Slutsky Matrix Norms and the Size of Bounded Rationality*

Victor H. Aguiar† and Roberto Serrano‡

This version: February 2014

Abstract

Given any observed demand behavior by means of a demand function, we quantify by how much it departs from rationality. Using a recent elaboration of the "almost implies near" principle, the measure of the gap is the smallest 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. Variants, examples, and applications are discussed, and illustrations are provided using several bounded rationality models.

JEL classification numbers: C60, D10.

Keywords: consumer theory; rationality; Slutsky matrix function; bounded rationality; "almost implies near" principle.

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:

- (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?

*We thank Bob Anderson, Mark Dean, Federico Echenique, Xavier Gabaix, Michael Jerison, Susanne Schennach, and the participants at the Brown Theory Lunch for helpful comments and encouragement.

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

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

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 such additive factor. Using Anderson (1986) "almost implies near" (AN) principle and its recent elaboration, developed by Boualem and Brouzet (2012), we establish that for every approximation of rational behavior (i.e., the observed demand function being "almost" rational), there exists a rational demand function such that the two Slutsky matrix functions are also close ("near" symmetry, singularity, and NSD). This result allows us to use the Frobenius norm of the correcting factor added to the Slutsky matrix function 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 norm, is simply the sum 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 two approaches. The first is a normalization method, through dividing the norm of the additive correcting matrix function by the norm of the Slutsky matrix function of the observed demand. The second translates the first norm into dollars, providing a monetary measure.

The rest of this paper is organized as follows. Section 2 presents the model. Section 3 goes over the "almost implies near" principle and applies it to our problem. Section 4 deals with the matrix nearness problem, and finds its solution, emphasizing its additive decomposition. Section 5 provides interpretations of the matrix nearness problem in terms of the axioms behind revealed preference and in terms of wealth compensations, and presents unit-independent measures. Section 6 presents several examples and applications of the result, including hyperbolic discounting, the sparse consumer model, and an econometric application to estimation of the Slutsky matrix with noisy data. Section 7 is a brief review of the literature, and Section 8 concludes. Some of the proofs of the more technical results are collected in an appendix.

2 The Model

Let $\tau \in \mathcal{T}$ be an element of the set of theories or models of behavior. Examples of τ include behavior derived by a certain utility function, and more generally, the class of rational consumer models, as well as the class of models satisfying the weak axiom of revealed preference. Consider a demand function $x^\tau : Z \mapsto X$, where $Z \equiv P \times W$ is the compact space of price-wealth pairs (p, w), $P \subseteq \mathbb{R}^L_{++}$, $W \subseteq \mathbb{R}_{++}$, 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 under a particular τ.

Moreover, assume that $x^\tau(p, w)$ is continuously differentiable and satisfies Walras' law. That is for $p \gg 0$ and $w > 0$ $p'x^\tau(p, w) = w$. Let the set of functions that satisfy these characteristics be $\mathcal{X} \subset \mathcal{C}^1$. Hence, define also $\mathcal{X}(Z) \subset \mathcal{C}^1(Z)$, 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 $Z \subset \mathbb{R}^{L+1}_{++}$, equipped with the norm $||f||_{C1} = max(\{||f_l||_{C1,1}\}_{l=1,\ldots,L})$, with $||f_l||_{C1,1} = max(||f_l||_{\infty,1}, ||\nabla f_l||_{\infty,L+1})$ where $f(z) = \{f_l(z)\}_{l=1,\ldots,L}$.[1]

Let $\mathcal{R}(Z) \subset \mathcal{X}(Z)$ be the set of rational demand functions $\mathcal{R}(Z) = \{x^r \subset \mathcal{X}(Z) | x^r(z) = x^r(p, w) \succeq x \quad \text{subject to} \quad p'x \leq w\}$ for some complete and transitive relation $\succeq : X \times X \mapsto X$.

Definition 1. Define for any $\tau \in \mathcal{T}$, the distance from $x^\tau \in \mathcal{X}$ to the set of rational demands $\mathcal{R} = \{x^r(p, w) | r \in R\} \subset \mathcal{X}$ by the "least" distance from an element to a set: $d(x^\tau, \mathcal{R}) = inf\{d_{\mathcal{X}}(x^\tau, x^r) | x^r \in \mathcal{R}\}$.

We shall refer to this problem of trying to find the closest rational demand to a given demand as the "behavioral nearness" problem. Observe that the behavioral nearness problem at this level of generality presents several difficulties. First, the constraint set $\mathcal{R}(Z)$, 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

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

partial differential equations system. To calculate analytically the solution to this program is computationally challenging.

Having noted these difficulties with the "behavioral nearness" problem, our next goal is to talk about Slutsky norms. Let $\mathcal{M}(Z)$ be the complete metric space of matrix-valued functions, $F : Z \mapsto \mathbb{R}^L \times \mathbb{R}^L$, equipped with the inner product $\langle F, G\rangle = \int_{z\in Z} Tr(F(z)'G(z))dz$. This vector space has a Frobenius norm $||F||^2 = \int_{z\in Z} Tr(F(z)'F(z))dz$. 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.[2] Consider the price change dp and a compensated wealth change $dw = dp'x^\tau(p, w)$. Then the Slutsky matrix function $S^\tau \in \mathcal{M}(Z)$ is defined pointwise: $S^\tau(z) = D_p x^\tau(z) + D_w x^\tau(z)x^\tau(z)' \in \mathbb{R}^{L\times L}$, with entry $s^\tau_{l,k}(p, w) = \frac{\partial x^\tau_l(p,w)}{\partial p_k} + \frac{\partial x^\tau_l(p,w)}{\partial w} x^\tau_k(p, w)$.

The Slutsky matrix function is well defined for all $f \in \mathcal{C}^1(Z)$. Restricted to the set of rational behaviors, the Slutsky matrix satisfies a number of regularity conditions. Specifically, when a matrix function $S^\tau \in \mathcal{M}(Z)$ 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 $\mathfrak{R}$, for short. The one-to-one relation between matrices satisfying $\mathfrak{R}$ and rational theories of behavior will be exploited to define a metric that represents $d(x^\tau, \mathcal{R}(Z))$ for every $z \in Z$.

Definition 3. For any Slutsky matrix function $S^\tau \in \mathcal{M}(Z)$, let its Slutsky norm be defined as follows: $d(S^\tau) = min\{||E|| : S^\tau + E \in \mathcal{M}(Z)$ having property $\mathfrak{R}\}$.

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

3 The "Almost Implies Near" Principle

Intuitively, there should be a close relationship between the "least" distance to the set of rational demand functions (the behavioral nearness problem) and the Slutsky matrix nearness problem just defined. In order to make this relationship explicit, we will make extensive use of Anderson (1986) "almost implies near" (AN) principle and its recent elaboration, developed by Boualem and Brouzet (2012).

We begin by establishing a technical claim, whose proof can be found in the appendix.

Claim 1. The set $\mathcal{X}(Z)$ is a compact subset of $\mathcal{C}^1(Z)$.

The "almost implies near" (AN) principle allows us to assert that for all $\epsilon > 0$ there exists $\delta > 0$ such that for all $x^\tau \in \mathcal{X}(Z)$ with $\mathfrak{R}$ being $\delta-$ (almost) satisfied by S^τ, one can find $x^r \in \mathcal{X}(Z)$ with associated $S^r \in \mathcal{M}(Z)$ having property $\mathfrak{R}$ such that $d_{\mathcal{X}(Z)}(x^\tau, x^r) < \epsilon$. In the name of the principle, the "almost" part refers to$\delta > 0$ (matrix nearness), and the "near" part to $\epsilon > 0$ (behavioral nearness).[3]

[2] Since Z is closed, we use the definition of differential of Graves (1956) that is defined non only in the interior but also on the accumulation points of Z.

[3] Instead of relying on nonstandard analysis, Boualem & Brouzet (2012) use functions between metric spaces to represent a property in Anderson's language, and a metric to represent his formulas. This treatment is also useful because it allows us to adapt our results in order to derive an explicit expression for $\epsilon(\delta)$ for an arbitrary $Z \subset P \times W$.

The matrix nearness problem allows us to represent property $\mathfrak{R}$ by a function a with the AN property, as defined next.

Definition 4. [[Boualem and Brouzet (2012)]] A function $a : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ (with $\mathcal{X}(Z)$ and $\mathcal{M}(Z)$ metric spaces) satisfies the "almost implies near" (AN) property at $C \in \mathcal{M}(Z)$, if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for every $x^\tau \in \mathcal{X}(Z)$, the inequality $d_{\mathcal{M}(Z)}(a(x^\tau), C) < \delta$ implies the existence of an element $x_0^r \in \mathcal{X}(Z)$ satisfying $a(x_0^r) = C$ and $d_{\mathcal{X}(Z)}(x_0^r, x^\tau) < \epsilon$.

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

$$a(x^\tau) = E = 0,$$

where 0 represents the zero matrix function of $L \times L$ dimension in the metric space $\mathcal{M}(Z)$, and $E = S^r - S^\tau$ denotes a solution of the program in the definition of the Slutsky norm. The analytical expression of the (unique) solution to such a problem, as well as its properties, will be derived in the sequel (Theorem 1).

We state a result that applies the "almost implies near" principle to our problem.

Proposition 1. [4] *For all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x^\tau \in \mathcal{X}(Z)$, the inequality $d_{\mathcal{M}(Z)}(a(x^\tau), 0) = ||E|| < \delta$ implies the existence of an element $x_0^r \in \mathcal{R}(Z)$ satisfying $a(x_0^r) = 0$ and $d_{\mathcal{X}(Z)}(x_0^r, x^\tau) = ||e||_{C1} < \epsilon$. Moreover if $d(S^\tau) \leq \delta$, then we have the bound*

$$\epsilon(\delta) = min_{x_o^r : S_o^r = S^r} ||x^\tau - x_o^r||_{C1}$$

Proof. The proof uses that the solution to the matrix nearness problem is $E = S^\tau - S^r$, as shown in Theorem 1. We want to show that $a(x^\tau) = S^\tau - S^r$ is AN at $0 \in \mathcal{M}(Z)$. By Claim 7, found in the Appendix, $a : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ is continuous. Additionally, the set $\mathcal{X}(Z)$ is compact under the differentiability assumption, by Claim 1.

Then we conclude (applying Proposition 3.1 in Boualem & Brouzet (2012)) that a is (AN) everywhere, i.e., $a(x^\tau) = S^\tau - S^r = C$ has the AN property for all $C \in \mathcal{M}(Z)$. In particular, a is AN at $0 \in \mathcal{M}(Z)$.

Moreover, it follows that $lim_{\delta \to 0} \epsilon(\delta) = 0$ (applying Proposition 2.6 in Boualem & Brouzet (2012)), which implies the bound $\epsilon(\delta) = min_{x_o^r : S_o^r = S^r} ||x^\tau - x_o^r||_{C1}$. $\square$

We underscore the fact that the AN property is stated for every$\epsilon > 0$, not necessarily arbitrarily small, and therefore, we are able to account for violations of rationality of any "size", where the size of the violation is made precise using the $\epsilon(\delta)$ function.[5]

A somewhat surprising, but direct consequence of Proposition 1 is presented below. That is, by focusing on a compact subset $Z \subset P \times W$, one makes the definition of behavioral nearness

[4] As suggested by Jerison & Jerison (1993), the proof is a special case of Anderson (1986), itself reworked in Boualem & Brouzet (2012), as already discussed.

[5] We can be now more specific on the importance of compactness of $\mathcal{X}(Z)$. Note that one can avoid requiring compactness and replace it with the condition that a is onto. More precisely, we need to solve the partial differential equation system $S^r(z) = D_p x^r(z) + D_w x^r(z) x^r(z)'$. The condition that the map a is surjective amounts then to guaranteeing existence of a solution of the PDE system. If a is onto, then it follows that $lim_{\delta \to 0} \epsilon(\delta) = 0$. In fact, if $\delta \to 0$ and a is onto, then $S^r(z) \to S^\tau(z)$, and it follows that $x^\tau(z) = x_o^r(z)$, the minimizer of $\epsilon(\delta)$ in the feasible set $\{x_o^r : d(a(x_o^r), 0) \leq \delta, \delta \to 0\}$. That is $x^\tau(z) \in \mathcal{R}(Z)$ and $S^\tau(z)$ has property $\mathfrak{R}$, leading to the desired $\lim_{\delta \to 0} \epsilon(\delta) = 0$. Then, applying Proposition 2.6 in Boualem & Brouzet (2012), we conclude that$a$ is (AN) at 0.

operational. Then the function $d_{\mathcal{X}(Z)}(x^\tau, x^r)$ is a distance between any two demand functions, induced by the norm of the complete metric space $\mathcal{C}^1(Z)$. Observe that for any arbitrary compact subset $Z \subset P \times W$, $\mathcal{C}^1(Z)$ and $\mathcal{M}(Z)$ are compact-valued sets, as they are the images through continuous mappings of a compact set. By the next result, Proposition 2, we are able to guarantee that the minimum is attained in $\mathcal{X}(Z)$ assuming only that the elements of $\mathcal{X}(Z)$ and their first-order derivatives are continuous. Therefore, we have a local metric for $x^\tau \in \mathcal{X}(Z)$ defined for each $z \in Z$ and $\mathcal{R}(Z)$ as $||e||_{C1} = d(x^\tau, \mathcal{R}(Z)) = min\{d_{\mathcal{X}(Z)}(x^\tau, x^r)|x^r \in \mathcal{R}(Z)\}$, where $x^r(z) = x^\tau(z) + e(z)$ with $p'e(z) = 0$ by Walras' law.

Formally, we have:

Proposition 2. *The infimum is attained in the distance from a behavioral demand $x^\tau \in \mathcal{X}(Z)$ to the set of rational demands $\mathcal{R}(Z)$. Equivalently, the "least" distance can be defined as $d(x^\tau, \mathcal{R}(Z)) = min\{d_{\mathcal{X}(Z)}(x^\tau, x^r)|x^r \in \mathcal{R}(Z)\}$.*

Proof. Note that $\mathcal{X}(Z) \subset \mathcal{C}^1(Z)$, and that $\mathcal{C}^1(Z)$ is a metric space with norm $||\cdot||_{C1,L} = ||\cdot||_{C1}$. Note also that$\mathcal{M}(Z)$ is a metric space with Frobenius norm $||\cdot||$. Furthermore, $\mathcal{X}(Z)$ is compact, by Claim 1, and property $\mathfrak{R}$ can be expressed with a continuous function $a : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ by Proposition 1. This allows us to apply Anderson's almost-near principle.

