)^2 w^2 \left(p_1^2 + p_2^2 + p_3^2\right) - \left(\beta \theta p_1^2 \frac{1}{p_3}\right) \left(\beta \theta p_2^2 \frac{1}{p_3}\right)}{2 \sigma^2 \left(p_2 \left(\beta \theta p_3 \frac{1}{p_2}\right)^{\frac{1}{2}} + p_2\right) - \left(\beta \theta p_1^2 \frac{1}{p_3}\right) \left(\beta \theta p_2^2 \frac{1}{p_3}\right)}.$$  
which readily gives us that: (i) when $\sigma = 1$ then $Tr(E^E) = 0$ and $\delta = 0$, that is the demand is rational, (ii) when $\beta = 1$ then $Tr(E^E) = 0$, (iii) finally when $\beta \to 0, \theta \to 0$, then $\delta \to 0$. In these three cases by the previous results $\epsilon \to 0$. In fact, in the limit cases the hyperbolic demand system is rational. Take for instance case (iii), because the agent consumes everything in the first period and gives no weight to the other time periods then it is trivially rational, with $S^* \to 0$ and $x_1^h \to \frac{w}{p_1}$ and $x_2^h, x_3^h \to 0$. In case (i), the logarithmic utility case, the hyperbolic discounter manages to keep his commitment and therefore his consumption is time consistent and $\|E\| = 0$.

To illustrate further the use of the tools developed here, we find an explicit value for $\delta$ in terms of the behavioral parameters $\beta, \theta$, for an arbitrary rectangle $Z$ of prices and wealth. Consider the region $Z = \{p_1, p_2, p_3 = 1, w \in [1, 2]\}$, with $\sigma = \frac{1}{2}$ we compute $\delta(\beta, \theta)$, which can be represented graphically in the box $\beta \in [0, 1]$, $\theta \in [0, 1]$. The analytical expression for $\delta^2 = \frac{\gamma_p^4 (\sigma^2 - 1)^2 \gamma_p^4}{2(\beta^2 \theta + 1)^2 (\beta \theta + 1)^2 (\beta^2 \theta + 1)^2}$.\footnote{One can also use the level curves in this example to identify pairs $(\beta, \theta)$ that are “equidistant” from rationality, capturing an interesting tradeoff between the short-run and the long-run discount factors and its effects on the violations of the Slutsky conditions.}

The level curves (figure 3) show that the hyperbolic discounter is very close to the rational consumer, in the matrix nearness sense, for very low values of $\beta, \theta$ and for values of $\theta \leq \frac{1}{2}$. This makes intuitive sense as a lower $\theta$ means heavier discount on the future and lower consumption of goods of time 2 and 3 that are the ones affected by self-control.

Another observation that we can draw from this example is that for any arbitrary compact region $Z$ of prices and wealth analyzed $\|E\|^2 = \|E^E\|^2$. That is, only the asymmetric part plays a role in the violation of property $\mathcal{R}$. In other words, under this numerical conditions the hyperbolic discounter violates symmetry but it satisfies singularity in prices and negative
semidefiniteness. The hyperbolic discounter fulfills WARP.

7.3 Sophisticated Quasi-Hyperbolic Discounting

The sophisticated quasi-hyperbolic discounter is intuitively closer to rationality. However, the Slutsky norm helps appreciate some of the subtleties and assess which conditions of rationality are fulfilled by this type of consumer. We build this example as a followup to the naive quasi-hyperbolic consumer. In this case, the consumer knows that in $t = 2$ he will not be able to keep his commitment and therefore will adjust his consumption at $t = 1$. Then the consumer maximizes

$$\max_{x_1^h} u(x_1^h) + \beta \theta u(x_2^h) + \beta^2 \theta^2 u(x_3^h)$$

where $x_2^h, x_3^h$ are known to his in $t = 1$ and depend on period 1 consumption. However, he can control only how much he consumes in the first period. Then, the first period consumption under sophisticated hyperbolic discounting is:

$$x_1^{sh} = \left[ \frac{p_1 \beta \theta + p_1 \beta^2 \theta^2 \left[ \frac{\beta \theta p_2}{p_3} \right]^{\frac{1-\sigma}{\sigma}}}{p_2 + p_3 \left[ \frac{\beta \theta p_2}{p_3} \right]^{\frac{1-\sigma}{1-\sigma}}} \right]^{\frac{1}{\sigma}} + p_1^{-1} w$$

The argument in the integral of the expression for $\delta$ for a generic $Z$ is given by the quantity:

$$Tr(E'E) = 25$$
\[
(\beta - 1)^2(\sigma - 1)^2w^2\theta^{1/\sigma}(p_1^2 + p_2^2 + p_3^2)\left(\beta p_2 + p_3 \left(\frac{\beta \theta p_2}{p_3}\right)^{\frac{1}{\sigma}}\right)^{\frac{2}{\sigma}} - 2\left(p_1 \left(\frac{\beta \theta p_2}{p_3}\right)^{\frac{1}{\sigma}} + p_2\right)
\]

As expected, this implies that: (i) when \(\sigma = 1, \delta = 0\) for any \(Z\); (ii) when \(\beta = 1, \delta = 0\); and (iii) when \(\beta = 0, \delta = 0\). Thus, in all these cases, \(\epsilon = 0\). Also, the decomposition of \(||E||^2 = ||E^\sigma||^2\), which means that only the symmetry property is violated, while the weak axiom and the homogeneity of degree zero in prices and wealth are preserved.

Finally, we compare this quantity with the case of the naive hyperbolic discounter. We consider the ratio \(r = \frac{T_r(E^{\sigma\theta, E^{\sigma\theta}})}{T_r(E^{\sigma\theta}, E^{\sigma\theta})}\). When \(r < 1\), the sophisticated hyperbolic consumer has a lower \(\delta\) for any \(Z\) and any parameter configuration. The first finding is that the ratio \(r\) and the implied values of \(\delta\) depend crucially on the parameter \(\sigma\). For \(\sigma = 1\), both \(E^{\sigma\theta}, E^{\sigma\theta}\) are equal to zero: this is a knife-edge case, in which the marginal rates of substitution yield optimal consumptions equal to the commitment baseline. For \(\sigma < 1\) (e.g., \(1/2\)), the sophisticated hyperbolic discounter has a uniformly lower \(\delta\) and thus \(r < 1\).

However, for \(\sigma > 1\) (e.g., equal to \(2\)), the naive hyperbolic discounter has a uniformly lower \(\delta\), hence \(r > 1\). Although this may seem counterintuitive, it tells us that the closest rational type (which need not be the commitment baseline) is closer for the naive than it is for the sophisticated consumer. To see this, note that the “rational” comparative statics analysis fails due to two possible sources: (i) The lack of self-control in the consumption of period 2 (overspending); and (ii) in the case of the sophisticated type, an additional effect due to the strategic change in period 1 consumption (overspending or underspending). So, the “rational” comparative statics analysis gives a higher failure than in the naive case for \(\sigma > 1\), because for an increase in prices for the naive case the change in period 1 consumption equals that in the rational case and only the second and third period demanded amounts are different. In contrast, for the sophisticated case, demanded amounts for the three goods all differ, in particular due to a higher level of wealth remaining at the end of period 1, leading to a higher overspending in period 2. This observation holds for such CRRA instantaneous utility with any parameter values of \(\delta, \theta \in [0, 1]\).\textsuperscript{20}

8 Literature Review

The canonical treatment of measuring deviations from rational consumer behavior was establish by Afriat (1973) with its critical cost-efficiency index. Afriat’s index measures the amount by which budget constraints have to be adjusted so as to eliminate violations of the Generalized

\textsuperscript{20}Further, to enhance the comparison for the case of \(\sigma = \frac{1}{2}\), we compute explicitly the expression for \(\delta\) in the same region \(Z = \{p_1, p_2, p_3 = 1, w \in [1, 2]\}\) as in the previous example for the naive discounter: \(\delta^2(\beta, \theta) = [14(\beta - 1)^2\beta^2\theta^2(\beta \theta^2 + 1)]^2/16\beta^2(\beta \theta^2 + 2) + 1\). The level curves of this \(\delta\) expression are very similar to the naive case, but it is closer to rationality uniformly for all parameter configurations. Evaluated at the values of \(\beta = 0.7\) and \(\theta = 0.9\), one gets the value \(\delta = 0.07069347\), which is slightly lower than the \(\delta\) of the naive hyperbolic case for the same \(\sigma = \frac{1}{2}\) (\(\delta = 0.074\), where the parameters are taken from the empirical literature.

26
Axiom of Revealed Preference (GARP). Varian (1985; 1990) refines Afriat’s measure by focusing on the minimum adjustment of the budget constraint needed to eliminate violations of GARP. Houtman and Maks (1985) measure deviations from GARP through identifying the largest subset of choices that is consistent with maximizing behavior. More recently, Echenique, Lee and Shum (2011), give a new measure of violations of revealed preference behavior called the “money pump index”. Also Jerison and Jerison (2012) propose a way to bound Afriat’s index of cost-efficiency using an index of violations of the symmetry and negative semidefiniteness Slutsky conditions. It would be interesting to compare our Slutsky matrix norm with these other approaches.

The closest treatment of the problem to our work is the approximately rational consumer demand proposed by Jerison and Jerison (1992; 1993). These authors are able to relate the violations of negative semidefiniteness and symmetry of the Slutsky matrix to the smallest distance between an observe smooth demand system and a rational demand. Russell (1997) proposes a notion of quasi-rationality. Russell’s argument links the Slutsky matrix anti-symmetry part with the lack of integrability of a demand system.

Our work takes a different methodological approach to this problem and generalizes the results to the case of violations of singularity of the Slutsky matrix. More importantly, this new approach allows to treat the three kinds of violations of the Slutsky conditions simultaneously. For instance, new behavioral models like the sparse-max consumer Gabaix (2014) suggest the presence of a money illusion such that prices are not in the null space of the Slutsky matrix. We also emphasize a positive measure of bounded rationality based on errors in making comparative statics analysis when one assumes (perhaps incorrectly) that the consumer is rational. The only other work we are aware of that analyzes comparative statics with behavioral consumers is Farhi & Gabaix (2015). In their interesting work they propose alternate ideas to the traditional Slutsky matrix including extensions of the Slutsky matrix for nonlinear budget constraints. It remains an open task how to extend our methodology to these new concepts.

9 Conclusion

By redefining the problem of finding the closest rational demand to an arbitrary observed behavior in terms of matrix nearness, we are able to pose the problem in a convex optimization framework that permits a better computational implementability and provide a tractable approach with a closed form solution. We define a metric in the space of smooth demand functions and finally propose a way to recover the best Slutsky approximation matrix function under a Frobenius norm. Our approach gives a geometric interpretation in terms of transformations of the Slutsky matrix or first order behavior of demand functions. As a result, a classification of the different kinds of violations of rationality is also provided. Comparative statics for a boundedly rational consumer is measured with our norm. Our approach is also suited to measure the mis-specification error from assuming a given form of rationality when the consumer’s behavior is actually not rational.

References


Appendix: Proofs

We begin this appendix with the following claim, which is an auxiliary result to be used in the sequel.

Proof of Claim 1

Proof. The problem is \( \min_{S'} \| S - S' \| \) subject to \( S'(z) \leq 0, S'(z) = S'(z)', S'(z)p = 0 \) for \( z \in Z \).