Consider the two programs:

Program (I), the *behavior nearness problem*:

$||e||_{C1,L} = min_{x^r}||x^\tau - x^r||_{C1,L}$

subject to

$x^r \in \mathcal{R}(Z)$.

Program (II), the *matrix nearness problem*:

$||E||^2 = min_{S^r \in \mathcal{M}_{\mathfrak{R}}(Z)} \int_{z\in Z} Tr([S^\tau(z) - S^r(z)]'[S^\tau(z) - S^r(z)])dz$

subject to

S^r having property $\mathfrak{R}$.

Here, as already noted, property $\mathfrak{R}$ stands for:

$S^r(z) \leq 0$

$S^r(z) = S^r(z)'$

$S^r(z)p = 0$.

Recall that $S^r \in \mathcal{M}(Z)$ has property $\mathfrak{R}$ if and only if $x^r \in \mathcal{R}(Z)$.

Applying Proposition 1, we can conclude that the behavioral nearness problem has at least one solution such that $||e||_{C1,L} < \epsilon(\delta)$ and $||E|| < \delta$. □

4 The Matrix Nearness Problem: Measuring the Size of Violations of the Slutsky Conditions

Having established the formal link between the solutions to the behavioral nearness and matrix nearness problems, we turn to the latter. In this section we provide in our main result the exact solution of 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.

We begin by reviewing some definitions.

It will be useful to denote the three regularity conditions of any Slutsky matrix function with shorthands. We shall use σ for symmetry, π for singularity with p in its null space ($p-$singularity) and ν 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, if $S = S^{\tau}$ (i.e., a Slutsky matrix function), then $S^{\sigma} = S^{sym}$. Equivalently, S^{σ} is the projection of the function S^{τ} 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.

Any symmetric matrix-valued function $S^{sym} \in \mathcal{M}(Z)$ can be pointwise decomposed into the sum of its positive semidefinite and negative semidefinite parts. Indeed, we can always write $S^{\sigma}(z) = S^{\sigma}(z)_{+} + S^{\sigma}(z)_{-}$, with $S^{\sigma}(z)_{+}S^{\sigma}(z)_{-} = 0$ for $z \in Z$, $E^{\nu} = S^{\sigma}(z)_{+}$ being positive semidefinite (PSD) and $S^{\sigma,\nu} = S^{\sigma}_{-}$ negative semidefinite (NSD) for all $z \in Z$. Thus, one can write $S^{\sigma}(z) = S^{\sigma}_{+}(z) + S^{\sigma}_{-}(z) = E^{\nu}(z) + S^{\sigma,\nu}(z)$ 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 $S^{\sigma,\nu} = S^{\sigma}_{-}$.

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.[6] In particular, S^{σ} can be diagonalized: $S^{\sigma}(z) = Q(z)\Lambda(z)Q(z)'$.[7] Here, $\Lambda(z) = Diag[\{\lambda_l(z)\}_{l=1,\ldots,L}]$, where $\Lambda(z) \in \mathcal{M}(Z)$, with $\lambda_l : Z \mapsto \mathbb{R}$ a real-valued function with norm $||\cdot||_s$ (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 $||\cdot||_s$ norm.[8] Let $Q(z) = [q_1(z) \ldots q_L(z)]$, where $Q \in \mathcal{M}(Z)$ and its columns $q_l \in \mathcal{C}^1(Z)$ are the eigenvector functions such that for $l = 1, \ldots, L$: $S^{\sigma}(z)q_l(z) = \lambda_l(z)q_l(z)$ 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 $S^{\sigma,\nu}(z) = S^{\sigma}(z)_{-} = Q(z)\Lambda(z)_{-}Q(z)'$ for $\Lambda(z)_{-} = Diag[\{\lambda_l(z)_{-}\}_{l=1,\cdots L}]$ with $\lambda_l(z)_{-}$ the negative part function for the $\lambda_l(z)$ function. We can write also $E^{\nu} = S^{\sigma}(z)_{+} = S^{\sigma}(z) - S^{\sigma}(z)_{-}$ for $z \in Z$, or $E^{\nu} = S^{\sigma}(z)_{+} = Q(z)\Lambda(z)_{+}Q(z)'$ with $\Lambda(z)_{+}$ defined analogously to $\Lambda(z)_{-}$. Finally, any matrix function that is singular with p in its null space will be denoted S^{π}, that is, $S^{\pi}(z)p = 0$.

We are ready to state our main result:

Theorem 1. *Given a Slutsky matrix $S^{\tau}(z)$ for $z \in Z$, the solution to the matrix nearness problem is $S^{r}(z) = S^{\sigma,\pi,\nu}(z)$, the negative semidefinite part of $S^{\sigma,\pi}(z)$ defined by:*

$$S^{\sigma,\pi}(z) = S^{\sigma}(z) + E^{\pi}(z),$$

where $S^{\sigma}(z) = \frac{1}{2}[S^{\tau}(z) + S^{\tau}(z)']$ and $E^{\pi}(z) = -\frac{1}{p'p}[S^{\sigma}(z)pp' + pp'S^{\sigma}(z) - \frac{[S^{\sigma}(z)p]'p}{p'p}pp']$.

[6]In fact, Kadison (1984) has shown that any matrix function $\mathbb{M}_L(U)$ with U a von Neumann algebra is diagonalizable. Let U be $L^{\infty}(Z)$, and notice that Z is a separable Hilbert space. Let Z be $\sigma-$finite measurable, a subset of a Borel algebra generated by closed rectangles in $\mathbb{R}^{L+1}_{++}$, then $L^{\infty}(Z)$ is a von Neumann algebra and all $F \in \mathcal{M}(C^1(Z)) \subset \mathbb{M}_L(L^{\infty}(Z))$ are diagonalizable. Observe that there is an injection from$\iota : C^1(Z) \mapsto L^{\infty}(Z)$, such that we can identify any element in $\mathcal{C}^1$ with the von-Neumann algebra with the supremum norm.

[7]As is standard, Λ represents the diagonal matrix of eigenvalues, and Q is an orthogonal matrix that lists each eigenvector as a column.

[8]The 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.

We elaborate at length on the different components of this solution 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 2. Its proof is in the appendix.

Claim 2. A 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* $min_A||S^\tau - A||$ *subject to* $A(z)p = 0$, $A(z) = A(z)'$ *for all* $z \in Z$ *is* $S^{\sigma,\pi}$.

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

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

with $\lambda \in \mathbb{R}^L$ and $U \in \mathbb{R}^L \times \mathbb{R}^L$. Using that the singularity restriction term is scalar ($\lambda' S^{\sigma,\pi}(z)p \in \mathbb{R}$), as well as the identity $Tr(A'B) = vec(A)'vec(B)$ for all $A, B \in \mathcal{M}(Z)$,[9]

one can rewrite the Lagrangian as:

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

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 $S^{\sigma,\pi}$, the pointwise first-order necessary and sufficient conditions in this convex problem (Euler Lagrange Equation) is:

$S^{\sigma,\pi}(z) = \frac{1}{2}[S^\tau(z) + S^\tau(z)' + \lambda p' - U + U']$;

$S^{\sigma,\pi}(z)p = 0$;

$S^{\sigma,\pi}(z) = S^{\sigma,\pi}(z)'$.

With some manipulation, one gets:

$S^{\sigma,\pi}(z)p = S^\sigma(z)p + \frac{1}{2}\lambda p'p - \frac{1}{2}Up + \frac{1}{2}U'p$.

Using the restriction $S^{\sigma,\pi}(z)p = 0$ we have:

$S^\sigma(z)p + \frac{1}{2}\lambda p'p - \frac{1}{2}Up + \frac{1}{2}U'p = 0$.

Then:

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

This result reduces the system of first order conditions to:

$S^{\sigma,\pi}(z) = S^\sigma(z) - \frac{1}{p'p}S^\sigma(z)pp' + \frac{1}{p'p}\frac{1}{2}[U - U']pp' - \frac{1}{2}[U - U']$;

$S^{\sigma,\pi}(z) = S^{\sigma,\pi}(z)'$.

Let $E^\pi(z) = -\frac{1}{p'p}S^\sigma(z)pp' + \frac{1}{p'p}\frac{1}{2}[U - U']pp' - \frac{1}{2}[U - U']$.

By imposing the symmetry restriction on $S^{\sigma,\pi}(z)$, it follows that $E^\pi(z)$ must be symmetric. Therefore,

$E^\pi(z) = -\frac{1}{p'p}pp'S^\sigma(z) + \frac{1}{p'p}\frac{1}{2}pp'[U' - U] - \frac{1}{2}[U' - U]$

Postulating that

$U - U' = \frac{2}{p'p}pp'S^\sigma(z)$,

we get

[9]The $vec(A)$ symbol stands for the vectorization of a matrix A of dimension $L \times L$ in a vector $a = vec(A)$ of dimension L^2 where the columns of A are stacked to form a. Observe that the symmetry restriction can be expressed in a sigma notation (entry-wise) but this matrix algebra notation help us to make more clear the use of the trace operator in the objective function.

$\frac{1}{p'p}\frac{1}{2}[U-U']pp' = \frac{1}{p'p}\frac{1}{2}[\frac{2}{p'p}pp'S^{\sigma}(z)]pp' = \frac{1}{p'p}[\frac{p[p'S^{\sigma}(z)p]p'}{p'p}] = \frac{1}{p'p}[\frac{p'S^{\sigma}(z)p}{p'p}pp']$
since $p'S^{\sigma}(z)p$ is a scalar.

Then, $S^{\sigma,\pi}(z) = S^{\sigma}(z) + E^{\pi}(z)$, where

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

along with the implied multipliers λ and U, satisfies all the first-order conditions of the problem. Since we can use arguments identical to those in Claim 2 –only not imposing NSD–, we know that the solution is unique. Hence, this expression describes the solution to the posed calculus of variations problem with the symmetry and singularity restrictions. The proof is complete. □

If $S^{\sigma,\pi} \leq 0$ then we are done, since it has property $\mathfrak{R}$ and minimizes $||E||^2$ (by Lemma 3. 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* $min_A||S^{\sigma,\pi} - A||^2$ *subject to* $A \in \mathcal{M}(Z)$ *having property* $\mathfrak{R}$.

Proof. Recall the matrix nearness problem: $min_{A\in\mathcal{M}(Z)}||S^{\tau} - A||^2$ subject to A satisfying singularity, symmetry, and NSD. This is equivalent, by manipulating the objective function to: $min_{A\in\mathcal{M}(Z)}||E^{\sigma} - E^{\pi} + S^{\sigma,\pi} - A||^2$. Writing out the norm as a function of the traces, and using the fact that E^{σ} is skew symmetric, while the rest of the expression is symmetric, we get that this amounts to writing: $min_{A\in\mathcal{M}(Z)}||E^{\sigma}||^2 + || - E^{\pi} + S^{\sigma,\pi} - A||^2$ subject to A having property $\mathfrak{R}$. This is in turn equivalent to: $min_{A\in\mathcal{M}(Z)}||E^{\sigma}||^2 + ||E^{\pi}||^2 - 2\langle E^{\pi}, [S^{\sigma,\pi} - A]\rangle + ||S^{\sigma,\pi} - A||^2$ subject to A having property $\mathfrak{R}$. Then, exploiting the fact that E^{π} and $S_+ = S^{\sigma,\pi} - A$ are orthogonal (as proved in Lemma 6), we obtain that the problem is equivalent to $min_{A\in\mathcal{M}(Z)}||E^{\sigma}||^2 + ||E^{\pi}||^2 + ||E^{\nu}||^2$ subject to A having property $\mathfrak{R}$.

Hence, since the objective function of the matrix nearness problem $||E||^2 = ||E^{\sigma}||^2 + ||E^{\pi}||^2 + ||S^{\sigma,\pi} - A||^2$, solving the program $min_{A\in\mathcal{M}(Z)}||E||^2$ subject to A having property $\mathfrak{R}$ is equivalent to solving $min_{A\in\mathcal{M}(Z)}||S^{\sigma,\pi} - A||^2$ subject to the same constraints. □

Now, the best NSD matrix approximation of the symmetric valued function $S^{\sigma,\pi}$ is $S^r = S^{\sigma,\pi,\nu}$. Then, the candidate solution to our problem is $A(z) = S^r(z)$ for all $z \in Z$. Notice that $S^r(z)$ 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 $S^r(z) = S^{\sigma,\pi,\nu}(z) = Q(z)\Lambda(z)_{-}Q(z)'$. Then it follows that $S^r(z)$ 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 $q_L(z) = p$ eigenvector. Then we have that $\lambda_L(z)_{-} = 0$ is also associated to the eigenvector p, and we can conclude that $S^r(z)p = 0$.

As just argued, $S^r(z)$ has property $\mathfrak{R}$, i.e., $S^r(z)$ is in the constraint set of the matrix nearness problem or Program II. We conclude that it is its 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. Indeed, as was evident in the previous proof, the objective function of the matrix nearness problem can be expressed as follows:

$$||E||^2 = ||E^{\sigma}||^2 + ||E^{\pi}||^2 + ||E^{\nu}||^2.$$

We should think of the three terms in this decomposition 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. 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 $S^{\tau}(z)$ satisfies property $\mathfrak{R}$ for all $z \in Z$, $S^{\tau}(z) = S^{\sigma}(z)$ and $E^{\sigma}(z) = 0$, $S^{\sigma,\pi}(z) = S^{\sigma}(z)$ and $E^{\pi}(z) = 0$, and $S^{r}(z) = S^{\sigma,\pi,\nu}(z) = S^{\sigma,\pi}(z)$ 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 5.1 below.

The violations of the property $\mathfrak{R}$ 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, $||E||^2 = ||E^{\sigma}||^2$.

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 $S^{\sigma}(z)$ can be used to bound $||e||^2_{C1}$ locally when NSD is violated and $E^{\sigma}(z)$ can be used to bound $||e||^2_{C_1}$ when symmetry is violated. Indeed, this is consistent with our solution to Program II. In this case $||E||^2 = ||E^{\sigma}||^2 + ||S^{\sigma}{}_{+}||^2$, where $max(\{\tilde{\lambda}_l(z)_+\}_{l=1,\ldots,L}) \leq ||S^{\sigma}{}_{+}||^2 \leq d \cdot max(\{\tilde{\lambda}_l(z)_+\}_{l=1,\cdots,L})$, with $d = Rank(S^{\sigma}(z)_+)$ (by the norm equivalence of the maximum eigenvalue and the Frobenius norm).

Remark 1. The strict convexity of the objective functional of Program II and the convexity of the constraint set suggest that the solution to Program II can be found by the alternating projection algorithm. Indeed, one can first project $S^{\tau}(z)$ on the set of symmetric matrices, then project the result on the set of singular matrices with p in their null space, and finally project this second result on the set of negative semidefinite functions. The alternating projection algorithm can only guarantee that $S^{r}(z)$ has property $\mathfrak{R}$, but it may not necessarily be the solution to the problem. However, in our case, this specific sequence of projections yields the solution because the procedure results in the additive decomposition of $||E||^2$ provided in Lemma 2.

Remark 2. Using Lemma 3, one can deduce the analytic expression of $E^{\pi}(z)$ as a projection on a convex set. The lemma says that$S^{\sigma,\pi}(z) = S^{\sigma}(z) + E^{\pi}(z)$ is also the nearest matrix function with p in the null space of $S^{\sigma}(z)$. Thus $E^{\pi}(z)$ must be the minimal matrix additive adjustment in the Frobenius norm such that $[S^{\sigma}(z) + E^{\pi}(z)]p = 0$. Then, for any fixed $z \in Z$ this problem is analogous to the matrix nearness problem of finding the nearest linear symmetric system. Defining the feedback error $r(z) = -S^{\sigma}(z)p$, it follows that $E^{\pi}(z) = \frac{r(z)p' + pr(z)'}{p'p} - \frac{(r^T(z)p)pp'}{(p'p)^2}$ (Dennis & Schnabel, 1979; Higham, 1989). The resulting matrix function $S^{\sigma,\pi}(z)$ is the projection of $S^{\sigma}(z)$ on the set of symmetric matrix functions with p in its null space as made precise in Claim 6.

We underscore that the compactness of$\mathcal{M}_{\mathfrak{R}}(Z)$ is inherited from the mild assumptions of continuity of the demand system and its derivatives if we limit ourselves to a compact set Z. Furthermore, with the supremum norm we guarantee that $S^{r}(z)$ is continuous. Indeed, the following is a property of the solution to our problem:

Claim 3. $S^r(z)$ is continuous.

Proof. This follows from the Theorem of the Maximum. Specifically, let $F : \mathcal{M}(Z) \mapsto \mathcal{M}(Z)$ be such that $F(\mathcal{M}(Z))$ has property $\mathfrak{R}$. This is a compact-valued correspondence with a closed graph. Also, F is continuous and the Frobenius norm $||\cdot||$ is a continuous functional. It follows that S^r is continuous. Alternatively, 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 $z \in Z$. □

5 Behavioral Interpretations of the Slutsky Matrix Nearness Norm

5.1 Connecting with Axioms of Revealed Preference

The Slutsky regularity conditions are implied by utility maximization, but they can also be derived from the axioms behind revealed preference. Roughly, each of the conditions can be related to an axiom. Some of these relations have been used by Jerison and Jerison (1992) in order to provide a behavioral interpretation to the Slutsky matrix distance from symmetry. Our aim is to generalize this link while providing a behavioral interpretation of the matrix nearness norm decomposition.

We briefly review connections between the different axioms. Since Houthakker (1950) it is known that, for the class of continuous demand functions, the strong axiom of revealed preference –SARP– implies that a demand can be rationalized. Hence, the Slutsky conditions are satisfied. Nonetheless, it is also known that SARP is indeed strong in the sense that it implies symmetry of the Slutsky matrix function and also implies WARP and therefore NSD of the Slutsky matrix function. 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. WARP implies that the Slutsky matrix function is NSD, and furthermore, the NSD and singularity in prices are equivalent to a weak version of the WARP (Kihlstrom et al., 1976). VARP is a differential axiom and does not imply SARP or WARP (Hurwicz & Richter, 1979). WARP does not imply VARP or SARP for dimensions greater than two.

A continuously differentiable demand function is said to be rationalizable if it fulfills SARP. However, we can also impose other weaker axioms to have the same result while making connections to the additive components of our Slutsky norm. In particular, VARP and WARP imply that a demand function is rationalizable. Finally, we can impose the Wald Axiom, homogeneity of degree zero and VARP, which also imply a rationalizable demand. Moreover, to appreciate our decomposition, the Slutsky symmetry condition is related to VARP, the singularity in prices is related to homogeneity of degree zero and the NSD is related to the Wald axiom.

Before stating the main result of this subsection, for completeness, it is useful to posit the axioms that we employ and their relevant implications for the class of demand functions that we are considering and for the associated Slutsky matrix functions. Our primitive is a member of the set of demand functions $\mathcal{X}(Z) \equiv \{x^\tau \in \mathcal{C}^1(Z) | p' x^\tau(z) \leq w\}$ with Z a compact set.

The first Slutsky condition (price is its left eigenvector) is given by the balance axiom or Walras' law.

Axiom 1. *(Walras' law) The first axiom requires that $p'x^\tau(p,w) = w$.*

We have that $x^\tau \in \mathcal{X}(Z)$ satisfies Walras' law if and only if its Slutsky matrix function $S^\tau \in \mathcal{M}(Z)$ has the following property: $p'S^\tau(z) = 0$ for $z \in Z$.

The second Slutsky condition (price is its right eigenvector or singularity in prices) is given by "no money illusion".

Axiom 2. *(Homogeneity of degree zero -HD0-) $x^\tau(\alpha z) = x^\tau(z)$ for all $z \in Z$ and $\alpha > 0$.*

It is easy to prove that $x^\tau \in \mathcal{X}(Z)$ satisfies HD0 if and only if $S^\tau(z)p = 0$ for $z \in Z$.

The symmetry of the Slutsky matrix function is given by VARP.[10] To state this axiom we need to define an income path as $y : [0,b] \mapsto W$ $(t \mapsto w)$ and a price path $\rho : [0,b] \mapsto P$. Let $(y(t), p(t))$ be a piecewise continuously differentiable path in Z. Jerison and Jerison (1992) define a rising real income situation as whenever $(\frac{\partial y}{\partial t}(t), \frac{\partial p}{\partial t}(t))$ exists, leading to $\frac{\partial y}{\partial t}(t) > \frac{\partial p}{\partial t}(t)'x^\tau(\rho(t), y(t))$. A Ville cycle is a path such that: (i) $(y(0), p(0)) = (y(b), p(b))$; and (ii) $\frac{\partial y}{\partial t}(t) > \frac{\partial p}{\partial t}(t)'x^\tau(\rho(t), y(t))$ for $t \in [0,b]$.

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

Hurwicz and Richter (1979) proved that $x^\tau \in \mathcal{X}(Z)$ satisfies VARP if and only if S^τ is symmetric.

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 is said to fulfill the Wald axiom when so do the whole parametrized family (for $w \in W$) of conditional demands. Formally, a demand function can be expressed as the parametrized family of conditional demands. That is: $x^\tau(p,w) = \{x^{\tau,w}(p)\}_{w \in W}$ where $x^{\tau,w} : P \mapsto X$ such that $p'x^{\tau,w}(p) = w$ for all $p \in P$.

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

The Wald axiom implies that $S^\tau \leq 0$ (John, 1995).

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 $z = (p,w)$ $\overline{z} = (\overline{p}, \overline{w})$: $\overline{p}'x^\tau(p,w) \leq \overline{w} \implies p'x^\tau(\overline{p}, \overline{w}) \geq w$.*

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 $x^\tau \in \mathcal{X}(Z)$ satisfies WARP if and only if $S^\tau \leq 0$ and $S^\tau(z)p = 0$.

These axioms interact in interesting ways with direct consequences on the properties of the Slutsky matrix function. VARP and Walras' law imply homogeneity of degree zero. WARP and

[10] 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.

Walras' law imply homogeneity of degree zero and the Wald Axiom. Ultimately, VARP, Walras' law and the Wald Axiom implies a rationalizable demand.

The relations to our three Slutsky components can be summarized as follows: If VARP holds then $E^{\sigma} = 0$, if the Wald Axiom holds then $E^{\nu} = 0$, if homogeneity of degree zero and Walras' law hold then $E^{\pi} = 0$. Finally, if WARP and Walras' law hold then $E^{\pi} = 0$ and $E^{\nu} = 0$.

We are ready to summarize the main point of this section in the following remark.

Remark 3. The square matrix nearness norm $||E||^2$ will be equal to zero if and only if x^{τ} satisfies VARP, is homogeneous of degree zero in prices and wealth, and satisfies the Wald Axiom.

Moreover,

(i) $||E||^2 = ||E^{\sigma}||^2$ only if VARP fails while either of the following two situations happens: (i) Walras' law or HD0 and the Wald Axiom hold. (ii) Warp holds.

(ii) $||E||^2 = ||E^{\pi}||^2$ only if HD0 and Walras' law fail while VARP and the Wald Axiom hold.

(iii) $||E||^2 = ||E^{\nu}||^2$ only if the Wald axiom fails while either of the following two situations happens: (i) Walras' law and VARP hold. (ii) HD0 and VARP hold.

We can also observe $||E||^2 = ||E^{\nu}||^2$ only if WARP fails while Walras' law and VARP hold, but as we saw we do not need the failure of WARP to be stronger than the Wald axiom in order to have this case. The first part of the remark follows from the Frobenius' theorem and from the equivalence of the Ville axiom, WARP and Walras' law or VARP, Wald Axiom and HD0 to the Slutsky conditions. Indeed if x^{τ} satisfies any of two groups of axioms, then it can be seen as rational and $||E||^2 = 0$. Conversely, if $||E||^2 = 0$ it follows that S^{τ} fulfills the Slutsky regularity conditions and that x^{τ} satisfies the axioms.

The second part of the remark can be explained as follows. If x^{τ} satisfies Walras' law and WARP, it follows that $E^{\pi}(z) = 0$ and $E^{\nu}(z) = 0$, leading to $||E||^2 = ||E^{\sigma}||^2$. Then, thanks to Jerison and Jerison (1992; 1996; 1993) we know that the degree of asymmetry of the skew-symmetric part of a Slutsky matrix function grows with the rate of real income growth along the worst ("steepest") Ville cycle. Along the same lines, due to Russell (1997), we know that the size of the skew-symmetric matrix is exactly the distance from integrability of x^{τ}. If x^{τ} satisfies the Ville axiom and homogeneity of degree zero, then it follows that $||E||^2 = ||E^{\nu}||^2$, which corresponds to the PSD part of $S^{\tau}(z) = S^{\tau}_{-}(z) + S^{\tau}_{+}(z) = S^{\nu}(z) + E^{\nu}(z)$ for $z \in Z$. Then it follows that the size of $||E||^2$ grows exactly according to the degree of violations of the differential form of WARP (or the Wald Axiom in this case). Finally, if x^{τ} satisfies the Ville axiom and the Wald Axiom but both the HD0 and Walras' law fail then $S^{\tau}(z) + E^{\pi}(z) \leq 0$, it follows that $||E||^2 = ||E^{\pi}||^2$; in this case, $E^{\pi}(z)p = -S^{\tau}(z)p$, which shows that the size of $E^{\pi}(z)$ grows in the same direction as the degree of violations of the differential version of the homogeneity of degree zero condition (or Walras' law due to symmetry of S^{τ}). In fact, by construction $||E^{\pi}||$ measures how far S^{τ} is from having the price vector in its null space (as its eigenvector associated to the null eigenvalue).[11] There are partial converse implications to those just presented: If $||E^{\sigma}|| = 0$ then VARP holds. If $||E^{\pi}|| = 0$ then Walras' law or HD0 hold. If $||E^{\nu}|| = 0$ and $||E^{\pi}|| = 0$ then WARP holds.

[11] The idea of measuring violations of axioms by translating their consequences to a metric space was done for decision making under uncertainty and expected utility in Russell (2003).

5.2 Slutsky Wealth Compensations

We now describe ways to obtain related quantifications, in wealth terms, of departures of rationality. This builds upon the ideas of Russell (1997). First, we define the set $B_p = \{q \in \mathbb{R}^L_{++} | q' x^\tau(p, \overline{w}) = \overline{w}\}$ that is, all the price vectors that belong to the budget hyperplane for a fixed wealth $\overline{w}$ and bundle $x^\tau(p, \overline{w}) = x^{\tau,w}(p)$. We also define a directional derivative for any function $f \in C^1(Z)$ with respect to prices in the direction of a vector $v \in \mathbb{R}^L$ as $D_{p,v} f(z) = D_p f(z) v$. Observe that $q \in B_p$, can be expressed as a function of $q : P \times \{\overline{w}\} \mapsto B_p$, and the following identity holds for any $q \in B_p$ and any $x^\tau \in \mathcal{X}(Z)$: $[x^\tau(p, \overline{w})' D_p q(p, \overline{w})] + q' D_p x^\tau(p, \overline{w}) = 0$. This identity is obtained under Walras' law and uses the definition of the set B_p (i.e. we differentiate $q(p, \overline{w})' x^\tau(p, \overline{w}) = \overline{w}$ with respect to p). We define the conditional Slutsky matrix function pointwise: $S^{\tau,w}(p) = D_p x^{\tau,w}(p) - \frac{1}{w} D_p x^{\tau,w}(p) p x^{\tau,w}(p)'$. It can be shown that all the results of the Slutsky matrix norms carry over to this conditional Slutsky matrix function (to see this, note that our derivations do not depend on the parameters p, w but only on the structure of the matrix).

Russell (1997) defines for price vectors $q, r \in B_p$, $|q' E^\sigma(z) r| = m_1$, where m_1 is twice the wealth compensation required by an agent (who fulfills homogeneity of degree zero and the weak axiom but not necessarily the Ville axiom –symmetry–) in order to move from r to q on the budget hyperplane instead of moving from q to r. This quantity is zero for the rational consumer. Our result builds upon this finding and extends it to the case of the three possible violations of the Slutsky regularity conditions.