Under the Frobenius norm, the minimization problem amounts to finding the solution to
\[
\min_{S'} \int_{z \in Z} Tr((S(z) - S'(z))'[S(z) - S'(z)])dz
\]
subject to the stated constraints.

The objective function is strictly convex, because of the use of the Frobenius norm. This norm is also a continuous functional.

The constraint set \( M_{\mathbb{R}}(Z) \) is convex and closed. In fact, the cone of negative semidefinite matrices is a closed and convex set. Also, the set of symmetric matrices is closed and convex, and finally the set of matrices with eigenvalue \( \lambda = 0 \) associated with eigenvector \( p \) is convex. To see the last statement, let \( A(z)p = 0, B(z)p = 0 \), and let \( C(z) = \alpha A(z) + (1 - \alpha)B(z) \) for \( \alpha \in (0, 1) \). It follows that \( C(z)p = 0 \). Then \( M_{\mathbb{R}}(Z) \) is the intersection of three convex sets and is therefore convex itself. It is also useful to note that all three constraint sets are subspaces of \( \mathcal{M}(Z) \) and the intersection \( M_{\mathbb{R}}(Z) \) is itself a subspace of \( \mathcal{M}(Z) \).

Now we prove that not only the symmetric and the NSD constraints sets are closed but so is the set of all \( M_{\mathbb{R}}(Z) \). Any matrix function in the constraint set is a symmetric NSD matrix with \( p \) in its null space. Therefore, every sequence of matrix functions in the constraint set has the form \( D^n(z) = Q^n(z)\Lambda^n(z)Q^n(z)' \), where \( \Lambda^n(z) = \text{Diag}[\lambda^n_i(z)]_{i \in 1, \ldots, L} \) with ascending ordered eigenvalues functions. It follows that the eigenvalue function in position \( L, L \) is the null eigenvalue \( \lambda_L = 0 \), or the null scalar function. That is, imposing an increasing order the position 1, 1 is then held by \( \lambda^n_1(z) \leq \lambda^n_2(z) \leq \ldots \leq 0 \) where the order is induced by the distance to the null function using the Euclidean distance for scalar functions defined over \( Z \).

The matrix function \( Q^n(z) = [q^n_1 \cdots q^n_L] \) is the orthogonal matrix with eigenvectors functions as columns. For all \( D^n(z) \in M_{\mathbb{R}}(Z) \), \( \lambda^n_L = 0 \) is associated with the price vector \( q^n_L = p \) always, to guarantee that \( p \) is in its null space. The eigenvectors are defined implicitly by the condition \( D^n(z)q^n_L(z) = \lambda^n_L(z)q^n_L(z) \) with pointwise matrix and vector multiplication and \( q^n_L \perp p \) or \( (q^n_L, p) = 0 \) using the inner product for \( C^0(Z) \) for \( i = 1, \ldots, L - 1 \) and for all \( n \in \mathbb{N} \). Take any sequence of \( \{D^n(z)\}_{n \in \mathbb{N}} \) with \( D^n(z) \in M_{\mathbb{R}}(Z) \) for each \( z \in Z \), with limit \( \lim_{n \to \infty} D^n(z) = D(z) \). We want to show that \( D(z) \in M_{\mathbb{R}}(Z) \). It should be clear that any \( D^n(z) \rightarrow D(z) \) converges to a symmetric matrix function (the symmetric matrix subspace is an orthogonal complement of a subspace of \( \mathcal{M}(Z) \) (the subspace of skew symmetric matrix functions) and therefore, it is always closed in any metric space). It is also clear that \( D(z)p = 0 \) since \( \lambda^n_L \rightarrow 0 \) for all \( n \) and certainly \( \lambda^n_L \rightarrow 0 \) with the associated eigenvector \( q^n_L = p \) for all \( n \) and \( q^n_L \rightarrow p \). This condition, along with symmetry, guarantees that \( q^n_{L \perp} \rightarrow q \perp p \). Finally, the set of negative scalar functions is closed. Then, \( \lambda^n_{L \perp} \rightarrow \lambda(z) \) with \( \lambda_{L \perp}(z) = \min(0, \lambda_{L \perp}(z)) \). This is a negative scalar function by construction, since if \( \lambda_{i \neq L}(z) > 0 \) then \( \lambda_{i \neq L}(z) = 0 \). Then \( M_{\mathbb{R}}(Z) \) is closed. Observe that

\(^{21}\)These eigenvalue functions can be labeled because the domain \( Z \) is simply connected. The order admits crossing eigenvalue functions.
$M_{\mathcal P}(Z)$ is closed since it is in the intersection of three closed sets. In conclusion, since the Frobenius norm in $M(Z)$ is a continuous and strictly convex functional and the constraint set is closed and convex, the minimum is attained and it is unique.

Next, we present a claim that is auxiliary to establish Lemma 2, in the proof of Theorem 1.

Claim 3

Claim 3. The matrix $E^{\pi}(z)$ is pointwise orthogonal to $S(z)_{+}$. That is $Tr(E^{\pi}(z)'S(z)_{+}) = 0$.

Proof. By definition $S^{\sigma,\pi}(z) = S^{\sigma}(z) + E^{\pi}(z)$, with $E^{\pi}(z)$ a symmetric matrix such that $E^{\pi}(z)p \neq 0$ when $S^{\sigma}(p)p \neq 0$ and $E^{\pi}(z) = 0$ when $S^{\sigma}(p)p = 0$. Thus, $E^{\pi}(z)$ is always singular.

One can then write the direct sum decomposition of the set $A(z)$ of symmetric singular matrix functions with the property that $p'A(z)p = 0$ as follows: $A(z) = P(z) \oplus N(z)$ for all $z \in Z$, where

$$P(z) = \{ E^{\pi}(z) : Tr(E^{\pi}(z)pp') = 0 \quad \text{and} \quad E^{\pi}(z)p \neq 0 \quad \text{for} \quad E^{\pi}(z) \neq 0 \}$$

and

$$N(z) = \{ N(z) : Tr(N(z)pp') = 0, \quad N(z)p = 0 \}.$$ 

To see that this is a direct sum decomposition, first observe that $P(z) \cap N(z) = \{ 0 \}$, with 0 denoting the zero matrix function, by construction. Furthermore, any $A(z) \in A(z)$ can be written as a sum of $A(z) = E^{\pi}(z) + N(z)$ since $A(z)p = 0$ or (exclusive) $A(z)p \neq 0$, for $A(z) \neq 0$. Furthermore, $p'A(z)p = p'E^{\pi}(z)p + p'N(z)p = 0$ for any $E^{\pi}(z), N(z)$. Then the decomposition is precisely $A(z) = E^{\pi}(z)$ when $A(z)p \neq 0$ and $A(z) = N(z)$ when $A(z)p = 0$. Since every direct sum decomposition represents the sum of a subspace and its orthogonal complement, and $N(z)$ is a subspace in the space of symmetric matrix-valued functions, it follows that $P(z)$ is its orthogonal complement. In particular, since $S(z)_{+}p = 0$ and $Tr(S(z)_{+}pp') = 0$, it follows that $Tr(E^{\pi}(z)S(z)_{+}) = 0$, for $z \in Z$. $\square$
Slutsky Matrix Norms and Revealed Preference Tests of Consumer Behaviour

Victor H. Aguiar† and Roberto Serrano‡

This version: October 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 the discrete revealed preference (RP) approach, showing that testing rationality under both approaches with finite data is equivalent. We propose a way to classify the departures from rationality in a systematic fashion in finite data sets, by connecting violations of the Slutsky matrix properties to the length of demand cycles. We test our methodology in simulations, and apply it to an experimental data set. We find that violations of symmetry of the Slutsky matrix are more prevalent than violations of negative semidefiniteness, lending support to an old theoretical conjecture.

JEL classification numbers: C60, D10.

Keywords: consumer theory; rationality; Slutsky matrix norm; 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 incomplete, 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 rationality violations. 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 study. Here we propose a methodology that may help overcome at least one of these drawbacks (i.e., a useful decomposition of bounded

†Department of Economics, Brown University, victor.aguiar@brown.edu
‡Department of Economics, Brown University, roberto.serrano@brown.edu

*We thank Mark Dean, Xavier Gabaix, John Quah, Jesse Shapiro, and the participants of the Brown University Theory Lunch and SAET 2015 for useful comments and encouragement.
In a previous effort, Aguiar and Serrano (2015) answer the question of how far is a given behavior (demand function) 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 violations of 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 data set environment.

In contrast to our previous SMN approach, the “revealed preference approach” (RP) born in the seminal contributions of Afriat (1973) and Varian (1983), is specifically designed to deal with finite data sets. In fact, the most used and accepted measure is the Afriat efficiency index. The prevalence of the RP approach is due to its implementation simplicity for small data sets (very common) and its nonparametric nature. However, for instance, the Afriat and 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. Also, in some cases, it is not clear how these measures could be transported to an infinite data set framework.\(^2\)

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 pros and cons, 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,

- Is testing rationality in the SMN approach with finite data equivalent to testing it under the RP approach?
- How can we “summarize” the departures from rationality in a systematic fashion?

The aim of this paper is to provide an answer to these 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 data set. Conversely, if a data set can be rationalized in the sense of Afriat (1973), there exists an extension thereof that offers a rationalization in the sense of Hurwicz.

\(^1\)The computational implementation of revealed preference tests is arguably less of an issue due to improving processor power and due to the improvement of algorithms to implement them such as the work of Heufer & Hjertstrand (2015) with the Houtman-Maks index and specially the Afriat index (Smeulders et al., 2014).

\(^2\)This is a theoretical question of interest, since this is needed to understand the connection of the Afriat approach with the Richter environment, as Nishimura et al. (2013) makes it clear in their recent work.
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. 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 (2015), is also shown to apply to finite data sets. Finally, our results seem to hold well in simulations, and they 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 failures of negative semidefiniteness (weak axiom). We close our set of results by extending our coverage to demand correspondences.

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 develops a new class of interpolators of consumer choice data sets that have a minimal Slutsky norm and have desirable properties. Section 5 implements a simulation exercise to study the numerical behavior of the minimal Slutsky norm interpolators. Section 6 develops an application using experimental data from Ahn et al. (2014) and the tools developed in this paper. Sections 7 and 8 contain extensions: Section 7 generalizes the SMN approach and presents a convergence result between the Afriat efficiency index and a particular norm of the Slutsky error matrix, and Section 8 extends our previous analysis to nondeterministic environments, in particular regarding correspondences. Finally, we present a brief literature review in Section 9 and conclude in Section 10. Several more technical claims are collected in an appendix.

2 The Model: Testing Rationality with Limited Data Sets

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

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

Consider a demand function $x : P \times W \mapsto \mathbb{R}^L_+$, going from $Z = P \times W$, 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 each given price-wealth pair. The demand functions considered can be thought (without loss of generality) to belong to a closed and bounded set of demand functions $\mathcal{X}(Z) \subset \mathcal{C}^1(Z, \mathbb{R}^L)$. We assume that this space has the $\mathcal{C}^1(Z, \mathbb{R}^L)$ norm. Indeed, we also assume that this is a compact set. When we have two vectors of the same length say $v1,v2 \in \mathbb{R}^L$ we denote its inner product as $v1v2$.

**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 $k \in \{1, \cdots, K\}$.

Given that our interest is to connect with approaches based on the entire (infinite) functional...
data, we introduce the following theoretical construct. We say that \( x \in \mathcal{X}(Z) \) is the “true” DGDF if for any data set \( O^K \) or “sample” of individual choices for a given price-wealth situation \( (p^k, w^k) \in P \times W \), it is the case that \( x \in \mathcal{X}(Z) \) is the unique function \( x : P \times W \rightarrow \mathbb{R}^L_+ \) such that \( x^k = x(p^k, w^k) \) for all \( k \in \{1, \cdots , K\} \) 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. In addition, we are beginning our analysis with cases where \( O^K \) is such that there is no \( p^k = p^m \) for any \( k, m \in \{1, \cdots , K\} \) and \( k \neq m \). Later, in Subsection 8.1, we consider the case of demand correspondences and relax these assumptions.

We also assume that Walras’ law holds throughout, for all \( O^K, p^k x^k = w^k \) and all \( x \in \mathcal{X} \) (i.e., \( px(p, w) = w \) for all \( (p, w) \in P \times W \)). Behaviorally, when we face deterministic 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, which corresponds to the extensions of a finite data set \( O^K \). Every finite data set \( O^K \) has an associated set of functions that describe all possible extensions, which are also elements of \( \mathcal{X}(Z) \).

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

The following claim is trivial:

**Claim 1.** For finite \( K \) and any \( O^K \) its set of extensions \( \mathcal{X}^K(Z) \neq \emptyset \) is nonempty, when \( p^k \neq p^m \) for all \( k, m \in \{1, \cdots , K\} \) and \( k \neq m \).

For instance we can always find a smooth \( x \in \mathcal{X}(Z) \) that interpolates \( O^K \), a polynomial interpolator or B-splines interpolator among others.

### 2.1 RP and SMN Approaches to Test Rationality

The RP (revealed preference) approach to testing and measuring departures from rationality started with the groundbreaking 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 data set \( O^K \). If a data set \( O^K \) is compatible with the system of inequalities, we say that the data set 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 data set \( O^K \) departs from being rationalizable, many paths can be followed. In fact, there is a large literature of measuring departures from rationality in finite data sets.

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 \( t \in \{1, \cdots , T\} \).

\(^4\) We sometimes write \( x(t) = x^t \) and \( p(t) = p^t \) for all \( t \in \{0, \cdots , T\} \) and \( T \in \mathbb{N}_+ \). Recall also that we assume Walras’ law: \( p^t x^t = w^t \) for all \( t \).
Thus, if we can find a revealed demand cycle $CD,T$ 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: $x^tRx^{t-1}$ 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: $x^tR x^t$ whenever one can find a sequence $x^t R x^2 R \cdots x^n R x^t$, which we read as $x^t$ being "strictly revealed preferred" to $x^t$. It is clear that in $CD,T$ we have $x^T R x^{T-1} R x^{T-2} \cdots x^1 R x^0$, 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 $(p^k, x^k), (p^m, x^m) \in O^K$). For the sampling scheme that we are considering, namely $p^k \neq p^m$ for all $m, k \in \{1, \ldots, K\}$ and $m \neq k$, SARP stated for strict revealed preferences as is done here, is equivalent to the usual SARP (that rules out cycles for weak revealed preferences relations or weak inequalities).

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

**Definition 4. (Afriat’s index)** The Afriat efficiency index of a data set $O^K$, with at least one $CD,T$ with $T \geq 2$ is:

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

and zero otherwise.

Afriat’s index is such that $G \in [0,1]$ and researchers interpret its measure as a “loss of efficiency” or distance from a rational behavior benchmark: the larger it is, the larger gap between actual expenditure and expenditure on a nonchosen bundle that becomes chosen for efficiency” or distance from a rational behavior benchmark: the larger it is, the larger gap and zero otherwise.

The SMN (Slutsky matrix norm) approach, developed in Aguiar and Serrano (2015), 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. The Slutsky matrix $S \in M(Z)$ is assumed to belong to the space of matrix-valued functions and its $i,j$ entry is defined as usual as $s_{ij}(p, w) = \frac{\partial E_i(p,w)}{\partial p_j} + \frac{\partial E_j(p,w)}{\partial w_i} E_{ij}(p,w)$. 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 data set.

**Definition 5. (SMN approach)** The measure of bounded rationality for a given Slutsky matrix function $S = D_p x + D_w x'$ is

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

Where $||E||^2_W = \int_{z \in Z} ||W(z) F(z) W(z) ||^2_{KL,\mu(z)} dz$ is the weighted Frobenious norm with $W(z)$ a weighting matrix (positive semi-definite and symmetric) and $\mu$ a probability measure on $Z$ that is assumed to be measurable. The Slutsky regularity conditions are symmetry (denoted $\sigma$), singularity with prices as the eigen-vector associated with the null eigen-value or in short singularity in $p$ ($\pi$) and negative semidefiniteness or NSD ($\nu$). The two main results of Aguiar

---

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

---

5
and Serrano (2015) are that any Slutsky matrix function can be decomposed into two parts, a rational one and a behavioral error part. In fact, $S = S^r + E$ where $S^r = S^r,\pi,v$ is the projection of $S$ in the space of matrix functions that satisfy all three Slutsky conditions and $E$ is the sum of three orthogonal complement matrix functions namely $E = E^\pi + E^\nu + E^\sigma$ where $E^\pi$ is the violation of symmetry, $E^\nu$ corresponds to the violation in of in $p$ and $E^\sigma$ captures the violations of NSD. Finally the Slutsky matrix norm is decomposed in three parts $\|E\|_W^2 = \|E^\pi\|_W^2 + \|E^\nu\|_W^2 + \|E^\sigma\|_W^2$.

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

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

where $S(\bar{p}, \bar{w})$ is the Slutsky matrix evaluated at a reference point $(\bar{p}, \bar{w}) \in P \times W$ and $\{v^i\}_{i=0}^T$ is a collection of $v^i \in \mathbb{R}^L$ 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, Weak WARP, 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 7. (Real income path)** A real income path consists of both a wealth path $w: [0,b] \to W$, and a price path $p: [0,b] \to P$, having that $(w(\tau), p(\tau))$ is a piecewise continuously differentiable path in $Z$.

Thus, we are allowing for continuity of the derivative to fail at a countable subset of points.

Jerison and Jerison (1992) define a rising real income path whenever $(\frac{\partial w}{\partial \tau}(\tau), \frac{\partial p}{\partial \tau}(\tau))$ exist, leading to $\frac{\partial w}{\partial \tau}(\tau) > \frac{\partial p}{\partial \tau}(\tau)'x(p(\tau), w(\tau))$. A Ville cycle is a rising real income path such that $(w(0), p(0)) = (w(b), p(b))$.

We define a Ville Cycle as follows:

**Definition 8. (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 $x: [0,b] \to \mathbb{R}_+^L$ such that $x(0) = x(b)$ and $x \in C^2([0,b]; \mathbb{R}_+^L)$ for $S \geq 1$ and $p(\tau)^2 (\frac{\partial x}{\partial \tau}(\tau)) > 0$ almost everywhere in $\tau \in [0,b]$, for any piecewise continuous price path $p: [0,b] \to \mathbb{R}_+^L$ such that $p(\tau)x(\tau) = w(\tau)$. 

6
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 $p(\tau)\frac{\partial w}{\partial \tau}(\tau) > 0$ for all $\tau \in [0, b]$ since $\frac{\partial w}{\partial \tau}(\tau) = p(\tau) \frac{\partial w}{\partial \tau}(\tau) + \frac{\partial p}{\partial \tau}(\tau)x(p(\tau), w(\tau))$.

**Axiom 1. Ville Axiom of revealed preference (VARP).** A demand function $x \in X$ is said to satisfy VARP if it does not have a Ville cycle $C^V(S)$ 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).** We say that a demand function $x \in X$ satisfies Weak WARP if when we have $\bar{p}x(p, w) \leq w$ then it follows that $px(\bar{p}, \bar{w}) \geq w$.

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

Additional conditions that consumers are required to fulfill are the no money illusion property or homogeneity of degree zero that holds for a demand $x \in X$ if $x(\alpha p, \alpha w) = x(p, w)$ for all $\alpha > 0$. Finally, we state an additional axiom, that is equivalent to Weak WARP under homogeneity of degree zero, the Wald Axiom. The Wald Axiom works on a special kind of data set 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 that $p^k x^k = \bar{w}$ for all $k \in \{1, \ldots, K\}$. We state the Wald Axiom as follows: $p^k x(p^l, \bar{w}) < \bar{w}$ implies $p^l x(p^k, \bar{w}) > \bar{w}$.

### 3 Connecting the SMN and RP Approaches

We are ready to make the connection between the finite and the infinite data cases. Rationality is fully characterized in our environment (i.e., continuously differentiable demands) by both the Slutsky conditions and by SARP. The Slutsky or 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 $Z \subseteq P \times W$. Thus both approaches have been used as the basis of indices of bounded rationality. The RP approach is concerned with measuring departures from SARP when there is finite data, while the SMN approach, using functional data (i.e., if we observe the actual DGDF), measures 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 data set $O^K = \{p^k, w^k, x^k\}_{k=0}^K$ to the whole region $Z$. 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 data set that violates SARP, when extended, produces a Slutsky matrix function that is nonrational. Also, we are interested in the question of whether a demand function with a

---

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


nonrational Slutsky matrix will be able to generate a data set that violates SARP. The answer in the affirmative will be provided by our main result in this section.

Furthermore, since evidence shows that SARP (and more generally GARP) is routinely violated by a significant fraction of subjects in experimental studies, there is an increasing need to classify experimental data sets 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 the Slutsky norm 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 (2015), we have the following relations:

- (i) If Weak 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 data sets. 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. It shows that testing rationality of a limited data set \( O^K \) is equivalent in the SMN and the RP approaches:

**Theorem 1.** The next three statements are equivalent:

- A given data set \( O^K \) with \( K \geq 2 \) has at least a revealed demand cycle.
- All elements \( x \in X^K(Z) \) of the set of extensions of the data set either violate Weak WARP or violate the Ville axiom.
- All elements \( x \in X^K(Z) \) in the set of extensions of the data set have an associated Slutsky matrix norm error that is strictly positive.

Moreover, in the decomposition of the Slutsky matrix norm error we have:

When there are \( C^{D,T} \) for \( T \geq 3 \) and no \( C^{D,T} \) for \( T = 2 \):

- (i) For all extensions, \( \|E^\pi\| > 0 \),
- (ii) and for some extensions, \( ||E||^2 = ||E^\pi||^2 > 0 \), or equivalently \( ||E^\pi||^2 + ||E^\nu||^2 = 0 \).

Similarly, when there are no \( C^{D,T} \) for \( T \geq 3 \) and there are \( C^{D,T} \) for \( T = 2 \):

- (i) For all extensions, \( ||E^\pi||^2 + ||E^\nu||^2 > 0 \),
- (ii) and for some extensions, \( ||E||^2 = ||E^\pi||^2 + ||E^\nu||^2 > 0 \), or equivalently, \( ||E^\pi||^2 = 0 \).

---

7The converse of the following statements hold under (sufficient) technical regularity conditions on \( Z \), that require no additional assumptions on \( O^K \), but that require that \( \mathcal{X}(Z) \) is a space of smooth, Lipschitz continuous (at least in wealth) demand functions. The space \( Z \) is required to be path-connected, convex and open. It is clear that for any set of finite pair of prices and wealth we can find a space \( Z \) with the given characteristics such that the price-wealth pair belongs to \( Z \).
The structure of the proof is as follows. First, we connect revealed demand cycles with axioms of consumer theory. This is done in Lemmas 1 and 2. After establishing that connection, Lemma 3 links types of revealed demand cycles to properties of the Slutsky matrix norm. All these results (and their associated corollaries for the true DGDF), building blocks for the proof of Theorem 1, are presented in the rest of the current section.