Before stating the main result of this section, we need some intermediate definitions. For a given observed demand $x^\tau \in \mathcal{X}(Z)$ consider the following three associated demand functions in $\mathcal{X}(Z)$ (that satisfy Walras law) that have special properties, linked to the Ville Axiom, homogeneity of degree zero, and the weak axiom. (i) Let $x^s \in \mathcal{X}(Z)$ be a demand that satisfies the Ville axiom, with Slutsky matrix function equal to $S^\sigma \in \mathcal{M}(Z)$ (i.e., the symmetric part of the $S^\tau \in \mathcal{M}(Z)$). (ii) Define $x^h \in \mathcal{X}(Z)$ as the demand function that satisfies the Ville axiom, homogeneity of degree zero in prices and wealth, and its Slutsky matrix function is equal to $S^{\sigma,\pi} \in \mathcal{M}(Z)$ (i.e., the projection of S^τ on the subset of matrix functions that are symmetric and have p in its null space). (iii) Finally, we have $x^r \in X(Z)$, the rational demand function that has its Slutsky matrix function equal to $S^r \in \mathcal{M}(Z)$ (i.e., the projection of $S^\tau \in \mathcal{M}(Z)$ on the subset of rational Slutsky matrix functions). The existence of a function that satisfies (iii) is guaranteed by the result proved in proposition 1. In the same spirit, we can guarantee the existence of $x^s, x^h \in \mathcal{X}(Z)$ by suitable straightforward modifications of the AN principle. The proof is direct when we notice that the properties of a demand function satisfying the Ville axiom or homogeneity of degree zero can be expressed in an analogous way as the rational case.[12] Then, we can conclude that there exists an $x^j \in \mathcal{X}(Z)$ for $j = s, h, r$ such that we can write $x^\tau + e^j = x^j$, where $e^j \in \mathcal{C}^1(Z)$ is a residual function that has the property: $p' e^j(p, w) = 0$ and $||e^j||_{C1} = \epsilon(\delta^j)$.

Proposition 3. *For any triple of vectors $p, q, r \in B_p$, define:*

- *(i) $|p' E^\sigma(p, \overline{w}) r| = \frac{1}{2} m_1$ if only symmetry is violated;*

[12] We note that the Ville Axiom and homogeneity of degree zero can be expressed using a continuous map $a^j : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$, and by finding an appropriate $\delta^j \geq 0$ for $j = s, h$ such that $||S^\sigma|| < \delta^s$ and $||E^\pi|| < \delta^h$. Also note the closedness of the symmetric and the singular in p matrix function sets in $\mathcal{M}(Z)$ (in this setting this is sufficient to guarantee the compactness of these sets).

- *(ii)* $|q'E^{\pi}(z)p| = m_2$ *if only singularity is violated;*

- *(iii)* $|q'E^{\nu}(p,\overline{w})r| = m_3$ *if only NSD is violated.*

Then,

- *(i)* $m_1 \in \mathbb{R}$ *is the wealth compensation that a non rational consumer with wealth* $\overline{w}$ *will have to be given in order for her to accept a price change from* r *in direction* q *instead of a change from* q *in direction* r.

- *(ii) The quantity* $m_2 \in \mathbb{R}$ *corresponds to the compensation that a consumer that satisfies the Ville axiom but that does not fulfill homogeneity of degree zero must receive to accept a price change from* q *in direction* p.

- *(iii) And* $m_3 \in \mathbb{R}$ *is the difference in wealth compensations that must be paid in order to accept a movement from* q *in direction* r *between the* $\epsilon-$*closest rational consumer and that of a consumer that satisfies the Ville axiom, homogeneity of degree zero but that does not necessarily fulfill the Wald axiom (or the WARP).*

Proof. We proceed to prove the three different parts:

(i) For fixed wealth $\overline{w}$:

$|q'E^{\sigma}(p,\overline{w})r| = |\frac{1}{2}[q'S^{\tau}(p,\overline{w})r - q'S^{\tau}(p,\overline{w})'r]|$.

$|q'E^{\sigma}(p,\overline{w})r| = |\frac{1}{2}[q'D_p x^{\tau,w}(p)r - q'D_p x^{\tau,w}(p)p - q'D_p x^{\tau,w}(p)'r + p'D_p x^{\tau,w}(p)'r]$

Since $p, r \in B_p$ it follows that

$|q'E^{\sigma}(p,\overline{w})r| = |\frac{1}{2}[[x^{\tau,w}(p)'D_p q(p)r + x^{\tau,w}(p)'D_p q(p)p] - [x^{\tau,w}(p)'D_p r(p)q + x^{\tau,w}(p)'D_p r(p)p]]|$.

Observe that $D_p r(p,\overline{w})q = D_p r(p)q$ is the directional derivative of r with respect to prices along that direction and with the magnitude of $q \in B_p$. Then the quantity $m_1$roughly measures the difference between the compensation in wealth from a price movement r in the direction q rather than the reverse.

For parts (ii) and (iii), we need an intermediate result, that takes into account the fact that $q'x^{\tau} + q'e^j = w + q'e^j$ for $j = s, h, r$ as defined above. That is, $q, r \in B_p$ may no longer fulfill Walras' law for $\epsilon-$closest demands that satisfies certain axioms. However, for any $q \in B_p$ this identity is satisfied: $q'x^j = w + q'e^j$. This in turn implies that: $q(p,\overline{w})'D_p x^j(p,\overline{w}) = q(p,\overline{w})'D_p e^j(p,\overline{w}) - [D_p q(p,\overline{w})'x^{\tau}(p,\overline{w})]'$ or equivalently $q'D_p x^j = D_p(q'e^j) - [D_p q'x^j]'$ for $j = s, h, r$. Here, $D_p(q'e^j) = [D_p q'e^j]' + q'D_p e^j$.

(ii) $|q'E^{\pi}(p,\overline{w})p| = |-q'\frac{1}{p'p}[S^{\sigma}(p,\overline{w})pp' + pp'S^{\sigma}(p,\overline{w}) - \frac{[S^{\sigma}(p,\overline{w})p]'p}{p'p}pp']p|$. It is proved in the appendix that if Walras' law holds then $|q'E^{\pi}(p,\overline{w})p| = 0$. We also know, that for singularity to be the only failure of the Slutsky conditions, Walras' law and homogeneity of degree zero must fail together. In this case, $|q'E^{\pi}(p,\overline{w})r| \geq 0$. To see a wealth interpretation of this quantity, we will choose $q = p$ and $r = p$:

Then, $|p'E^{\pi}(p,\overline{w})p| = |p'S^{\sigma}(p,\overline{w})p|$. Moreover if the failure of Walras' law is independent of the level of prices $p'x^{\tau}(p,\overline{w}) = c < w$ for $c \in \mathbb{R}_{++}$, then it follows that $|p'E^{\pi}(p,\overline{w})p| = |w - x^{s,w}(p)'p| = m_2$. Where $m_2 = |w - c|$, that is, m_2 is exactly the absolute size of the violation of Walras' law.

Then m_2, quantifies the wealth extraction that this non rational consumer accepts at prices p. Also, we can derive the following measure (modifying the proof in the appendix accordingly): $|q'E^{\pi}(p,\overline{w})p| = |\frac{q'D_p x^{s,w}(p)p}{w}[w - p'x^{s,w}(p)]| \propto |w - c|$.

To complete this part of the proof we show the existence of $x^s \in \mathcal{X}(Z)$. It suffices to recall that the AN principle can be appropriately modified by letting $||E|| = ||E^\sigma|| \leq \delta$ and by noticing that the cone of symmetric matrix functions is closed, and it is contained in the compact set $\mathcal{M}(Z)$. Then, there exists an $\epsilon-$closest demand function $x^s \in \mathcal{X}(Z)$ with the desired characteristics.

(iii) $|q'E^\nu(p, \overline{w})r| = |q'S^{\sigma,\pi}(z)r - q'S^r(z)r|$ since $S^{\sigma,\pi}(z) = S^{\sigma,\pi,\nu}(z) + E^\nu(z) = S_+^{\sigma,\pi}(z) + S_-^{\sigma,\pi}(z)$ by the direct sum decomposition of the space of symmetric matrix functions. For fixed $\overline{w}$:

$|q'E^\nu(p, \overline{w})r| = |q'S^\sigma(p, \overline{w})r - q'S^r(p, \overline{w})r| = |[q'D_p x^s(p, \overline{w})r - q'D_p x^r(p, \overline{w})r]|$. We know that $q'E^\pi r = 0$. Then it follows that:

$|q'E^\nu(p, \overline{w})r| = [q'D_p x^{s,w}(p)r - q'D_p x^{r,w}(p)r] + [q'\frac{1}{w}D_p x^{r,w}(p)px^{r,w}(p)'r - q'\frac{1}{w}D_p x^{s,w}(p)px^{s,w}(p)'r]$

Notice that, $D_p x^{r,w}(p)p = D_p x^{s,w}(p)p + D_p e^{r,w}(p)p$. Note that $S^r = [D_p x^s + D_w x^s x^{s'}] + [D_p e^r + D_w e^r e^{r'} + D_w x^s e^{r'}] = S^\sigma + E^\nu$.

Since, $E^\nu p = 0$ it follows that $D_p e^{r,w} p = 0$ because $D_p e^{r,w} p + D_w e^{r,w} e^{r,w'} p + D_w x^s e^{r,w'} p = 0$ and $e^{r,w'} p = 0$ by construction.

Then, $|q'E^\nu(p, \overline{w})r| = |q'D_p x^{s,w}(p)r - q'D_p x^{r,w}(p)r| = |q'D_p e^{r,w}(p)r| = m_3$.

Finally, by using the identity for $q, r \in B_p$, and the fact that $x^s = x^h$ when Walras' law hold we obtain the desired result:

$|q'E^\nu(p, \overline{w})r| =$

$|[x^r(p, \overline{w})'[D_p q(p, \overline{w})r] - x^h(p, w)'[D_p q(p, w)r]] - [q(p, \overline{w})'D_p(q'e^r(\cdot))r - q(p, \overline{w})'D_p(q(\cdot)'e^h(\cdot))r]|$,

that is, the difference between the wealth compensation that has to be made for the consumer to accept a change from q with respect to prices in the direction r when she is rational versus when she satisfies the Ville axiom and homogeneity of degree zero but not necessarily the weak axiom. To this quantity, a correction term is subtracted that measures the difference between the product of q times the marginal change of $q'e^j = w - x^\tau$ for $j = h, r$ with respect to prices in the direction of r for both kinds of consumers (rational and Ville Axiom plus homogeneity of degree zero). Equivalently, $|q'E^\nu(p, \overline{w})r| = |[q(p, \overline{w})'D_p e^h(p, \overline{w})r - q(p, \overline{w})'D_p e^r(p, \overline{w})r]| = m_3$. That is, a measure of the difference between the marginal change of the correction term e^j in the direction r when initial prices are q between the two types of consumers. In particular, when $x^h(p, \overline{w}) = x^\tau(p, \overline{w})$ this simplifies to $|q'E^\nu(p, \overline{w})r| = |-q(p, \overline{w})'D_p e^r(p, \overline{w})r| = m_3$.

Of course, the existence of $x^r \in \mathcal{X}(Z)$ follows from our result in Proposition 2. By an analogous argument to part (ii) of this proof we modify the $E = -[E^\sigma(z) - E^\pi(z)] = -[S^\tau(z) - S^{\sigma,\pi}(z)]$ ($||E|| \leq \delta$) and note that the intersection of the cone of symmetric matrix functions with the set of matrix functions with p in its null space is closed and thus compact under our assumptions. Furthermore, the AN principle guarantees under these conditions that there exists an $\epsilon-$closest demand $x^r \in \mathcal{X}(Z)$ with the required properties. □

Remark 4. The extension of Russell's (1997) idea for using a Slutsky residual matrix to our case, which covers all three possible violations, makes heavy use of the AN principle, modified appropriately in each case to guarantee the existence of a "corrected" demand system that fulfills certain axioms. This extension comes at a cost. Indeed, the wealth compensation measure does not depend only on the primitive $x^\tau \in \mathcal{X}(Z)$, but it must incorporate corrections for price changes that may not belong to the budget hyperplane of the relevant demands. In geometric terms, one can think of the correction term as measuring the change in the cosine of the angle between the price vector q and the residual e^j with respect to prices in the direction r times the initial vector of prices q. This measure is converted in wealth (i.e., for fixed $p = \overline{p}$ and $\overline{w}$, $q'e = cos(\theta_{q,e})||q|| \cdot ||e||$ and $D_p(q'e) = D_p(cos(\theta_{q,e})||q|| \cdot ||e||)$ with the euclidean norm in $\mathbb{R}^L_{++}$).

The proposition serves mainly as a blueprint on how to compute measures to the size of bounded rationality that are expressed in wealth terms. These measures have the advantage of having an intuitive interpretation, in the sense that a rational consumer should have always for all $p, q, r \in B_p$ a measure of $q'E(z)r = 0$ where r may be equal to p, since $E(z) = 0$. The measures are imperfect, though: observe that there is a family of measures for each $q, r \in B_p$ and for each $\overline{w}$ that must be fixed for a particular application. In addition, this measure of the size of bounded rationality does not induce a metric, as nothing prevents that there exists a $q, r \in B_p$ such that $q'E(z)p = 0$ when $E(z) \neq 0$.

Remark 5. Consider two consumers with demands $x^{\tau 1}(p, \overline{w})$ and $x^{\tau 2}(p, \overline{w})$ satisfying Walras' law and exhibiting identical violations of the Wald Axiom and homogeneity of degree zero (or violations of WARP). Suppose further that the first consumer violates the Ville axiom, but the second satisfies it. We write $x^{\tau 1}(p, \overline{w}) + e^{s1}(p, \overline{w}) = x^{s1}(p, \overline{w})$ and $x^{\tau 2}(p, \overline{w}) = x^{s2}(p, \overline{w})$. When forced to satisfy the Ville axiom, both have the same first order behavior for compensated wealth changes. In that case, for prices $q, p \in B^1_p \cup B^2_p$, assuming there is a q vector in the budget line of both consumers different from p,we have:

$2|q'E^{\sigma 1}(p, \overline{w})p| + |q'E^{\pi 1}(p, \overline{w})p| > |q'E^{\pi 2}(p, \overline{w})p|$. That is, $|m^{\tau 1}_1| + |m^{\tau 1}_2| > |m^{\tau 2}_2|$ since $m^{\tau 1}_2 = q'S^{\sigma}(p, \overline{w})p = m^{\tau 2}_2$ because both consumers fulfill Walras' law, and in that case $q'E^{\pi 1}(p, \overline{w})p = q'E^{\pi 2}(p, \overline{w})p = q'S^{\sigma}(p, \overline{w})p$ as shown in proposition 3. Also observe that $|m^{\tau 2}_1| = 0$ by construction.

In other words, the wealth measure of bounded rationality is larger for $x^{\tau 1}$ than for $x^{\tau 2}$. Observe that this can be concluded only when assuming the equality in first order compensated behavior of the "corrected" (Ville axiom) demand x^{s1} and $x^{\tau 2}$. But this assumption can be justified in an interesting way when we think of $x^{\tau 2}$ as a "minimally perturbed" version of $x^{\tau 1}$ (with Slutsky matrix function S^{σ}), such that $x^{\tau 2}$ is the demand of the first consumer when forced to satisfy the Ville axiom.

5.3 Normalizations and Relative Matrix Nearness

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 faced by 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).

Therefore, we next propose a relative matrix nearness norm that, while keeping most of the features of the absolute measure of bounded rationality, is unit-free.

Definition 5. For any non null Slutsky matrix function $S^{\tau} \in \mathcal{M}(Z)$, let its relative Slutsky norm be defined as follows: $\rho(S^{\tau}) = \frac{d(S^{\tau})}{||S^{\tau}||}$, where $d(S^{\tau}) = ||E||$ is the absolute matrix nearness distance to rationality.

Observe that we have excluded from the definition the case of null Slutsky matrix functions (i.e., $S^\tau = 0 \in \mathcal{M}(Z)$). This, however, is just a technicality since the null S^τ satisfies property $\mathfrak{R}$, and hence, one can postulate $\rho(0) = 0$.

Claim 4. The $\rho : \mathcal{M}(Z) \mapsto \mathbb{R}_{++}$relative error is positive, unit-free, and has the following componentwise bounds:

Let $E^\sigma \neq 0$, $E^\pi \neq 0$, $E^\nu \neq 0$. Then, $\frac{||E^\sigma||}{||S^\tau||} \leq 1$, $\frac{||E^\pi||}{||S^\tau||} \leq 1$, and $\frac{||E^\nu||}{||S^\tau||} \leq 2$, leading to $\rho(S^\tau) \leq \sqrt{6}$.

The proof of this claim can be found in the appendix. The following equation, used in it, is of importance: $\rho(S^\tau)^2 = \frac{||E^\sigma||^2}{||S^\tau||^2} + \frac{||E^\pi||^2}{||S^\tau||^2} + \frac{||E^\nu||^2}{||S^\tau||^2}$.

The bounds established in the claim can sometimes be made tighter. For instance, if $E = E^\nu$,

$$\rho(S^\tau) = \frac{||E^\nu||}{||S^\tau||} = \frac{||S_+^\tau||}{||S^\tau||} \leq \frac{max\{||\lambda_+||_s\}}{max\{||\lambda||_s\}} \leq 1.$$

This is because S_+^τ shares the same non negative eigenvalue functions as S^τ, and then $||S_+^\tau|| < ||S^\tau||$.

This claim shows that all violations of rationality in the consumer choice setting are indeed bounded above and we have computed the exact upper bound for the relative matrix nearness error: $\rho(\cdot) \leq \sqrt{6}$.

However, it is interesting to think of $\rho(S^\tau) = 1$ as being an important threshold for bounded rationality, in the sense that it is the upper bound for violations of VARP alone and WARP alone. It is also interesting to note that the violations of WARP have a higher upper bound for the relative measure than the other two axioms.

Another useful approach to deal with the unit dependence of the Slutsky matrix norm that we have built is to consider a normalized Slutsky matrix that is expressed in dollars.[13]

Definition 6. For any Slutsky matrix function $S^\tau \in \mathcal{M}(Z)$, let its normalized Slutsky matrix function be $\overline{S}^\tau = Diag(p) S^\tau Diag(p) = \Lambda_p S^\tau \Lambda_p$ or element-wise $\overline{S}^\tau_{ij} = S^\tau_{ij} p_i p_j$.

Observe that this normalized Slutsky matrix function is expressed in dollar terms, and that its Frobenius norm is $||\overline{S}^\tau(z)||^2 = \int_{z\in Z} Tr(\overline{S}^\tau(z)' \overline{S}^\tau(z)) = \int_{z \in Z} \sum_{i,j} [S^\tau_{ij}(z) p_i p_j]^2 dz$. We can reformulate our matrix nearness problem using a dollar-norm for any Slutsky matrix function defined as $||S^\tau||_\$ = ||\overline{S}^\tau(z)||$, which is the Frobenius norm applied to the weighted matrix function. Under this new norm choice we have $max_A ||S^\tau - A||_\$$ where A has property $\mathfrak{R}$.

We state a technical remark that underscores how the main results carry over to this modified problem.

Remark 6. [14]

$A^* = S^\sigma$ when only the VARP is violated but $||E||_\$ = ||\Lambda_p E^\sigma \Lambda_p||$. Also, one has the same $S^{\sigma,\pi} = S^\sigma + E^\pi$, but the negative semidefinite matrix nearness solution is no longer the same as under the Frobenius norm.

The solution to this problem is given implicitly by:

$S^{\overline{\tau}} = argmin_{A \leq 0} ||\Lambda_p [S^{\sigma,\pi} - A] \Lambda_p||^2 - 2\langle \Lambda_p E^\pi \Lambda_p, \Lambda_p [S^{\sigma,\pi} - A] \Lambda_p \rangle$, where $S^{\overline{\tau}}$ has property $\mathfrak{R}$ and has an associated rational demand that is ϵ- close to the observed demand by the almost implies near principle.

[13]We thank Xavier Gabaix for suggesting the use of this norm and pointing out its importance.

[14]The proof of this technical remark is in the appendix.

The exact solution to this problem is given by $S^{\overline{\tau}} = \Lambda_p^{-1}(\Lambda_p S^{\sigma,\pi}\Lambda_p)_-\Lambda_p^{-1}$ with the quadratic dollar norm given by:

$$||E||_\$^2 = ||\Lambda_p E^\sigma \Lambda_p||^2 + ||\Lambda_p E^\pi \Lambda_p||^2 + ||\Lambda_p E^\mu \Lambda_p||^2,$$

where $E^\mu = \Lambda_p^{-1}(\Lambda_p S^{\sigma,\pi}\Lambda_p)_-\Lambda_p^{-1}$.

We use μ to make clear that the projection under the dollar norm such that the Slutsky matrix function fulfills the NSD constraint is different from the solution under the Frobenius norm (ν). Here we provide the main elements of the proof or this technical remark while leaving the details to the appendix. To solve this problem we use the alternating projection technique. Therefore, we postulate $S^{\sigma,\pi,\mu}$ as a solution to the problem above, where $S^{\sigma,\pi,\mu} = \Lambda_p^{-1}(\Lambda_p S^{\sigma,\pi}\Lambda_p)_-\Lambda_p^{-1}$. Define $E^\mu = \Lambda_p^{-1}(\Lambda_p S^{\sigma,\pi}\Lambda_p)_+\Lambda_p^{-1}$.

We can easily check that $S^{\sigma,\pi,\mu}$ is a symmetric matrix and since for every matrix function such that $S(p,w)p = 0$ then $S(p,w)\Lambda_p\mathbf{1} = 0$ where $\mathbf{1} \in \mathbb{R}^L$. It also follows that $S^{\sigma,\pi,\mu}(p,w)p = 0$ since $(\Lambda_p S^{\sigma,\pi}(p,w)\Lambda_p)_-\mathbf{1} = 0$ and $\Lambda_p^{-1}p = \mathbf{1}$. We conclude that this must be the solution, and since it is unique, we let $S^{\overline{\tau}} = S^{\sigma,\pi,\mu}$. Moreover,

$$||E||_\$^2 = ||\Lambda_p E^\sigma \Lambda_p||^2 + ||\Lambda_p E^\pi \Lambda_p||^2 + ||\Lambda_p E^\mu \Lambda_p||^2.$$

The weighted norm is important in applications. In principle, the interested reader can tailor a different symmetric positive definite weighting matrix instead of Λ_p, as a function of the specific application. The alternating projection algorithm will still yield the solution.

It is also important to observe that the objective function using the dollar-norm is bounded above by $||S^\tau - A||_\$ \leq ||\Lambda_p||^2 \cdot ||S^\tau - A||$, so we know that the bounds that we have established for the relative norm carry over to this case.

6 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.

6.1 The Sparse-Max Consumer Model of Gabaix (2012)

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:
$x_i^{CD} = \frac{\alpha_i w}{p_i}$ for $i = 1, 2$.
$x_{i,p_i}^{CD} = -\frac{\alpha_i w}{p_i^2}$
$x_{i,w}^{CD} = \frac{\alpha_i}{p_i}$
$s_{i,i}^{CD} = -\frac{\alpha_i w}{p_i^2} + \frac{\alpha_i}{p_i}\frac{\alpha_i w}{p_i} = -\frac{\alpha_i(1-\alpha_i)w}{p_i^2}$

$s_{i,j}^{CD} = \frac{\alpha_i}{p_i}\frac{\alpha_j w}{p_j}$.

The Slutsky matrix function is:

$$S^{CD}(p,w) = \begin{bmatrix} -\frac{\alpha_1\alpha_2 w}{p_1^2} & \frac{\alpha_1}{p_1}\frac{\alpha_2 w}{p_2} \\ \frac{\alpha_1}{p_1}\frac{\alpha_2 w}{p_2} & -\frac{\alpha_1\alpha_2 w}{p_2^2} \end{bmatrix}.$$

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

$x_i^G = \frac{\alpha_i}{p_i^G}\frac{w}{\sum_j \alpha_j \frac{p_j}{p_j^G}}$ 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 $p_i^G(m) = m_i p_i + (1-m_i)p_i^d$. The vector of attention to price changes m, weights 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:

$$S^G(p,w) = S^{CD}(p,w)M$$

$$S^G(p,w) = \begin{bmatrix} -\frac{\alpha_1\alpha_2 w}{p_1^2} & 0 \\ \frac{\alpha_1}{p_1}\frac{\alpha_2 w}{p_2} & 0 \end{bmatrix}.$$

This matrix is not NSD, nor singular with p in its null space. Applying Theorem 1, the nearest Slutsky matrix when $p = p^d$ is:

$$S^r(p,w) = \frac{p_2^2}{p_1^2+p_2^2}\begin{bmatrix} -\frac{\alpha_1\alpha_2 w}{p_1^2} & \frac{\alpha_1}{p_1}\frac{\alpha_2 w}{p_2} \\ \frac{\alpha_1}{p_1}\frac{\alpha_2 w}{p_2} & -\frac{\alpha_1\alpha_2 w}{p_2^2} \end{bmatrix}$$

Also, one has

$$E(p,w) = S^\tau(p,w) - S^r(p,w)$$

$$E(p,w) = \begin{bmatrix} [1-b(p)][-\frac{\alpha_1\alpha_2 w}{p_1^2}] & b(p)(-\frac{\alpha_1}{p_1}\frac{\alpha_2 w}{p_2}) \\ [1-b(p)][-\frac{\alpha_1}{p_1}\frac{\alpha_2 w}{p_2}] & b(p)(\frac{\alpha_1\alpha_2 w}{p_2^2}) \end{bmatrix}$$

with

$$b(p) = \frac{p_2^2}{p_1^2+p_2^2}.$$

Now, we compute a useful quantity:

$$Tr(E'E) = \frac{w^2{\alpha_2}^2\alpha_1^2}{p_1^2p_2^2}.$$

It is convenient to compute the contributions of the violations of symmetry and singularity in p separately.

$Tr(E'E) = Tr(E^{\sigma\prime}E^\sigma) + Tr(E^{\pi\prime}E^\pi) = \frac{1}{2}\frac{w^2{\alpha_2}^2\alpha_1^2}{p_1^2p_2^2} + \frac{1}{2}\frac{w^2{\alpha_2}^2\alpha_1^2}{p_1^2p_2^2}$. 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: $||E||^2 = \frac{1}{2}\int_{\underline{w}}^{\overline{w}} Tr(E^\sigma(w)'E^\sigma(w))dw + \frac{1}{2}\int_{\underline{w}}^{\overline{w}} Tr(E^\pi(w)'E^\pi(w))dw = \left(\frac{\overline{w}^3 - \underline{w}^3}{3}\right)\frac{{\alpha_2}^2\alpha_1^2}{p_1^2p_2^2}$, 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 Gabaix does not satisfy WARP and its Slutsky matrix function violates NSD, we conclude that the 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 $m = [1, 0]'$.

$Tr(E'E) = \frac{w^2 {\alpha_2}^2 \alpha_1^2}{p_1^2} \frac{[p_2^d]^2}{[p_2 + [p_2^d - p_2]\alpha_1]^4}$, with $Tr(E^{\sigma'} E^{\sigma}) = Tr(E^{\pi'} E^{\pi})$ and $E^{\nu} = 0$.

The expression above has a positive derivative with respect to p_2^d for $\alpha_1 + \alpha_2 = 1$, this indicates that $\frac{\partial}{\partial p_2^d} \delta(\cdot) > 0$ for any p. A sparse consumer that does not pay attention to p_2 will be further from rationality when the default price p_2^d is high. Furthermore, the power of our approach lies in the decomposition of $||E||^2 = ||E^{\sigma}||^2 + ||E^{\pi}||^2$. 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.