Indeed, our first lemma concerns Weak WARP:

**Lemma 1.** A given data set $O^K$ with $K \geq 2$ has at least a revealed demand cycle of order $T = 2$ if and only if all elements $x \in X^K(Z)$ of the set of extensions of the data set 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 data set $O^K$ 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 extension of the set of extensions $X^T = \{x \in X| x^t = x(p^t), p^1x^1, \forall t \in \{0,1,2\}\}$. Then, it follows from Walras’ law that $p^tx^t = p^tx(p^t, w^t)$, $w^t$ for $t = 0, 1, 2$. It is straightforward to see that we have $p^2x(p^1, w^1) < w^2$ and $p^1x(p^2, w^2) < w^1$, which is a violation of Weak WARP. Thus the DGDF violates weak WARP.

Conversely, if all elements $x \in X^K(Z)$ of the set of extensions of the data set violate 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 $x \in X^K(Z)$ violate the aforementioned axiom.

**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) $x \in X$ violates Weak WARP.

**Proof.** The related proof is omitted.

Next, Lemma 2 asserts that, under Weak WARP, 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 $C^{V,b}_S$. $S \geq 1$.

**Lemma 2.** Let the DGDF satisfy weak WARP. Then:

(i) A given data set $O^K$ 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 $x \in X^K(Z)$ of the set of extensions of the data set have a Ville Cycle $C^{V(S),b}_S$ for $S \geq 1$ and $b > 0$.

(ii) If all elements $x \in X^K(Z)$ of the set of extensions of the data set have a Ville Cycle $C^{V(S),b}_S$ for $S \geq 1$ and $b > 0$, then one can observe a revealed demand cycle of order $T \geq 3$ in some $O^K$ of size $K \geq 3$.

If 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.
Proof. First, we prove necessity. Say we observe a finite data set \( O^K, K \geq 3 \), 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:

\[
\mathcal{X}^T = \{ x \in \mathcal{X}([0, b]) : \quad x^t = x(p^t, p^{t+1}) = x(\tau_t) \quad \forall t \in \{0, \cdots, T\} \quad \tau_t \in [0, b]\}.
\]

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 \( C^{V(S)} \) for \( S \geq 1 \). Assume to the contrary that there is at least one \( x \in \mathcal{X}^T \) that satisfies the Ville axiom. In particular, this implies that such a function can generate the data set \( O^K \) and at the same time exhibit no Ville cycles.

The absence of Ville cycles \( C^{V(S)} \) 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 \( z : x([0, b]) \to \mathbb{R} \) such that \( \partial z(\tau) \equiv \lambda(\tau)p(\tau)\partial x(\tau) \), with \( \lambda(\tau)p(\tau) = \nabla_z z(x(\tau)) \) and with the property that \( x(0) = x_0, x(1) = x^1, \cdots, x(T) = x^T \) 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 \( \tau_{t+1} - \tau_t = ||x^{t+1} - x^t||_{\infty} \), fix \( \tau_0 = 0 \), and define recursively \( \tau_t \) for \( t \geq 1 \) (i.e., define \( b = \sum_{t=0}^{T-1} ||x^{t+1} - x^t||_{\infty} \)).

Integrating along the path, we have \( \int_0^{T^*} \partial z(\tau) d\tau = \int_0^{T^*} \lambda(\tau)[p(\tau)\partial x(\tau)]d\tau = z(x(T)) - z(x(0)) = 0 \) by the observation that \( x^T = x_0 \). However, by linearity of the integral over an interval, we have

\[
\int_0^{T^*} \partial z(\tau) d\tau = \int_0^{T^*} \partial z(\tau) d\tau + \int_{\tau_1}^{\tau_2} \partial z(\tau) d\tau + \cdots + \int_{\tau_{T-1}}^{\tau_T} \partial z(\tau) d\tau = 0.
\]

Recall that by definition, \( \int_{\tau_{t-1}}^{\tau_t} \partial z(\tau) d\tau = z(x^t) - z(x^{t+1}) \). Now, since we have \( p^t x^t > p^{t+1} x^{t+1} \) for all \( t \in \{0, \cdots, T\} \) and \( T \geq 3 \) and by assumption there are no cycles \( C^{D,T} \) (\( T = 2 \)). Then there is a concave utility function \( z \) (in fact, strictly concave since we ruled out revealed indifference) such that \( p^t x^t > p^{t+1} x^{t+1} \) which implies that \( \sum_{t=0}^{T-1} \int_{\tau_{t+1}}^{\tau_t} \partial z(\tau) d\tau = \sum_{t=0}^{T-1} [z(x^{t+1}) - z(x^t)] > 0 \). This contradicts the result that \( \int_0^{T^*} \partial z(\tau) d\tau = z(x(T)) - z(x(0)) = 0 \) 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 \( C^{V(S)} \) for \( S \geq 1 \) and some \( b > 0 \).

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.

A related result follows, linking the previous lemma 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 (GDGF) \( x \in \mathcal{X} \) has a Ville Cycle \( C^{V(S)} \) for \( S \geq 1 \) and \( b > 0 \).

Proof. The related proof is omitted. \( \square \)

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:

**Lemma 3.** Any given data set \( O^K = \{p^k, w^k, x^k\}_{k=1}^K \) for \( K \geq 2 \) contains at least one revealed demand cycle if and only if all elements \( x \in \mathcal{X}^K \) in the set of extensions have an associated Slutsky matrix 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 data set \( O^K \) contains at least one cycle. To show this we use the contrapositive. When there are no revealed demand cycles \( C^{D,T} \) contained in \( O^K \), then \( O^K \) satisfies SARP. By Chiappori & 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, a contradiction.

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

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

We argue by contradiction. Assume that the data set contains a revealed demand cycle, yet at least one \( x \in \mathcal{X}_L(Z) \) has a Slutsky matrix norm \( ||E|| = 0 \). Then, pick any of such \( x \in \mathcal{X}_L(Z) \) with \( ||E|| = 0 \). Then, by Hurwicz and Uzawa (1971), we know there exists a locally nonsatiated continuous utility function defined on the region \( x(Z) \subseteq X, u : X(Z) \mapsto \mathbb{R} \) such that \( x(z) \) for \( z \in Z \), \( x \in \mathcal{X}_L \) is the unique maximizer of \( u \) subject to the budget constraint \( \{px(p, w) \leq w | (p, w) \in Z \subseteq P \times W \} \). 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 \( x \in \mathcal{X}_L(Z) \) 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 \( t \in T \). 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 \( t \in \{1, \ldots, T\} \). 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 \( x \in \mathcal{X}_L(Z) \) must have a positive Slutsky matrix norm \( ||E|| > 0 \).

If we have a data set 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 \( E^v = 0 \) 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 \( p^1 x_2 > p^1 x_1 \) with \( (p^0, x^0) = (p^2, x^2) \), implying a violation of Weak WARP.

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

**Proof.** The related proof is omitted.

Finally, the details about the decomposition of the Slutsky norm in the extensions are left to the reader. They can be easily shown appealing to the lemmata in this section and to the results in Aguiar and Serrano (2015).

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

Remark 2. The proofs of Lemma 3 and Corollary 3 use Hurwicz and Uzawa (1971) integrability theorem to define a utility function for an arbitrary compact set \( Z \subseteq P \times W \) instead of locally relying on a specific cycle. In other words, for the case of a data set that can be rationalized, there is at least one element \( x \in \mathcal{X}(Z) \) that has \( ||E|| = 0 \). Therefore, we can identify such an element \( x \in \mathcal{X}(Z) \) as the outcome of maximizing a unique concave utility function. Hence, we can think of an injection between the set of rational extensions \( \mathcal{R}(Z) = \{ x \in \mathcal{X}(Z) | ||E|| = 0 \} \) and \( \mathcal{U}(Z) \) where \( Z = \cup_{x \in \mathcal{R}(Z)} x(Z) \) a set of utility functions defined on the union of the 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.

To sum up, 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 data set satisfies SARP and 0 otherwise, and another binary test that yields 1 as its output when at least one of the elements \( x \in \mathcal{X}(Z) \) has a Slutsky matrix norm \( ||E|| = 0 \) and 0 otherwise. Then, both tests are equivalent for the same data set, and an output of 1 signifies that the data set \( O^K \) can be rationalized. More on such Slutsky norm tests follows.

4 Minimal Slutsky Norm Interpolators

Theorem 1 has the strong implication that, if the true DGDF violates one of the axioms of revealed preference, then all of the extensions of \( O^K \) for any \( K \) have a positive value for the corresponding Slutsky matrix norm. We shall pursue such extensions as an interpolation exercise given the limited data set, and most of the time, we shall seek interpolators that minimize the Slutsky norm. In doing so, the actual value of the decomposition may be also of interest, in order to help classify different kinds of bounded rationality. We shall close the section with a different approximation exercise: instead of seeking extensions in the set of all allowable demands, we shall approximate this functional space by means of Sieves spaces, which will be of help for computational purposes.

4.1 Interpolators as Extensions

The actual way of testing rationality in this interpolation exercise is captured by the following results. While their proofs are not hard or follow from previous work, it is important to state them explicitly.

**Proposition 1.** A SMN test for rationality in limited data sets corresponds to the solution to the optimization problem \( \alpha^* = \min_{x \in \mathcal{X}(Z)} ||E_x||^2 \), where \( \alpha^* = 0 \) means that the data set can be
rationalized, and it cannot otherwise, with $\alpha^* = \|E_\sigma\|^2 + \|E_\pi\|^2 + \|E_\nu\|^2 > 0$, the lower bound of the true value of $\|E\|^2$ for the underlying DGDF. Moreover, $\alpha^* = \|E_\sigma\|^2 > 0$ when there is a $C^{D,T}$ for $T \geq 3$ and no $C^{D,T}$ with $T = 2$, and $\alpha^* = \|E_\pi\|^2 + \|E_\nu\|^2 > 0$ when there is a $C^{D,T}$ for $T = 2$ and no $C^{D,T} \geq 3$.

Proof. Observe that the set of extensions $X^K(Z) = \{x \in X(Z) | x^k = x(p^k, u^k) \forall k \in \{1, \cdots, K\}\}$ 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 $E^x = 0$ for some rational demand $x^r \in X^K(Z)$ and since $\|E^x\| \geq 0$ for all $x \in X^K(Z)$, then $x^r = \arg\min_{x \in X^K(Z)} \|E^x\|^2$. When $O^K$ is not generated by a rational DGDF, then all $x \in X^K(Z)$ have $\|E\| > 0$ and by construction the value $\alpha^* > 0$ 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, Lemma 2, and Aguiar and Serrano (2015).

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

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

Following Aguiar and Serrano (2015), 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 $a' = a'(x) = g \circ s'$ where $g$ is a projection map and $s'(x, v) = D_p v + D_{v^i} [x^i]$, along direction $v \in C^1(Z)$.

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

Proposition 2. For every $O^K$, with $x$ its true DGDF and any extension $x^K \in X^K(Z)$, with $\|x - x^K\| \leq \epsilon^K$. 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 $\|a(x) - a(x^K)\| \leq \|a'\| \cdot \|x - x^K\|$.

Let $\|a(x) - a(x^K)\| = \eta$, then $\eta \leq \|a'\| \cdot \epsilon$, with $\|a'\| \leq \|g\| \cdot \|s'\|$.
Remark 4. We note that a sharp upper bound for the norm of the map \( a' \) can be estimated, as follows. Notice that \( \varphi \), being a projection map, has by definition \( ||\varphi|| \leq 1 \). Also, \( ||s'|| = \frac{||D_{w+D_{x'}x'}||}{||s'||} \leq \frac{||D_{w}v||}{||s'||} + \frac{||D_{w}x''||}{||s'||} \). With the norm of differential operators under the norm chosen, then \( \kappa \) \( ||s'|| \leq 1 \), \( ||D_{w}x''|| \leq \kappa \). Here, \( \kappa = ||x|| + ||D_{w}[x]|| \), because \( ||D_{w}x''|| \leq \frac{||D_{w}[x]||}{||s'||} \). The third property follows from Quah (2005), where it is proved that a demand that has a singular utility function subject to a linear constraint. By Kim and Richter (1986) and Mas-Colell (1974), this implies that there is a strictly concave locally non satiated utility function. The second property follows from Proposition 1, where it is proved that a demand that has a singular Slutsky matrix that is NSD can be generated by maximizing a twice continuously differentiable utility function subject to a linear constraint. By Kim and Richter (1986) and Mas-Colell (1974), this implies that there is a strictly concave locally non satiated utility function. The second property follows from the proof of Hurwicz & Uzawa (1971) that uses only symmetry of the Slutsky matrix, thus \( ||\text{argmin}_{x \in X} x_{K} || \text{ argmin}_{x \in \text{ } x_{K}} ||E^x||^2 \). The third property follows from the proof of Hurwicz & Uzawa (1971) that uses only symmetry of the Slutsky matrix, thus \( ||\text{argmin}_{x \in X} x_{K} || \text{ argmin}_{x \in \text{ } x_{K}} ||E^x||^2 \) that can generate the data by optimization. This means that while the demand system is integrable, in this case \( \phi \) is not a utility function and in fact its Hessian is not NSD.

Proof. The first property follows from Hurwicz & Uzawa (1971) because \( Z \) is assumed to be compact, Walras’ law holds and \( E_{K}^{x_{K}^{*}} = 0 \) which means that the Slutsky regularity conditions hold. Then it follows that \( x_{K}^{*} \) can be generated by maximizing a twice continuously differentiable utility function subject to a linear constraint. By Kim and Richter (1986) and Mas-Colell (1974), this implies that there is a strictly concave locally non satiated utility function. The second property follows from Quah (2005), where it is proved that a demand that has a singular Slutsky matrix that is NSD can be rationalized by preferences that are complete and non-transitive, but that are regular as defined by the author.\(^9\) The third property follows from the proof of Hurwicz & Uzawa (1971) that uses only symmetry of the Slutsky matrix, thus \( ||\text{argmin}_{x \in X} x_{K} || \text{ argmin}_{x \in \text{ } x_{K}} ||E^x||^2 = 0 \) to construct a scalar function \( \phi \) that can generate the data by optimization. This means that while the demand system is integrable, in this case \( \phi \) is not a utility function and in fact its Hessian is not NSD.

4.2 Restricted Approximation Results: Sieves Spaces

Next, we present a result that differs from the previous ones in one important respect. Namely, we keep the data set \( O^K \) fixed, and do not consider all its unrestricted extensions. Rather, we approximate the space of functions where we search for the minimal Slutsky norm interpolator demand function. Such an approximation result helps in the practical implementation of the results provided in Proposition 1. The test provided in that result is a variational problem that has no closed formed solution. To be able to implement this test numerically, we have to ensure that we can approximate the result with a limited functional space. In fact, consider an increasing sequence of sets of demand functions \( \{X_{h}\}_{h \in \mathbb{N}} \) such that \( X_{h} \subset X \) and \( X_{h_{1}} \subset X_{h_{2}} \) if \( h_{2} > h_{1} \) such

\(^9\)The upper contour set of \( x \), \( B(x) \) must be closed, convex and fulfill two additional (mainly technical) conditions explained in De nition 2.1 in Quah (2005).
that for any \( x \in \mathcal{X} \), there is a mapping \( \pi_h \circ v \in \mathcal{X}_h \) such that \( \pi_h \circ v \to v \) as \( h \to \infty \). This is usually called a Sieves space (e.g., think of a space of polynomial functions approximating our space of demand functions, in which \( h \) is the highest degree of the polynomial). We define analogously the allowable set of extensions of a data set \( \mathcal{X}^K_h(Z) = \mathcal{X}^K(Z) \cap \mathcal{X}_h \) under the \( h^{th} \) approximating space. We only consider the case that \( \mathcal{X}^K_h(Z) \) is nonempty, in other words, \( h \) is sufficiently large given this data set. The next result shows that the correspondence of Slutsky norm minimizers is “continuous at \( h = \infty \)”, or more formally, a consistent approximation.

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

**Proof.** The result follows from the same reasoning as Proposition 2. By the intermediate value theorem for the Gateaux derivative between Banach spaces, we have \( ||a(x) - a(x^h)|| \leq ||a'|| ||x - x^h|| \) with \( x^h \in \mathcal{X}^K_h(Z) \) and \( x \in \mathcal{X}^K(Z) \). By the definition of the Sieves approximating spaces, there exists \( \pi_h \circ x \in \mathcal{X}_h \) such that \( \pi_h \circ x \to x \), satisfying that \( ||x - \pi_h \circ x|| \to 0 \) when \( h \to \infty \). We know also by Wong (1984) (Lemma 2.2) that there is a constant \( C > 0 \) independent of the data set \( O^K \) such that for any \( x \in \mathcal{X}^K \) there exists a \( x^h \in \mathcal{X}^K_h \) satisfying \( ||x^h - x|| \leq C||x - \pi_h \circ x|| \).

We have established so far that any \( x \in \mathcal{X}^K \) can be approximated by a sequence with elements \( x^h \in \mathcal{X}^K_h \). Now, we notice that \( a \) is continuous under the \( C^0 \) norm on \( \mathcal{X} \) (claim (4) in the appendix), and that means that \( ||E_x|| : \mathcal{X}^K \to \mathbb{R}_+ \) is too a continuous mapping under these conditions, by Theorem 3.1 in Chen (2007)\(^{11}\) we conclude that \( \alpha_h^* \to \alpha^* \) as \( h \to \infty \), and by Wong (1984)\(^{12}\) (theorem 2.5) we also conclude that for any \( x^* \in \arg \min_{x \in \mathcal{X}^K(Z)} ||E_x||^2 \) there is a sequence \( x^{*,h} \to x^* \) as \( h \to \infty \) where \( x^{*,h} \in \arg \min_{x \in \mathcal{X}^K(Z)} ||E_x||^2 \).

**Remark 5.** The results in this section provide guidelines for practical work. If we have a large data set (with \( K \) large enough) we could use any interpolator, possibly a fast one for computational reasons, and it will be close to the true DGDF and the distance to the true Slutsky matrix norm will be small as well. In contrast, if \( K \) is small, we could use a large Sieves space (\( h \) large enough) to find an interpolator that has a Slutsky matrix norm that is close enough to the lower bound of the true value.

\(^{10}\)The environment in Wong (1984) is different from ours in that the author requires a convex and closed space of functions and we do not. But Lemma 2.2 does not use these properties. Note also that we consider only the point interpolation case.

\(^{11}\)Our results are nonstochastic, thus we only require conditions 3.1-3.4 in Chen (2007), namely identification, Sieves space convergence, continuity, and compact sieave space. The first condition holds vacuously because we always have \( ||E_x|| < +\infty \). Continuity of \( ||E_x|| \) holds also, and the conditions on Sieves space convergence and compactness of the sieves space hold by assumption.

\(^{12}\)The result in Wong (1984) (theorem 2.5) is proved in an environment with the space of functions that is convex, for example by requiring Lipchitz continuous differentiability in our demand functions \( \mathcal{X} \), but this assumption is not used in the convergence result. Its only role is to establish uniqueness. As we see from the argument in this proof, we can approximate any \( x \in \mathcal{X} \), including \( x^* \in \arg \min_{x \in \mathcal{X}^K(Z)} ||E_x||^2 \) there is an approximation \( x^h \in \mathcal{X}^K_h \), the fact that \( x^h = x^{*,h} \in \arg \min_{x \in \mathcal{X}^K(Z)} ||E_x||^2 \) is a consequence of continuity of \( ||E_x|| \) and the convergence properties of the sieves space increasing sequence \( \mathcal{X}_h \).
5 Simulation Study

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

- $x_1^{sf}(p, w) = \frac{w[p_1 + (1-c)p_2]}{2p_1(p_1+p_2)}$,
- $x_2^{sf}(p, w) = \frac{w[p_1 + (1-c)p_2]}{2p_2(p_1+p_2)}$,
- $x_3^{sf}(p, w) = \frac{wc_{p_1}}{p_1(p_1+p_2)}$.

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

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

Notice that $\mu$ is known to the experimenter because it is the sampling scheme generator process, the choice of this particular distribution is irrelevant and we do not need to estimate it. Our exercise is deterministic.

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

The budget share form of the demand system expressed in terms of logarithms of prices and logarithm of wealth, that we take as the true DGDF is: (i) $b_1^k(ln(p), ln(w)) = \frac{p_1 + (1-c)p_2}{2(p_1+p_2)}$. (ii) $b_2^k(ln(p), ln(w)) = \frac{p_1 + (1-c)p_2}{2(p_1+p_2)}$. (iii) $b_3^k(ln(p), ln(w)) = 1 - b_1^k(ln(p), ln(w)) - b_2^k(ln(p), ln(w))$, where the vector entry $ln(p)_i = ln(p_i)$. The Slutsky matrix at a point $(p, w)$ in its budget share elasticity form, can be computed for numerical stability purposes from the budget shares expressed in terms of logged prices and wealth as follows: $s_{ij}(p, w)p_{ij}/w = \frac{\partial b_i(ln(p), ln(w))}{\partial ln(p_j)} + \frac{\partial b_j(ln(p), ln(w))}{\partial ln(w)} + b_j(ln(p), ln(w))b_i(ln(p), ln(w)) - \delta_{ij}b_i(ln(p), ln(w))$, where $\delta_{ij} = 1$ if $i = j$ and zero otherwise.

Using corollary (3) we approximate the lower bound of $\|E^{sf}\|_W$ for the fixed sample $O^K$. The lower bound is $\alpha^K = \min_{x \in X^K(Z)} \|E^x\|_W$, where $X^K(Z)$ is the set of extensions of $O^K$ that
belongs to $X(Z)$. In practice, we cannot optimize over the whole $X^K(Z)$ numerically. First, we recast the problem in terms of budget shares and we call the functional space of budget shares corresponding to $X^K(Z)$ as $X^{K,b}(Z)$. We replace this space by the approximate sieve space $X^{K,b}_{X}(Z)$ that corresponds to the set of extensions that are B-splines or piece-wise polynomials of degree $h$. A multivariate B-splines is obtained by the tensor product of $L+1$ univariate B-splines of degree $h$, with the same degree for all dimensions, and they are evaluated at the data set. We obtain the matrix $B_l$ of dimension $K \times H$ given fixed knots and the given degree $h$ for each good $l$.