By Proposition 1, we can automatically conclude that there exists a $x^r \in \mathcal{X}(Z)$ with Slutsky matrix function $S^r \in \mathcal{M}_{\mathfrak{R}}(Z)$ such that $||E|| < \delta$ and $||x^G - x^r|| < \epsilon(\delta)$, i.e., there is a rational demand system $x^r \in \mathcal{R}(Z)$ that is ϵ-close in the behavioral sense to $x^G \in \mathcal{X}(Z)$ that is different from the underlying Cobb-Douglas model. It must be underlined that the Cobb-Douglas model and x^r are related to x^G in different ways. The first one is a distortion of a rational model using the sparse max operator, while the second is $\epsilon-$closest behaviorally, as defined here. Our approach helps to complement the understanding of how much this particular bounded rationality model differs from the standard rational one.

Indeed, fixing $p = p^d$, one can compare the $Tr(E'E)$ to the trace of the residual matrix of the distance between S^G and S^{CD} from the Cobb-Douglas consumer:

$Tr(E'_{CD} E_{CD}) = \frac{p_1^2 + p_2^2}{p_2^2} \frac{w^2 {\alpha_2}^2 \alpha_1^2}{p_1^2 p_2^2}$, with $E_{CD} = S^G - S^{CD}$ or

$Tr(E'_{CD} E_{CD}) = \frac{1}{b(p)} Tr(E'E)$, notice that $0 < b(p) < 1$ for $p >> 0$. Then, as expected due to our theoretical results, the $S^r(p, w)$ obtained using Theorem 1 is uniformly closer (under the Frobenius norm) to $S^G(p, w)$ than the Cobb-Douglas matrix in any compact space Z of pairs $z = (p, w)$ where $||E||$ and $||E_G||$ are defined. This result says that both $||E||^2 = \frac{1}{b(p)} ||E_{CD}||^2$ are proportional for any segment of wealth $[\underline{w}, \overline{w}]$ due to linearity of the definite integral operator.

To finish 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 α_1 and p_2^d. Let $Z = \{w, p_1 = 1, p_2 \in [1, 2]\}$, then $\delta(\alpha_1, p_2^d) = \frac{1}{3}(\alpha_1 - 1){\alpha_1}^2 [p_2^d]^2 \left(\frac{1}{(\alpha_1(p_2^d - 2) + 2)^3} - \frac{1}{(\alpha_1(p_2^d - 1) + 1)^3} \right)$. One can now visualize this δ in the α, p_2^d space, that is at $p_2^d \in [1, 2]$ and $\alpha_1 \in [0, 1]$ (figure 1). We can observe that α_1 has a non-linear effect on δ, and the distance toward the rational matrix goes to zero when either $\alpha_1 \to 0$ or $\alpha_1 \to 1$ for all $p_2^d \in [1, 2]$.

6.2 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 $\mathfrak{R}$ by naive and sophisticated quasi-hyperbolic discounters.

The optimization problem for a consumer that can pre-commit is:

$max_{\{x_i^p\}_{i=1,2,3.}} u(x_1^p) + \beta\theta u(x_2^p) + \beta\theta^2 u(x_3^p)$

subject to the budget constraint

$\sum_{i=1}^{3} p_i x_i^p = w.$

The first order conditions are:

$u'(x_1^p) = \lambda p_1$

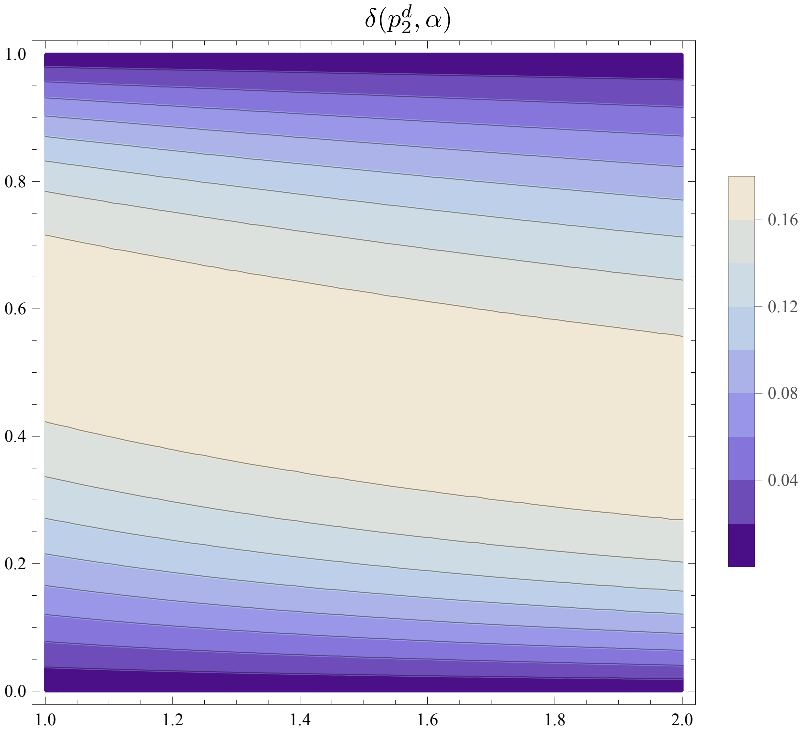

Figure 1: Level curves sparse-max consumer example . The figure plots $\alpha \in [0,1]$ on its vertical axis and $p_2^d \in [1,2]$ on its horizontal axis.

$\beta\theta u'(x_2^p) = \lambda p_2$
$\beta\theta^2 u'(x_3^p) = \lambda p_3$
With CRRA utility with relative risk aversion σ:
$u'(x_i^p) = [x_i^p]^{-\sigma}$
$x_2^p = [\beta\theta \frac{p_1}{p_2}]^{\frac{1}{\sigma}} x_1^p$
$x_3^p = [\beta\theta^2 \frac{p_1}{p_3}]^{\frac{1}{\sigma}} x_1^p$
Then, imposing the budget constraint:
$p_1 x_1^p + p_2[\beta\theta \frac{p_1}{p_2}]^{\frac{1}{\sigma}} x_1^p + p_3[\beta\theta^2 \frac{p_1}{p_3}]^{\frac{1}{\sigma}} x_1^p = w,$
which gives the demand system:
$x_1^p = [p_1 + p_2[\beta\theta \frac{p_1}{p_2}]^{\frac{1}{\sigma}} + p_3[\beta\theta^2 \frac{p_1}{p_3}]^{\frac{1}{\sigma}}]^{-1} w$
$x_2^p = [\beta\theta \frac{p_1}{p_2}]^{\frac{1}{\sigma}} x_1^p$
$x_3^p = [\beta\theta^2 \frac{p_1}{p_3}]^{\frac{1}{\sigma}} x_1^p.$
The naive quasi-hyperbolic discounter will have the following demand system:

In the first period, the consumer assumes she will stick to his commitment in the second period and consumes the same amount as in the pre-commitment case:

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

However, when period two arrives, she re-optimizes taking as given the remaining wealth $w - p_1 x_1^h$.

$x_2^h = \frac{w - p_1 x_1^h}{p_2 + p_3[\beta\theta \frac{p_2}{p_3}]^{\frac{1}{\sigma}}}$
$x_3^h = [\beta\theta \frac{p_2}{p_3}]^{\frac{1}{\sigma}} x_2^h.$
The analytical result for the matrix nearness norm has a nice expression:

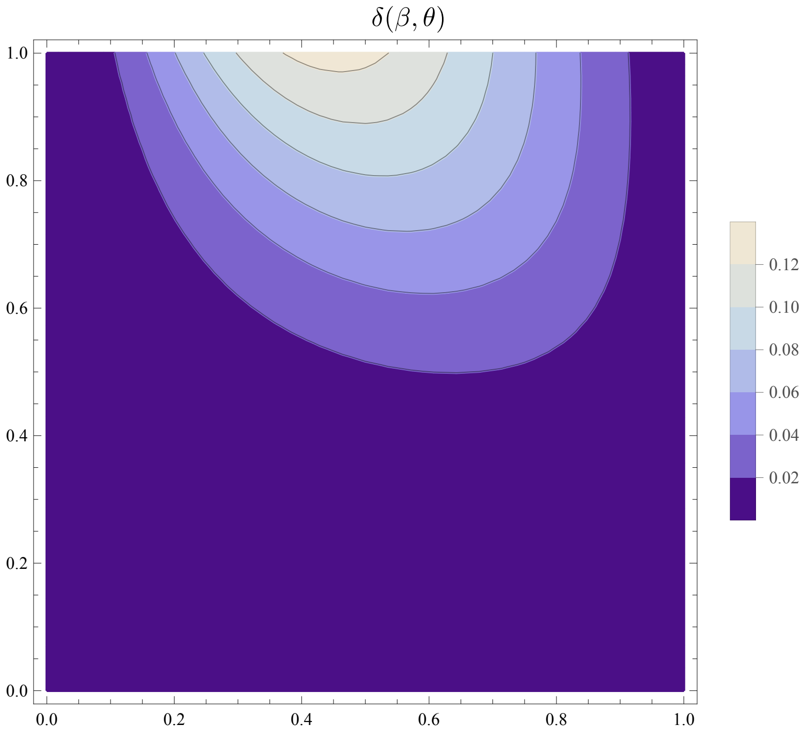

Figure 2: Level curves hyperbolic discounter example . The figure plots $\beta \in [0,1]$ on its horizontal axis and $\theta \in [0,1]$ on its vertical axis.

$$Tr(E'E) = \frac{(\sigma-1)^2 w^2 \left(p_1^2+p_2^2+p_3^2\right)\left(\left(\frac{\beta\theta^2 p_1}{p_3}\right)^{\frac{1}{\sigma}} - \left(\frac{\beta\theta p_1}{p_2}\right)^{\frac{1}{\sigma}}\left(\frac{\beta\theta p_2}{p_3}\right)^{\frac{1}{\sigma}}\right)^2}{2\sigma^2\left(p_3\left(\frac{\beta\theta p_2}{p_3}\right)^{\frac{1}{\sigma}}+p_2\right)^2\left(p_2\left(\frac{\beta\theta p_1}{p_2}\right)^{\frac{1}{\sigma}}+p_3\left(\frac{\beta\theta^2 p_1}{p_3}\right)^{\frac{1}{\sigma}}+p_1\right)^4},$$

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 $S^r \to 0$ and $x_1^h \to \frac{w}{p1}$ 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 her consumption is time consistent and $||E|| = 0$.

To illustrate further the use of the tools developed here, we find an explicit value for δ in terms of the behavioral parameters β, θ, 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 $\delta^2 = \frac{7\beta^4\left(\beta^2-1\right)^2\theta^8}{2\left(\beta^2\theta^2+1\right)^2\left(\beta^2\left(\theta^4+\theta^2\right)+1\right)^4}$.

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

The analytical expression for δ is messy. When evaluated at $\beta = 0.7$ and $\theta = 0.9$, then $\delta = 0.074$, where the parameters are taken from the empirical literature.

Another observation that we can draw from this example is that for any arbitrary compact region Z of prices and wealth analyzed $||E||^2 = ||E^\sigma||^2$. That is, only the asymmetric part plays a role in the violation of property $\mathfrak{R}$. 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.

Finally, one can also use this example to identify pairs (β, θ) 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.

6.3 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 $t = 2$ she will not be able to keep her commitment and therefore will adjust her consumption at $t = 1$. Then the consumer maximizes

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

where x_2^h, x_3^h are known to her in $t = 1$ and depend on period 1 consumption. However, she can control only how much she consumes in the first period. Taking first order conditions and keeping the assumption of the naive quasi-hyperbolic case, the first order conditions are:

$$u'(x_1^{sh}) + \beta\theta u'(x_2^h)\frac{\partial x_2^h}{\partial x_1^{sh}} + \beta\theta^2 u'(x_3^h)\frac{\partial x_3^h}{\partial x_1^{sh}} = 0$$

Under the parametric assumptions made in the previous example, the new demand system of the sophisticated hyperbolic discounter is:

$u'(x_i^p) = [x_i^p]^{-\sigma}$

$$[x_1^{sh}]^{-\sigma} + \beta\theta[x_2^h]^{-\sigma}\frac{\partial x_2^h}{\partial x_1^{sh}} + \beta\theta^2[x_3^h]^{-\sigma}\frac{\partial x_3^h}{\partial x_1^{sh}} = 0$$

$x_2^h = \frac{w - p_1 x_1^h}{p_2 + p_3[\beta\theta\frac{p_2}{p_3}]^{\frac{1}{\sigma}}}$

$x_3^h = [\beta\theta\frac{p_2}{p_3}]^{\frac{1}{\sigma}} x_2^h$.

$\frac{\partial x_2^h}{\partial x_1^{sh}} = \frac{p_1}{p_2 + p_3[\frac{p_2\beta\theta}{p_3}]^{\frac{1}{\sigma}}}$

$\frac{\partial x_3^h}{\partial x_1^{sh}} = -\frac{p_1[\frac{p_2\beta\theta}{p_3}]^{\frac{1}{\sigma}}}{p_2 + p_3[\frac{p_2\beta\theta}{p_3}]^{\frac{1}{\sigma}}}$.

Then the first period consumption under the sophisticated hyperbolic discounting is:

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

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

$$Tr(E'E) =$$

$$\frac{(\beta-1)^2(\sigma-1)^2 w^2 \theta^{4/\sigma}\left(p_1^2+p_2^2+p_3^2\right)\left(\beta p_2 + p_3\left(\frac{\beta\theta p_2}{p_3}\right)^{\frac{1}{\sigma}}\right)^{\frac{2}{\sigma}-2}\left(\frac{\beta p_1}{p_3\left(p_3\left(\frac{\beta\theta p_2}{p_3}\right)^{\frac{1}{\sigma}}+p_2\right)}\right)^{2/\sigma}}{2\sigma^4\left(\left(\frac{\theta p_1\left(p_3\left(\frac{\beta\theta p_2}{p_3}\right)^{\frac{1}{\sigma}}+p_2\right)^{\sigma-1}\left(\beta p_2+p_3\left(\frac{\beta\theta p_2}{p_3}\right)^{\frac{1}{\sigma}}\right)}{p_2}\right)^{\frac{1}{\sigma}}+p_1\right)^4}$$

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 want also to compare this quantity with the case of the naive hyperbolic discounter. The ratio of $r = \frac{Tr(E'_{sh}E_{sh})}{Tr(E'_h E_h)} < 1$ means that the sophisticated hyperbolic consumer has a lower δ for any Z and any parameter configuration. To simplify expressions, we let $p_i = 1$ for $i = 1, 2, 3$. The first finding is that the relation between the naive and the sophisticated discounter δ's depends crucially on the parameter σ. For $\sigma = 1$, they 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 $\sigma = \frac{1}{2}$, the sophisticated hyperbolic discounter has a uniformly lower δ. However, for $\sigma = 2$, the naive hyperbolic discounter has a uniformly lower δ. 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. More precisely, in light of the interpretation of the matrix nearness norm, the slope of the steepest Ville cycle changes with the amount of wealth remaining after the first period. Consequently, if there is a larger amount of remaining wealth for "re-optimization" in the second period, the Ville cycle slope is greater. This is the case when $\sigma < 1$, which increases the consumption in period 1 of the sophisticated hyperbolic discounter leaving less residual wealth and thus a limited rate of growth of the real income path along the Ville cycle. Furthermore, to enhance the comparison for the case of $\sigma = \frac{1}{2}$, we compute explicitly the expression for δ 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.

$$\delta^2(\beta, \theta) = \frac{14(\beta-1)^2\beta^6\theta^8\left(\beta\theta^2+1\right)^2}{\left(\beta^2\theta^2\left(\beta\theta^2\left(\beta\theta^2+2\right)+2\right)+1\right)^4}$$

The level curves of this δ expression are very similar to the naive case, but it is even closer to rationality everywhere. Evaluated at the typical values of $\beta = 0.7$ and $\theta = 0.9$, one gets the value $\delta = 0.0703847$, which is slightly lower than the δ of the naive hyperbolic case for the same $\sigma = \frac{1}{2}$.

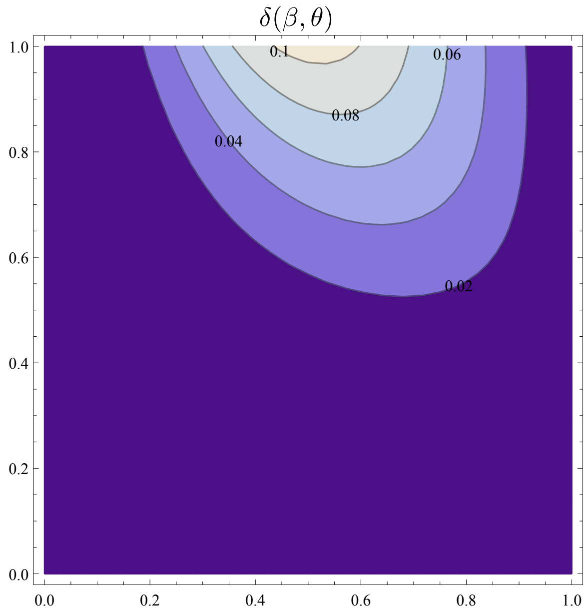

Figure 3: Level curves hyperbolic discounter example . The figure plots $\beta \in [0, 1]$ on its horizontal axis and $\theta \in [0, 1]$ on its vertical axis.

6.4 Semiparametric and Nonparametric Estimation of the Slutsky Matrix Norm with Noisy Data

Even though our Slutsky matrix norm main attractiveness is its analytical tractability, our approach can be useful when applied to observed consumer behavior where nonparametric estimates of the demand function are obtained from noisy data. Its main advantage over alternatives that use a least distance approach to measure violations to rationality such as Blundell et al. (2008) or Varian (1990) is its closed-form solution. Moreover, it not only quantifies by how much a behavior departs from rationality, but also suggests why (e.g. the violations to the Slutsky conditions and its axiomatic counterparts). It also has a potential advantage with respect to tests on the Slutsky regularity conditions such as those proposed by Hoderlein (2011) and Haag et al. (2009) because our approach deals with the condition in a unified manner. Here we propose a blueprint of a consistent estimator for departures from rationality using our results of the matrix nearness problem. The estimator of our functional is semiparametric and uses the plug-in principle. Its main interest rests in its simplicity.

In most applications, one observes noisy versions of our primitive $x^\tau \in \mathcal{X}(Z)$. Indeed, the practitioner may have access to a noisy version of the expenditure of households given the sampled prices and incomes and budget shares for each commodity. In this setting, we formally define these alternative functions of consumer behavior as $\alpha^\tau \in \mathcal{X}(Z)$ (e.g. when the demand is modeled as a budget share, then α^τ is defined componentwise $\alpha_l^\tau(p, w) = \frac{p_l x_l^\tau(p,w)}{w}$). Also, we assume that we observe N identically and independently distributed draws of the random vector $Y = (A \quad P \quad W)'$, where $A = f(P, W, V)$ is the noisy measure of the budget shares that depends on (random) price and wealth (P, W) pairs and on a vector V of noise or unobserved heterogeneity. In this simple framework, we assume that (P, W) is independent of V. Ideally, the practitioner should use experimental data for individual choice with a bundle set that is infinite $X \subseteq \mathbb{R}^L$ and a random linear budget constraint with the objective of inferring the shape of x^τ. However, the typical situation that a practitioner faces does not involve experimental data. In that case, the interested reader should use the Slutsky matrix function estimator that addresses the issues of endogeneity and heterogeneity (Hoderlein, 2011) while imposing strong assumptions about the population demand behavior (Lewbel, 2001).

To retain generality, we will only assume that there is some functional of the observed random variable A such that it is equal to the deterministic α^τ for given observables (prices and wealth). A classical assumption is that $\mathbb{E}[A|P = p, W = w] = \alpha^\tau(p, w)$ (Haag et al., 2009). More recently, Blundell et al. (2013) use the conditional quantile of A, assumed to correspond to the actual deterministic behavior component such that $Q_\tau(A|P = p, W = w) = \alpha^\tau(p, w)$, where we interpret τ both as a type of behavior and a population quantile. In this environment, we redefine our problem without loss of its structure. We can compute the Slutsky matrix function associated to α^τ, and call it $S^{\alpha,\tau}$. It must be the case that $S^{\alpha,\tau}$ inherits all the properties of S^τ when τ is a rationalizable behavior. In particular, if it is symmetric, p-singular and NSD then our results hold for α^τ. The typical budget share estimation is done under Walras' law, so we assume it here as well.[15] Interestingly, we find that our semiparametric estimator can be seen as a generalization of the well-known Slutsky matrix symmetry test, and it is close to an L_2 distance test of symmetry proposed by Hagg et al. (2009).

Our aim is to provide a blueprint to estimate the distance from rationality of α^τ (or x^τ). We define the quantity of interest $\gamma = ||E^\alpha||^2 = \int_Z g(z)dz$ (or δ in the case of x^τ) where the function g is defined pointwise as $g(p, w) = Tr[E^{\alpha'}(p, w)E^\alpha(p, w)]$.[16] Abusing notation, we omit α from the following derivations.

The first step in the estimation of γ, is the nonparametric estimation of the function g. Note that g is constructed using only the shares α^τ and its partial derivatives (e.g., $D_p\alpha^\tau$ and $D_w\alpha^\tau$) that we can readily estimate in this environment. [17]

(i) Plug-in estimator. By the plug-in principle, we propose the estimator $\tilde{\gamma} = \int_Z \hat{g}(z)dz = \Gamma(\hat{g})$, where $\Gamma : \mathcal{C}(Z, \mathbb{R}_+) \mapsto \mathbb{R}$ is a linear functional and $\int_Z ||\hat{g}(z)||^2 < \infty$. This functional is Fréchet differentiable (or Hadamard differentiable), then the plug-in estimator converges.

In particular, we will take advantage of a sufficient condition (i.e. the fact that $\gamma = \Gamma(g)$, and that Γ is a linear bounded functional of g). To see this, observe that for all $f \in C(Z, \mathbb{R})$ we have: $|\Gamma(g) - \Gamma(f)| \leq c \quad ||g(z) - f(z)||_{\mathcal{C}}$ for some constant $c > 0$.[18] Since Γ is a bounded functional, it follows that if $\hat{g} \xrightarrow{p} g$, and then $\Gamma(\hat{g}) \xrightarrow{p} \Gamma(g)$.

In the remaining part of this application, we propose an estimation procedure for the first nonparametric step for this estimator. This procedure takes advantage of our closed-form solutions to the matrix nearness problem. The emphasis of this procedure is on obtaining an

[15] It is desirable to use estimators for x^τ when heteroscedasticity is not an issue for the estimation technique.

[16] We can write $\gamma = \int_Z \frac{g(z)}{h(z)} h(z)dz$ or $\gamma = \mathbb{E}[\frac{g(P,W)}{h(P,W)}]$ where h is the probability density function (pdf) of (P, W) that is assumed to be bounded away from zero and whose support is Z. This will be useful to simplify the algorithm to compute the estimator as explained in the appendix. The estimator can be written as $\hat{\gamma} = \frac{1}{N}\sum_{i=1}^n \frac{1}{\hat{h}(z^i)}\hat{g}(z^i)$ where $\hat{g}(z^i)$ can be computed only at the data points.

[17] In particular, following Hagg et al. (2009), in the case where $\mathbb{E}[A|P = p, W = w] = \alpha^\tau(p, w)$, we can use the Naradaya-Watson local polynomial estimator of the conditional mean for α^τ and the partial derivative of $\hat{\alpha}^\tau$ for the corresponding functions of interest. Then we are ready to construct an estimator of $\hat{g}$. To estimate h we can use a traditional kernel estimator for probability density functions $\hat{h}$. We must underline that Haag et al. (2009) shows that the asymptotic properties of an object such as $\hat{g}$ are driven by the rate of convergence of the derivatives of α^τ, and therefore, we can treat$\hat{\alpha}^\tau$ as known in the asymptotic theory.

[18] In fact, note that by linearity of the integral operator, we have
$|\Gamma(g - f)| \leq c \quad max_{z\in Z}|g(z) - f(z)|$. Then we can construct a c using the fact that $|\Gamma(g - f)| \leq \Gamma(|g - f|)$ and fixing $c = max_{f \in L_2(Z,\mathbb{R}_+)}\{\frac{\Gamma(|g-f|)}{max_{z\in Z}|g(z)-f(z)|}\}$ for $f \neq g$.

algorithm that provides a consistent estimator of g in general settings, when we have consistent estimators of α^τ and its derivatives.

The procedure can be summarized in two steps:

- Use a local smoother on $\{p^i, w^i, \overline{g}^i\}_{i\in\{1,\cdots,N\}}$to obtain and estimate of $\hat{g}(p,w)$ for all $(p,w) \in Z$.
- Use the plug-in principle and obtain $\hat{\gamma} = \Gamma(\hat{g}) = \int_Z \hat{g}(z)dz$.

(ii) Properties of the estimator. The advantage of the semiparametric estimation procedure is that it is consistent if we have a uniform consistent estimator as an input. Formally, we state the following remark.

Remark 7. The semiparametric estimator $\hat{\gamma} \xrightarrow{p} \gamma$ as $n \to \infty$ if the estimators of α^τ and its derivatives converge in probability to their true value.

The asymptotic distribution of this statistic can be done under regularity conditions using the classical results of Newey (1994). The interested reader might want to check the following: (i) Mean square continuity. To achieve this we have to use an $o(n^{-1/4})$ consistent estimator of g. (ii) Linearity. This condition holds trivially when we know h then γ is a linear functional of g. In fact, the functional derivative of $\mathbb{E}[m(z,\gamma,g)]$ with respect to g is $\mathbb{E}[c(P,W)\delta g(P,W)]$ where $c(p,w) = \frac{1}{h(p,w)}$. (iii) Stochastic equicontinuity. Proving stochastic equicontinuity if possible depends on the specific estimator used by the practitioner. Under these conditions, the estimator converges at $n^{-1/2}$ rate.

(iii) Symbolic or algebraic algorithm for the estimation of g.[19] In the case of the first nonparametric step, we propose a simple polynomial estimator for all the components of our Slutsky matrix norm. These components (matrix functions) will be used to compute $\hat{\gamma}$ using the plug-in principle and other quantities of interest. We assume that we have polynomial consistent estimators of α^τ and its partial derivatives such that $\hat{S}^\tau(p,w) = \overline{S}_{r,q} p^r w^q$ (i.e., we use the Einstein notation where $p^r = p_1^{r_1} p_2^{r_2} \cdots p_L^{r_L}$, and w^q is the q-th power of w and $\overline{S}_{r,q} p^r w^q = \sum_{r_1,\cdots,r_L;q} \overline{S}_{r,q} p_1^{r_1} \cdots p_L^{r_L} w^q$ with $-k \leq r_l, q \leq k$ and $\overline{S}_{r,q}$ constant matrices). Notice that our map $a : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ is closed to rational polynomials in the sense that if α^τ is a rational polynomial then E has rational polynomial entries. Indeed, if $\alpha^\tau$and its derivatives are polynomials then S^τ has polynomial entries. Moreover, $S^{\sigma,\pi}$ having rational polynomial entries implies that the eigen-values and eigen-vectors (of $S^{\sigma,\pi}$) are rational polynomials and so is S^r (entrywise). Finally, E itself has rational polynomial entries. In this case, we can apply our analytical solutions in a very simple and numerically efficient way using readily available algorithms.

Notice that the estimator of E^σ, $\hat{E}^\sigma = -\frac{1}{2}(\overline{S}_{r,q} - \overline{S}'_{rq}) p^r w^q = \overline{E}^\sigma_{rq} p^r w^q$ is computed using only constant matrix subtraction. The estimator of E^π has rational polynomial entries instead of just polynomial ones due to the presence of $\frac{1}{p'p}$. However, we can factor the estimator in two components, a scalar rational polynomial and a polynomial matrix: $\hat{E}^\pi = \frac{1}{p'p}[\overline{E}^\pi_{s,t} p^s w^t]$. The change of subscripts denotes that this is a new polynomial matrix due to the following operation $\overline{E}^\pi_{s,t} p^s w^t = \overline{E}^{\sigma,\pi}_{r,q}(p,w) p^r w^q$ where $\overline{E}^{\sigma,\pi}_{r,q}(p,w) = -(pp'\overline{S}^\sigma_{r,q} + \overline{S}^\sigma_{r,q} pp')$[20] is a polynomial matrix itself and $\overline{E}^\pi_{s,t}$ is a constant matrix coefficient of the polynomial expansion of the matrix factor of $\hat{E}^\pi$ after rearranging the polynomial terms.