We can obtain the vector $b_l = \{h^f(p^k, w^k)\}^K_{k=1}$, for any element of the set of extensions of the data set $X^{K,b}_{X}(Z)$, by multiplying $B a = b$ for a given $a \in \mathbb{R}^H$ that is a vector of weights. In the same spirit, we can obtain the partial derivatives of any extension in $X^{K,b}(Z)$ by the following automatic procedure. We choose a variable that we want to differentiate with respect to, say $ln(p_1)$, derive the univariate B-spline corresponding to $ln(p_1)$ and obtain the tensor product with respect to the remaining B-splines (for $ln(p_2), ln(p_3), ln(w)$) to obtain a matrix $B_{l,ln(p_1)}$ of dimension $K \times H$, the partial derivative of the budget share $b_l$ with respect to $ln(p_1)$ at the data points, or the vector $b_{l,1} = \{\frac{\partial h_l(ln(p),ln(w))}{\partial ln(p_1)}|\}^K_{k=1}$ can be obtained by multiplying $B_{l,ln(p_1)}a = b_{l,a}$ for the same vector of weights as before. Using this procedure for all partial derivatives, we can obtain the Slutsky matrix evaluated at the data points, we name it $S^b(p^k, w^k)$ to denote its dependence of a given extension in the restricted space. We obtain the Slutsky error matrix $E^b(p^k, w^k)$ for each data point and we approximate the integral $||E^b||_W$, by the numerical approximation $||E^b||_W = \frac{1}{K} \sum_{k=1}^{K} ||E^b(p^k, w^k)||_{M,W}$ for any given $b^h \in X^{K,b}_{h}(Z)$. Use a Spectral Projection Gradient (SPG) large scale optimization algorithm to solve this optimization problem. The reader can notice that the minimization is with respect to the weights with a fixed $X^{K,b}_{h}(Z)$.

The effects of the number of observations and polynomial degrees in the accuracy of the estimation: We obtain a random sample of size $K \in \{50, 100\}$ i.i.d. draws from $\mu(z)$ and $x^f$ with its budget share $b^f$ such that we obtain: $O^K = \{p^{m,k}, b^{m,sf}|(p^{m,k})\}^K_{k=1}$ for each value of $K$. We fix a budget share space of extensions of $O^K$, $X^{K,b}_{h}(Z)$ for $h \in \{3, 5, 7\}$.

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

The effects of sample size on the decomposition of violations: A key feature of our methodology is the decomposition of the size of bounded rationality in its components ($\sigma, \pi, \nu$), by construction we can obtain the lower bounds of each type of violation. In this example the approximated values will be contrasted with the true values to show that the performance of our procedure is encouraging here too. For concreteness, we are going to present only the results for $h = 5$ for each element of the decomposition in a vector. The results of the simulation are provided in Table 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.

6 Application

We use in this section an experimental data set gathered by Ahn et al. (2014). We refer the reader to that paper for the experimental design and details about the data set. Here, we provide only the necessary information for this section to be self-contained. The data set 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 $p \cdot x = 1$ because prices and wealth are normalized so that this identity holds. 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 set-up). The main objective of this section is to measure the Slutsky matrix norm (in its unit-free version, as described in Aguilar and Serrano (2015)) 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 \( O^K \) when rationality is rejected have a positive SMN, and its decomposition is related to revealed demand cycles of different lengths and to the underlying Ville axiom and WARP. By the experimental design choice to normalize prices and wealth, we can always find demand extensions of the data set where wealth is fixed to \( \overline{w} = 1 \). Thus, we limit our attention to the conditional demand function \( x(p) = x(p, \overline{w}) \), for \( \overline{w} = 1 \). Then we compute the Budget Share Slutsky matrix conditional on a fixed wealth, let \( b_l(p, \overline{w}) = p_l x_l(p, \overline{w}) / \overline{w} \) for \( l = \{1, 2, 3\} \) such that \( \sum_{l=1}^{3} b_l(p, \overline{w}) = 1 \). The budget-share elasticity conditional Slutsky matrix (John, 1995) is \( S^h(p, \overline{w}) = [D_{ln}(p) b(p, \overline{w})] - [D_{ln}(p) b(p, \overline{w})] b(p, \overline{w})' \) where \( D_{ln}(p) b(p, \overline{w}) = D_{ln}(p) b(p, \overline{w}) - \text{Diag}(b(p, \overline{w})) \). Notice that \( S^h(p, \overline{w}) = \frac{1}{\overline{w}} \text{Diag}(p) S(p, \overline{w}) \text{Diag}(p) \) is the unit-free Slutsky matrix budget share from Aguilar and Serrano (2015), and \( S(p, \overline{w}) = D_p x(p, \overline{w}) - \frac{1}{\overline{w}} D_p x(p, \overline{w}) x(p, \overline{w})' \) is the Slutsky matrix for fixed wealth when homogeneity of degree zero holds because \( D_p x(p, \overline{w}) p + D_w x(p, \overline{w}) \overline{w} = 0 \). The importance of this, is that given the experimental set-up after fixing the wealth and due to the normalization of prices we can always find extensions that satisfy homogeneity of degree zero. For that reason we can focus on the conditional Slutsky matrix and in particular on its budget share elasticity form. All our results go through, but the decomposition of the norm will have its \( \pi \) component equal to zero:

\[
\|E\|_W = \|E^u\|_W + \|E^v\|_W \quad \text{with} \quad W(p, \overline{w}) = \frac{1}{\overline{w}} \text{diag}(p).
\]

The Slutsky matrix norm is computed only for the finite set of observations so \( \|E\|_W = \frac{1}{2} \sum_{k=1}^{50} \|E^h(p^k, \overline{w})\|_W^2 \), where \( E^h(p^k, \overline{w}) = E^{h \cdot o}(p^k, \overline{w}) + E^{h \cdot u}(p^k, \overline{w}) \) is the antisymmetric part of \( S^h(p^k, \overline{w}) \) and \( E^{h \cdot u}(p^k, \overline{w}) \) is the positive semidefinite part of the symmetrized \( S^h(p^k, \overline{w}) \). Using corollary (3) we approximate the lower bound of \( \|E\|_W \) for the fixed sample \( O^K \). The lower bound is \( \alpha^K = \min_{b_h \in \mathcal{X}_h} \|E^h\|_W \) where \( \mathcal{X}_h \) is the functional space of budget shares corresponding to \( \mathcal{X}_h \). We replace this space by the approximate sieve space \( \mathcal{X}_h \) that corresponds to the set of extensions that are B-splines or piecewise polynomials of degree \( h \). The compact domain \( Z \) is chosen to cover the observed prices in its interior. The degrees are fixed to be \( h = 5 \) for all \( L \) dimensions. The knots of the B-splines are fixed such that they provide 10 or more degrees of freedom per dimension. If there is no solution for the initial level of degrees of freedom, we increase them until reaching 15 degrees of freedom. The derivatives are obtained automatically taking advantage of the B-spline differentiation formulas.

The main difference of this approach with the traditional work of Haag et al. (2009) is that in the Ahn et al. (2014) data set, there is no measurement error and so we use interpolators of the data set instead of estimators of some statistic of the distribution of choices such as a conditional expectation of demand given prices and wealth. Any departure from rationality in this set-up
Table 3: Testing rationality using the Slutsky Matrix Norm estimates.

<table>
<thead>
<tr>
<th>Passes SARP</th>
<th>Symmetry $\sigma$</th>
<th>NSD $\nu$</th>
<th>Rationality $\sigma, \nu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes (12.98%)</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>No (87.02%)</td>
<td>8%</td>
<td>17%</td>
<td>5%</td>
</tr>
<tr>
<td>Total</td>
<td>19.9%</td>
<td>27.7%</td>
<td>17.3%</td>
</tr>
</tbody>
</table>

is the combination of an error of approximation of the derivative, that can be made arbitrarily small by increasing the “size” of the functional space $X^K,b(Z)$ and the behavior of the subjects. Due to the moderately large sample, and on the basis of our simulation study, the SMN that we compute here is taken as a precise measure of the lower bound of the SMN.

Testing Rationality: We implemented the test based on the SMN that corresponds to the solution and implementation of the class of estimators delineated in Proposition 1 and corollary 3. Recall that this exercise is deterministic but contains error due to the use of an approximate sieve space $X^K,b(Z)$. Thanks to the simulation study, we know that the “size” of the chosen functional space is satisfactory. In order to verify the quality of this approximation, we compare it here with the sharp and exact test of SARP. In particular, we compute the square of the Slutsky matrix norm $\alpha^K,h = \min_{b \in X^K,b,h(Z)} ||E^b||W$ for the Sieve space and the data set $O^K$ for each individual and fix three digits of precision. We denote $\alpha^K,h,j = ||E^b||W$ for $j \in \{\sigma, \nu\}$ and $b^* \in \arg\min_{b \in X^K,b,h(Z)} ||E^b||W$. We say that a data set $O^K$ satisfies the $\sigma$ property if $\gamma^K,h,\sigma = 0$ at the level of precision, else the data set does not satisfy this property. Similarly, $O^K$ is said to satisfy $\nu$ if $\gamma^K,h,\nu = 0$ at the level of precision. If $O^K$ satisfies both $\sigma$ and $\nu$, it is said to satisfy rationality.

The results in table (3) are encouraging, only about 5% of the subjects that do not pass SARP (8 subjects) are misclassified by the SMN based test with the given precision. The given precision was calibrated to obtain a 100% precision of the portion of experimental subjects that do pass SARP. The error is expected due to the use of the approximate functional space and the precision we have chosen, and we could make it smaller at a higher computational test. Regarding the behavior of the classification, both WARP and SARP are passed by the exact same subjects and the misclassification is higher in the case of the violations of NSD at the precision level. We recommend using the standard SARP test for a binary test. However, our approach allows to quantify departures, thereby shedding light on the kinds of bounded rationality exhibited.

The Stochastic Dominance of Violations of Symmetry w.r.t. NSD: Next, we compute the distribution of the violations of the symmetry and NSD conditions. We find an interesting stylized fact in this data set. The distribution of the violations of symmetry stochastically dominate the violations of negative semidefiniteness. The stochastic dominance of the violations of the symmetry condition can be visualized in Figure 1. The average Afriat index of this data set is $G = 0.947$.\textsuperscript{14} The mean of $\alpha^K,h,\sigma$ is 4.07% 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 error amounts to 0.0407 tokens for a compensated price change). The mean of $\alpha^K,h,\nu$ is 3.54%. This evidence supports the theoretical insights of Kihlstrom et al. (1976) that symmetry is more unstable than NSD.

\textsuperscript{14}The minimum Afriat efficiency is 0.74681 and the maximum is 1 – 12% of subjects satisfy SARP ($G = 1$). The standard deviation is 0.06.
The Empirical Relation of Afriat’s Index with SMN Measures: Using this framework, we provide the first study of the empirical 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 4). This finding 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 experimental subjects show a tendency to make errors in the same way for both σ and ν, even when quantitatively the errors in ν are of a lesser magnitude. In fact, 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
Table 4: Order -Spearman- correlation and Pearson Correlation of the Slutsky Matrix Norm estimates.

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, 2015). Finally, the ordinal correlation of the SMN and Afriat’s efficiency index surpasses 50%, again coexisting with a comparable but lower 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.

7 Generalized SMN and its Connections to the Afriat Approach through the Jerison-Jerison Index

An important question is how the Slutsky matrix norm relates numerically to Afriat’s index. The connection is established through the Jerison-Jerison modified Slutsky matrix 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. 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 (2015) for the general case). Our intent in this section is to focus on a generalized SMN norm that has as a special case the index of Jerison and Jerison (2012), and relate it to the Afriat index.

7.1 A Generalization of the Slutsky Matrix Norm

The space of matrix-valued functions $\mathcal{M}(Z)$ has some nice properties associated with the Slutsky regularity conditions $\sigma, \pi, \nu$. As shown in Aguiar and Serrano (2015) the spaces $\mathcal{M}(Z)_\sigma$ of symmetric matrices, $\mathcal{M}(Z)_{\pi,\pi}$ of symmetric matrices that are singular in prices, and $\mathcal{M}(Z)_{\sigma,\pi,\nu}$ of symmetric, singular in prices and negative semidefinite matrices are closed and convex subspaces under the topology induced by the Weighted Frobenius norm. The convexity does not depend on the norm and the closedness can be kept under more general set-ups and coarser topologies. However, we choose to keep the weighted inner product $\langle F, H \rangle_W$ in the space and we define the projections of any Slutsky matrix $S \in \mathcal{M}(Z)$ on the spaces of interest using it. With it, we have the same unique orthogonal decomposition as in Aguiar and Serrano (2015) with $S = S^\sigma + E$ with $E = E^\sigma + E^\pi + E^\nu$. However, it is possible to change the norm, in fact we can use any pseudo-norm or quasi-norm to compute our measure of bounded rationality.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha^{K,h,G}$ vs. $G$</td>
<td>SMN vs. Afriat’s</td>
<td>−0.517</td>
<td>−0.407</td>
</tr>
<tr>
<td>$\alpha^{K,h,\sigma}$ vs. $G$</td>
<td>SMN symmetry vs. Afriat’s</td>
<td>−0.504</td>
<td>−0.404</td>
</tr>
<tr>
<td>$\alpha^{K,h,\nu}$ vs. $G$</td>
<td>SMN NSD vs. Afriat’s</td>
<td>−0.502</td>
<td>−0.394</td>
</tr>
<tr>
<td>$\alpha^{K,h,\sigma}$ vs. $\alpha^{K,h,\nu}$</td>
<td>SMN symmetry vs. NSD</td>
<td>0.871</td>
<td>0.927</td>
</tr>
</tbody>
</table>
**Definition 9.** (Generalized Slutsky error semi-norm) A generalized Slutsky error norm is a mapping \( \Gamma = n \circ a \) such that is the composition of a \( n : \mathcal{M}(Z) \mapsto \mathbb{R}_+ \) seminorm on the space \( \mathcal{M}(Z) \) and the Slutsky error mapping \( a : \mathcal{A}(Z) \mapsto \mathcal{M}(Z) \) is defined as \( a(x) = E \).

We establish now that the JJ index can be obtained as a subcase of the Generalized Slutsky error seminorm. The JJ index has previously been written as being dependent on the whole Slutsky matrix \( S \), not only its error matrix \( E \). Ideally, we want a measure of violations of rationality that depends only on the departures of rationality and on nothing else (i.e. on the matrix \( E \) and not on the matrix \( S \)). This is indeed the case for the supremum of the JJ index. We next prove a nice feature of the discrete JJ index that, to the best of our knowledge, had not been previously pointed out.

Recall from Aguiar and Serrano (2015) that \( E = E^r + E^* + E^o \) is defined as the minimal perturbation matrix function that yields a rational Slutsky matrix function, decomposed in its symmetry, singularity and NSD error parts, \( E^r \) is the residual of Slutsky integrability, and \( S^{r^r} \) is the Slutsky matrix of the nearest rational demand function. Then:

**Lemma 4.** The JJ index is equal to

\[
\gamma(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w})) = \gamma(\{v_t\}_{t=1}^T, S^{r^r}(\bar{p}, \bar{w})) + \sum_{j \in \{r, r', \nu \}} \gamma(\{v_t\}_{t=1}^T, E^{r^r}(\bar{p}, \bar{w})).
\]

**Proof.** In the appendix.

This result shows that the JJ index can be decomposed additively into the JJ index of a rational demand, which is always nonpositive \( \gamma(\{v_t\}_{t=1}^T, S^{r^r}(\bar{p}, \bar{w})) \), because the eigen-vectors of \( S^{r^r} \) are all non positive and the JJ index is an average of quadratic forms, and an unsigned, sometimes nonnegative, part (i.e., it may be nonnegative for some cycle \( \{v_t\}_{t=1}^T \)). Observe also that an analogous decomposition holds for the computationally simpler decomposition \( S = S^r + E \), where \( \gamma(\{v_t\}_{t=1}^T, S^r(\bar{p}, \bar{w})) \) is always nonpositive and \( \gamma(\{v_t\}_{t=1}^T, E(\bar{p}, \bar{w})) \) is unsigned (sometimes –for some cycle– nonnegative).

We now look at the case of infinite data sets. 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 \( V^{T(r)} \) as the set of sequences of the type \( \{v^t\}_{t=0}^T \) where \( v^t \in V \subset \mathbb{R}^T \), \( V \) 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 \( t \in \{0, \cdots, T\} \), 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 \( V^{T(r),(r)} \) 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

\[
\mathcal{G}((\bar{p}, \bar{w}) + V^{T(r),(r)}) = \sup\{\{v^t, u^t\}_{t=0}^T \in V^{T(r),(r)} \mid \min_{t \in \{0, \cdots, T\}} \{p^t[x^t - x^t - 1]/|p^t x^t|\} |(\bar{p} + v^t, \bar{w} + u^t x^t) = C^{D,T}(\{v^t, u^t\}_{t=0}^T)\} \text{ (i.e., the supremum over the Afriat inefficiency index if we are allowed to sample price and wealth at will).}
\]

Now we establish the desired result regarding the supremum of the JJ index in \( V^{T(r)} \) and its relation with the Generalized Slutsky error matrix norm.

**Lemma 5.** Given a set of cycles \( V^{T(r)} \) and for a fixed \( T \geq 1 \) we have the supremum of the JJ index \( \sup\{\gamma(v, S(\bar{p}, \bar{w})) - \gamma(v, S^{r^r}(\bar{p}, \bar{w})) : v \in V^{T(r)}\} = \sup\{\gamma(v, S(\bar{p}, \bar{w})) : v \in V^{T(r)}\} \).
where $S^x(\bar{p}, \bar{w})$ is the Slutsky matrix of the nearest rational demand (under the Slutsky matrix norm). Moreover, $\sup \{ \gamma(v, S(\bar{p}, \bar{w})) : v \in V^{T(\epsilon_p)} \} = \sup \{ \gamma(v, E(\bar{p}, \bar{w})) : v \in V^{T(\epsilon_p)} \}$.

Proof. In the appendix.

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

7.2 Convergence Results

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

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

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

Proof. In the appendix.

Remark 6. Lemma 4 shows that locally the Afriat inefficiency index behaves as an index of errors in comparative statics due to bounded rationality with a very particular loss function: $\lim_{\epsilon_p \to 0} \epsilon_p^{-2} G((\bar{p}, \bar{w}) + V_w^{T(\epsilon_p, \eta)}) = \sup \{ \gamma(v, S(\bar{p}, \bar{w})) - \gamma(v, S^x(\bar{p}, \bar{w})), v \in V^{T(\epsilon_p)} \}$. This means that the Afriat inefficiency index at least locally can also be interpreted in a positive way (i.e., errors in comparative statics) instead of the usual welfare loss interpretation.

8 Demand Correspondences

The extension of the RP approach to data sets that allow for indifferences (with data generating demand correspondences –DGDC) instead of functions is straightforward. 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 (2015). Let $F$ be a closed and bounded set of demand correspondences.
Definition 10. (Data Generating Demand Correspondence -DGDC-). A demand correspondence \( X \in \mathcal{F} \) is a DGDC of \( O^K = \{p^k, w^k, x^k\}_{k=1}^K \) if \( x^k \in x(p^k, w^k) \) for all \( k \in \{1, \ldots, K\} \).

We say that \( X \in \mathcal{F} \) is the true DGDC if for any data set \( O^K \) or “sample” of individual choices for a given price-wealth situation \((p^k, w^k) \in P \times W\), it is the case that \( X \in \mathcal{F} \) is the Walrasian demand \( X : Z \to 2^\mathbb{R}_+ / \emptyset \) such that \( x^k \in X(p^k, w^k) \) for all \( k \in \{1, \ldots, K\} \) for all \( O^K \).

A true DGDC always exists. Of course, we have to impose some restrictions in order 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 \( X \in \mathcal{X} \) can be generated by maximizing a strict order \( \succ \in \mathbb{R}_+^3 \times \mathbb{R}_+^2 \) that is convex, monotone, and upper hemicontinuous. In practice, such restrictions are very weak in the case of finite data (i.e., to reject the assumption that \( X \) is compact-valued will require to observe all possible choices).

8.1 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.\(^\text{15}\) 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 \( d_X(x^\ast, X) < \eta \). The norm used in Mas-Colell (1974) is the Hausdorff distance defined on the set of graphs of the correspondences (\( d_X : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+ \)). Moreover, when dealing with the subset of demand functions, this metric is equivalent to the sup metric with associated norm \( ||\cdot||_\infty \).\(^\text{16}\)

It turns out that the density of smooth demands in the set of the demand correspondences that we consider is preserved for the case of boundedly rational demand. In fact, we first point out that we can approximate \( X \in \mathcal{X} \) with a smooth demand function \( x \in \mathcal{X} \cap \mathcal{C}_\infty \) arbitrarily closely.

Claim 2. For every \( \eta > 0 \) and \( X \in \mathcal{X} \) there exists a smooth demand function \( x \in \mathcal{X} \cap \mathcal{C}_\infty \) such that \( d_X(x, X) \leq \eta \).

Proof. First, we observe that:

(i) Thanks to De Blasi and Myjak (1986), we know that for every \( \nu_1 > 0 \) and \( X \in \mathcal{X} \), there exists a continuous demand function \( x \) such that \( d_X(x, X) < \nu_1 \).

(ii) Thanks to Azagra and Cepedello (2004), we know that for every \( \nu_2 > 0 \) and \( x \in \mathcal{X}(Z) \), there exists a \( x^* \in \mathcal{X} \cap \mathcal{C}_\infty \) such that \( ||x - x^*||_\infty \leq \nu_2 \).

Now, we observe that for \( \mathcal{X}(Z) \subseteq \mathcal{X} \), the metric \( d_X \) and the metric induced by \( ||\cdot||_\infty \) in \( \mathcal{X}(Z) \) are equivalent. Then, for every \( \eta > 0 \) and \( X \in \mathcal{X} \), there exists \( x \in \mathcal{X}(Z) \) such that

\(^{15}\)Mas-Colell (1974) proves this under the closed convergence topology and using a norm defined on the set of graphs of the correspondences.

\(^{16}\)In 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 \( z^k \in Z = P \times W \) and \( x^k \in A = \mathbb{R}_+^2 \), with the norm \( \rho(z^k, a^k, (x^m, a^m)) = \max\{|dz(z^k, z^m)|, |dA(a^k, a^m)|\} \) in \( Z \times A \), and the distance defined as the Hausdorff metric \( d_X(X, Y) = \max\{\sup_{a \in \mathcal{X}(Z)} d_X(z, X(z)) \max_{A \in \mathcal{Y}(Z)} d_A(a, Y(z)) \rho(a, b) \} \).
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 data set $O_{X,K} = \{p^k, w^k, \mathbf{X}(p^k, w^k)\}_{k=1}^K$ generated by $\mathbf{X} \in X$. Then there is a data set $O_{x,K} = \{p^k, w^k, x(p^k, w^k)\}_{k=1}^K$, where $x \in X \cap C^\infty$ and $d_H(O_{X,K}, O_{x,K}) \leq \eta$. In this sense, all the data sets $O_K$ that have a DGDC have a $\eta$−DGDF, that is, all data sets generated by sampling a DGDC have an arbitrarily close data set generated by sampling a DGDF.

However, an important 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 data set fails any of them, it cannot be rationalized, but to ensure that the data set can be rationalized, it must be the case that the test is passed for all $\eta > 0$. The practitioner can of course fix a small enough $\eta > 0$ and she will be sure that $\alpha^*_\eta > 0$ if the $O_K$ does not satisfy GARP because it is impossible that the data set has a DGDC that can be rationalized.

9 Literature Review

The current research is related to the work initiated by Chiappori and Rochet (1987), Matskin and Richter (1991) and culminated by 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 focuses on the realm of quantifying departures from rationality. In particular, we provide a minimal Slutsky norm interpolator of any finite data set, in the form of a continuously differentiable demand. When the data set satisfies SARP, then the minimal Slutsky norm is zero and the interpolator can be generated by maximizing a twice continuously differentiable utility function as in Hurwicz and Uzawa (1971). Our result provides a more general sense of integrability, since if the data set fails SARP but satisfies WARP (known also as Weak WARP), the minimal Slutsky norm interpolator can be generated by regular preferences in the sense of Quah (2005), that are nontransitive (i.e., nonrational). In fact, it provides a generalized form of integrability for cases that are less stringent than rationality (when at least one property of the Slutsky regularity conditions hold). The recent work of Nishimura et al. (2013) provides a bridge between the revealed preference approach for limited data sets and the Richter congruence axiom (1996) for complete data sets. Here we connect the limited data set environment with the results of Hurwicz and Uzawa (1971) and Quah (2005).

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. However, our work differs fundamentally from Jerison and Jerison (2012), among other things, in that we assume a finite data set of prices and observed choices, while the former authors assume an infinite amount of observations. 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), Apesteguia & Ballester (2015) and others. Our measure differs from this tradition in that ours is an objective measure of bounded rationality, in contrast with the more common welfare or subjective measures of bounded rationality that implicitly assume that the consumer wants to optimize, but fails to do it. In contrast, the SMN
is concerned with failures in doing comparative statics analysis for any given consumer behavior when we assume rationality (possibly contrary to the fact).

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

10 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) The violations of rationality identified in the SMN decomposition correspond to demand cycles of different length in the RP approach; (iii) Locally, a quadratic form of the Slutsky error norm converges to zero proportionally to Afriat’s inefficiency index. Our unified approach is smooth and easy to implement. 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 tested it in simulations and 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 its NSD. To the best of our knowledge, this is the first evidence brought forward on the claims of the higher robustness of WARP vis-à-vis the Ville axiom, as already suggested in Kihlstrom et al. (1976).

References


11 Appendix:

11.1 Continuity and Gateaux Differentiability of the Slutsky error map

Claim 3. The map $s: X(Z) \mapsto M(Z)$ defined as $s(x) = S$ is continuous.

Proof. First, we will prove that $D_p : X(Z) \mapsto M(Z)$ and $D_w : X(Z) \mapsto C(Z, \mathbb{R}^L)$ are not only closed linear operators, but are also continuous maps. In general, differential operators are closed but not continuous. However, in this specific domain, $D_p, D_w$ are defined everywhere by assumption, additionally $D_p$ and $D_w$ are closed operators, and finally $X(Z), M(Z)$ are Banach spaces with the norms $\| \cdot \|_{C_1}$ and $\| \cdot \|$ respectively, and so is $C(Z, \mathbb{R}^L)$, the space of continuous functions $f : Z \mapsto \mathbb{R}^L$ with supremum norm $\| \cdot \|_{\infty,L}$. Then, by the closed graph theorem, we can conclude that $D_p$ and $D_w$ are continuous maps.

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

Claim 4. The map $a: X(Z) \mapsto M(Z)$ defined elementwise as $a(x) = S - S'$ is continuous.

Proof. The continuity of the map $a$ follows directly from the continuity of the Slutsky map $s$ and the continuity of the projection maps that generates $S'$. Formally, $a = \varrho \circ s$ where $\varrho = \varnothing = (\varnothing - \rho_v \circ \rho_x \circ \rho_o)$ is a projection mapping, $\varnothing$ is the identity map in $M(Z)$ and $\rho_o$ is the projection on the space of symmetric matrices, $\rho_v$ is the projection in the space of symmetric matrices that are singular with prices in its eigen-space associated with the null eigen-value, and $\rho_p$ is the projection on the space of symmetric, singular and negative semidefinite matrices. By Claim 3, we know that $s: X(Z) \mapsto M(Z)$ is continuous. It remains to show that the projection maps are indeed continuous. For this we need that the range of the projection map is a closed subspace under the metric induced by the norm of $M(Z)$. In Aguiar and Serrano (2015) this is proven to be the case for each projection, in fact each range is convex and closed. It follows that $a$ is continuous.

Claim 5. The map $a$ derivative is $a' = \varrho \circ s'$.

Proof. The derivative of the mapping $s$ is $\frac{d}{dt}[s(x + tv)]_{t=0} = D_p v + D_w [x'v']$ and $\frac{d}{dt}[s(x + tv)]_{t=0} = D_w [vv']$ since $s(x + tv) - s(x) = D_p x + tD_p v + tD_w [x'v'] + t^2 D_w [vv'] + D_w xx' - D_p x - D_w xx'$. The derivative of $\varrho$ is the same $\varrho$ since it is linear. By the Chain rule of the Gateaux differential we have that $a' = \varrho \circ s'$.

11.2 Proof of lemma (4)

Proof. Aguiar and Serrano (2015) prove that we can decompose any observed Slutsky matrix in two components $S = S'' + E' + E$ locally. Then we can write $\gamma(\{v_1\}_{t=1}^T, S) = \gamma(\{v_1\}_{t=1}^T, S'' (\bar{p}, \bar{w}) + E' (\bar{p}, \bar{w}) + E(\bar{p}, \bar{w}))$ at the point $(\bar{p}, \bar{w})$. Then we observe that the JJ index is linearly separable $\frac{1}{T} \sum_{t=0}^T e'_t S (\bar{p}, \bar{w}) [v_t - v_{t-1}] = \frac{1}{T} \sum_{t=0}^T e'_t [S'' (\bar{p}, \bar{w})] [v_t - v_{t-1}] + \frac{1}{T} \sum_{t=0}^T e'_t [E' (\bar{p}, \bar{w}) + E(\bar{p}, \bar{w})] [v_t - v_{t-1}]$ so we can write $\gamma(\{v_1\}_{t=1}^T, S(\bar{p}, \bar{w})) = \gamma(\{v_1\}_{t=1}^T, S'' (\bar{p}, \bar{w})) + \gamma(\{v_1\}_{t=1}^T, E' (\bar{p}, \bar{w}) + E(\bar{p}, \bar{w}))$. □
11.3 Proof of lemma (5)

Proof. Notice that by subadditivity of the supremum \( \sup_{v} \gamma(v, S(p, \pi) - S^{x^r}(p, \pi)) \leq \sup_{v} \gamma(v, S(p, \pi)) - \sup_{v} \gamma(v, S^{x^r}(p, \pi)) \) and by the properties of the JJ index (Remark 2 in Jerison and Jerison (2012)) we have \( \sup_{v} \gamma(v, S^{x^r}(p, \pi)) = 0 \) because \( x^r \) is a rational demand, thus \( S^{x^r} \) is symmetric and NSD. Then \( \sup_{v} \gamma(v, S(p, \pi) - S^{x^r}(p, \pi)) \leq \sup_{v} \gamma(v, S(p, \pi)) \). Now, we let \( \hat{S}(p, \pi) = S(p, \pi) - S^{x^r}(p, \pi) \), and we want to prove that \( \sup_{v} \gamma(v, \hat{S}(p, \pi)) \geq \sup_{v} \gamma(v, S(p, \pi)) \).

In fact, this inequality follows from the properties of the JJ index (Remark 2 in Jerison and Jerison (2012)), because the following two conditions hold: (i) \( \hat{S}(p, \pi) - S(p, \pi) = -S^{x^r}(p, \pi) \) is both symmetric and PSD. (ii) \( \hat{S}(p, \pi)'p = S(p, \pi)'p - S^{x^r}(p, \pi)'p = 0 \) because \( p'S(p, \pi) = 0 \) and \( p'S^{x^r}(p, \pi) = 0 \) by Walras’ law. Thus we conclude that \( \sup_{v} \gamma(v, S(p, \pi) - S^{x^r}(p, \pi)) \geq \sup_{v} \gamma(v, S(p, \pi)) \). Because of the previous two inequalities we establish the result. For the “moreover” part of the statement, we notice that \( E = S - S^{x^r} \) and also that \( S^{x^r}(p, \pi) \) has the Slutsky regularity conditions (in particular symmetry and NSD) and therefore \( \sup_{v} \gamma(v, S^{x^r}(p, \pi)) = 0 \). Also by construction \( -S^{x^r}(p, \pi) \) is symmetric and PSD and \( (S(p, \pi) - S^{x^r}(p, \pi))'p = 0 \) thus the same argument as before applies and we have \( \sup_{v} \gamma(v, S(p, \pi)) : v \in V^{(\epsilon)} = \sup_{v} \gamma(v, S(p, \pi) - S^{x^r}(p, \pi)) : v \in V^{(\epsilon)} \) by the additivity of \( \gamma \) established in Lemma 4.

11.4 Proof of proposition (5)

Proof. By Jerison and Jerison (2012) we know that \( \lim_{\epsilon \to 0} \epsilon^{-2} G_T((p, \pi) + V_T^{(\epsilon, \eta)}) = \sup_{V_T^{(\epsilon, \eta)}} \gamma(v, S(p, \pi)) / \pi : v \in V_T^{(\epsilon, \eta)} \). By Lemma 4 we know that \( \gamma(v, S(p, \pi)) - \gamma(v, S^{x^r}(p, \pi)) = \gamma(v, E(p, \pi)) \).

Finally by Lemma 5 we know that \( \sup_{v} \gamma(v, E(p, \pi)) / \pi : v \in V_T^{(\epsilon, \eta)} \) and also \( \sup_{v} \gamma(v, S(p, \pi)) - \gamma(v, S^{x^r}(p, \pi)) / \pi : v \in V_T^{(\epsilon, \eta)} \) = \( \sup_{v} \gamma(v, S(p, \pi)) / \pi : v \in V_T^{(\epsilon, \eta)} \). Therefore, \( \lim_{\epsilon \to 0} \epsilon^{-2} G_T((p, \pi) + V_T^{(\epsilon, \eta)}) = \sup_{v} \gamma(v, E(p, \pi)) / \pi : v \in V_T^{(\epsilon, \eta)} \).