[19]For a numerical algorithm, the reader is referred to the appendix.
[20]This formula works under the assumption of Walras' law or adding-up.

Finally, the computation of the estimator of E^{ν} is more involved. We can apply the algorithms of Kitamoto (1994) on $\hat{S}^{\sigma,\pi,\rho} = (p'p)\hat{S}^{\sigma,\pi}$ if the polynomials have no interaction terms. More generally, we can exploit an additional structural property of $\hat{S}^{\sigma,\pi,\rho}$. Most modern algorithms designed to obtain an eigen-decomposition or a spectral decomposition of a polynomial matrix need that $\hat{S}^{\sigma,\pi,\rho}$ belongs to the class of para-hermitian polynomial matrices (i.e. $\hat{S}^{\sigma,\pi,\rho}(-p,-w) = \hat{S}^{\sigma,\pi,\rho}(p,w)')$. If this is not the case, one can always obtain the para-hermitian matrix transformation $\hat{S}^{\sigma,\pi,h}(z) = \hat{S}^{\sigma,\pi,\rho}(-z)'\hat{S}^{\sigma,\pi,\rho}(z)$, where $\hat{S}^{\sigma,\pi,h}(z) = \overline{Q}(-z)'\overline{\Lambda}(-z)'\overline{\Lambda}(z)\overline{Q}(z)'$, and where $\overline{Q}(z)$ is a para-unitary polynomial matrix such that $\overline{Q}(-z)'\overline{Q}(z) = I$ and $\Lambda(z)$ is a diagonal polynomial matrix. Then, the estimator is defined pointwise as $\hat{E}^{\nu}(z) = \frac{1}{p'p}[\overline{Q}(-z)'\overline{\Lambda}(z)_{+}\overline{Q}(z)']$. Observe that one can recover $\Lambda(z)$ from $\Lambda(-z)'\Lambda(z)$ (Foster et al., 2010). To exploit the para-hermitian property when possible is generally recommended, because of the increasing availability of algorithms to compute it. In particular, by using a decomposition algorithm for multivariate polynomial matrices with this property proposed by Avelli and Trentelman (2008) to obtain algebraic solutions to the PSD projection of $\hat{S}^{\pi,\sigma}$.[21]

We want to underscore that, in the process of deriving an estimator for $\hat{E}$, we have reduced the computation of the matrix function E to a simple algebraic or symbolic algorithm when the primitives are polynomials. Therefore, this process can also be applied when the practitioner has access to deterministic polynomial evaluator of x^{τ} or a Taylor approximation thereof. The benefits may be important computationally because we transformed the infinite-dimensional problem into a finite-dimensional one (except for the $\hat{E}^{\nu}$ that requires to apply a specialized algorithm).

We can use the approximate function $\hat{E}$ to calculate an estimator of γ. We start by obtaining $\hat{g}^{j} = Tr(\hat{E}^{j;}\hat{E})$ for $j \in \{\sigma, \pi, \nu\}$, and we use the plug-in estimator to compute the approximate $\tilde{\gamma} = \int_{Z} \hat{g}(z)dz$.[22]

As we can see, our closed-form solutions for the departures of rationality are advantageous when proposing estimators and algorithms for the computation of our Slutsky matrix norm. The asymptotic distribution and other properties of the matrix function polynomial estimator (S^{τ},S^{r} and E), are potentially interesting avenues for future research. Also, it seems natural to use these techniques to impose rationality in nonparametric demand estimation when using sieves techniques.[23]

[21] In the case of para-hermitian polynomial matrix functions, there is a spectral decomposition algorithm with para-unitary matrices such that $S^{\sigma,\pi}(z) = \overline{Q}(-z)'\overline{Q}(z)$ where $\overline{Q}(z) \in \mathbb{R}^{m\times L}$. Denote $\overline{Q}(z) = [\overline{q}_1(z) \quad \overline{q}_2(z)\cdots\overline{q}_m(z)]$, in other words, $S^{\sigma,\pi}(z) = \sum_{r=1}^{m}\overline{q}_r(-z)\overline{q}_r(z)'$. In this setting, we can obtain the PSD part $E^{\nu}(z) = \sum_{k\in K}\overline{q}_k(-z)\overline{q}_k(z)'$, where $K \equiv \{k \in \{1\cdots m\}|\overline{q}_k(-z)'\overline{q}_k(z) \geq 0\}$. Certainly, $E^{\nu}(z)$ is symmetric, PSD, and singularity in p is preserved since $\overline{Q}(-z)'\overline{Q}(z)p = 0$ and $p'\overline{Q}(-z)'\overline{Q}(z) = 0$, which implies that $\overline{q}_l(z)'p = \overline{q}_l(-z)'p = 0$ for all $l \in \{1,\cdots,m\}$. Uniqueness of the solution ensures that this is the PSD projection. Alternatively, we can obtain the PSD part by noticing that this decomposition can be seen also as a J-spectral decomposition where $S^{\sigma,\pi}(z) = \overline{F}(-z)'J\overline{F}(z)$ with $\overline{F}(z) \in \mathbb{R}^{L\times L}$ and $J = \begin{bmatrix} I_{\alpha} & 0 & 0 \\ 0 & -I_{\beta} & 0 \\ 0 & 0 & 0 \end{bmatrix}_{L\times L}$, where I_t denotes an identity matrix of dimension t, $t = \alpha$ is the number of positive eigen-values, $t = \beta$ is the number of negative eigen-values, and $m - (\alpha + \beta)$ is the number of null eigen-values. Then $E^{\nu} = \overline{Q}(-z)'J_{+}\overline{Q}(z)$, where $J_{+} = \begin{bmatrix} I_{\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{L\times L}$.

[22] The consistency of the plug-in estimator stems from the boundedness of the operator Γ (and Hadamard differentiability). Observe that the consistency of the coefficients of $\hat{E}$ polynomial matrix are guaranteed in the case of polynomial primitives. For general cases, we have approximate solutions that nevertheless can be made arbitrarily precise at a computational cost.

[23] Recent work on two-step estimators by Ackerberg et al. (2012) may prove useful in this setting to compute the asymptotic variance because it assumes that the first step is done using polynomial approximations

7 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 Axiom of Revealed Preference (GARP). Varian (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.

The closest treatment of the problem to our work is the approximately rational consumer demand proposed by Jerison and Jerison (1993) These authors are able to relate the violations of negative semi-definiteness and symmetry of the Slutsky matrix to the smallest distance between an observe smooth demand system and a rational demand. Russell (1997) proposes a measure of quasi-rationality. Russell's measure corresponds to Slutsky matrix symmetry violations. He uses exterior calculus and obtains a measure of non integrability that corresponds to the residual of a symmetric decomposition of the Slutsky matrix.

Our work gives 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, 2012) suggest the presence of a money illusion such that prices are not in the null space of the Slutsky matrix.

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 (2001) 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.

8 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 both a better computational implementability and the derivation of semiparametric consistent estimator and tests. 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.

9 References

References

Ackerberg, D., Chen, X., & Hahn, J. (2012). A practical asymptotic variance estimator for two-step semiparametric estimators. *Review of Economics and Statistics*, 94(2), 481–498.

Afriat, S. N. (1973). On a system of inequalities in demand analysis: an extension of the classical method. *International economic review*, 14(2), 460–472.

Anderson, R. M. (1986). Almost" implies "near. *Transactions of the American Mathematical Society*, 296(1), 229–237.

Avelli, D. N. & Trentelman, H. L. (2008). Algorithms for multidimensional spectral factorization and sum of squares. *Linear Algebra and its Applications*, 429(5–6), 1114–1134.

Blundell, R., Browning, M., & Crawford, I. (2008). Best nonparametric bounds on demand responses. *Econometrica*, 76(6), 1227–1262.

Blundell, R., Horowitz, J., & Parey, M. (2013). *Nonparametric estimation of a heterogeneous demand function under the Slutsky inequality restriction*. Technical report, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Boualem, H. & Brouzet, R. (2012). On what is the almost-near principle. *The American Mathematical Monthly*, 119(5), 381–397.

Dennis, J. & Schnabel, R. B. (1979). Least change secant updates for quasi-newton methods. *Siam Review*, 21(4), 443–459.

Echenique, F., Lee, & Shum, M. (2011). The money pump as a measure of revealed preference violations. *Journal of Political Economy*, 119(6), 1201–1223.

Foster, J., McWhirter, J., Davies, M., & Chambers, J. (2010). An algorithm for calculating the QR and singular value decompositions of polynomial matrices. *IEEE Transactions on Signal Processing*, 58(3), 1263–1274.

Gabaix, X. (2012). A sparsity-based model of bounded rationality, applied to basic consumer and equilibrium theory. *manuscript, New York University*.

Graves, L. M. (1956). *The Theory of Functions of Real Variables...*, volume 121. New York.

Haag, B. R., Hoderlein, S., & Pendakur, K. (2009). Testing and imposing slutsky symmetry in nonparametric demand systems. *Journal of Econometrics*, 153(1), 33–50.

Higham, N. J. (1989). Matrix nearness problems and applications. In *Applications of Matrix Theory* (pp. 1–27).: University Press.

Hoderlein, S. (2011). How many consumers are rational? *Journal of Econometrics*, 164(2), 294–309.

Houthakker, H. S. (1950). Revealed preference and the utility function. *Economica*, 17(66), 159–174.

Houtman, M. & Maks, J. (1985). Determining all maximal data subsets consistent with revealed preference. *Kwantitatieve Methoden*, (19), pp. 89–104.

Hurwicz, L. & Richter, M. K. (1979). Ville axioms and consumer theory. *Econometrica*, 47(3), 603–619.

Jerison, D. & Jerison, M. (1992). Approximately rational consumer demand and ville cycles. *Journal of Economic Theory*, 56(1), 100–120.

Jerison, D. & Jerison, M. (1993). Approximately rational consumer demand. *Economic Theory*, 3(2), 217–241.

Jerison, D. & Jerison, M. (1996). A discrete characterization of slutsky symmetry. *Economic Theory*, 8(2), 229–237.

Jerison, D. & Jerison, M. (2001). Real income growth and revealed preference inconsistency. *Documentos de trabajo. Economic series (Universidad Carlos III. Departamento de Economía)*, (29), 1.

John, R. (1995). The weak axiom of revealed preference and homogeneity of demand functions. *Economics Letters*, 47(1), 11–16.

Kadison, R. V. (1984). Diagonalizing matrices. *American Journal of Mathematics*, 106(6), 1451–1468.

Kihlstrom, R., Mas-Colell, A., & Sonnenschein, H. (1976). The demand theory of the weak axiom of revealed preference. *Econometrica*, 44(5), 971–978.

Kitamoto, T. (1994). Approximate eigenvalues, eigenvectors and inverse of a matrix with polynomial entries. *Japan Journal of Industrial and Applied Mathematics*, 11(1), 73–85.

Lewbel, A. (2001). Demand systems with and without errors. *American Economic Review*, (pp. 611–618).

Newey, W. K. (1994). The asymptotic variance of semiparametric estimators. *Econometrica: Journal of the Econometric Society*, (pp. 1349–1382).

Russell, T. (1997). How quasi-rational are you?: A behavioral interpretation of a two form which measures non-integrability of a system of demand equations. *Economics Letters*, 56(2), 181–186.

Russell, T. (2003). How quasi rational are you II? chern curvature measures local failure of the expected utility maximization axioms. *Economics Letters*, 81(3), 379–382.

Varian, H. R. (1990). Goodness-of-fit in optimizing models. *Journal of Econometrics*, 46(1), 125–140.

Appendix

Proof of Claim 1:

Proof. First, we show that$\mathcal{X}(Z)$ is closed. Take any sequence $\{x^n\}_{n\in\mathbb{N}}$ of demand functions with $x^n \in \mathcal{X}(Z)$. For any $n \in \mathbb{N}$, we have $p'x^n(z) = w$ by assumption. Let the arbitrary convergent sequence be such that $x^n(z) \to x(z)$ to some function $x \in \mathcal{C}^1(Z)$. Then we want to show that, for $x^n(z) = [x_1^n(z)\cdots x_L^n(z)]$, if $x_l^n(z) \to x_l(z)$ then $p'x(z) = w$. By Walras' law, we have $p'x^n(z) = w$. Define the function $g^n(z) = p'x^n(z) - w$, taking limits on this new function $lim_{n\to\infty} g^n(z) = lim_{n\to\infty}[\sum_{l=1}^{L} p_l x_l^n(z) - w]$ it follows that $g^n(z) \to 0$ and since $z = (p, w)$ and $z \in \mathbb{R}_{++}^{L+1}$ is given we can conclude that $p'x(z) = w$. This implies that $x \in \mathcal{X}(Z)$.

We also show that$\mathcal{X}(Z)$ is uniformly bounded. Since $x^\tau \in \mathcal{X}(Z)$ is continuously differentiable and Z is compact, then every $x^\tau(z)$ is compact-valued, and so is $\nabla x^\tau(z)$. Then, there exists a bound $M_{x^\tau} > 0$ such that $||x_l^\tau||_{C1,1} \leq M_{x^\tau}$ for $l = 1, \dots, L$. Then let $M = max\{M_{x^\tau}\}_{x^\tau\in\mathcal{X}(Z)}$. It follows that $||x_l^\tau||_{C_1,1} \leq M$ for $l = 1, \dots, L$, for all $z \in Z$, and for all $x^\tau \in \mathcal{X}(Z)$.

Finally, we demonstrate that $\mathcal{X}(Z)$ is equicontinuous. This is a direct consequence of the assumption of continuous differentiability and the compactness of Z. In fact, under the assumptions of continuity of $Dx^\tau(z) = [D_p x^\tau(z) \quad D_w x^\tau(z)]$, then set $\{\nabla x_l^\tau(z) \in \mathcal{C}(Z)\}$ is uniformly bounded in Z (by the same uniform boundedness argument of $\mathcal{X}(Z)$). Also, by the multivariate mean value theorem applied to each function $x_l^\tau(z)$, it follows that for $l = 1, \dots, L$, for every $u > 0$ there exists a $v_l > 0$ ($v_l = M|$ such that for $\overline{z} \in [z_1, z_2]$: $||\nabla x_l^\tau(\overline{z})||_{\infty,L+1} < M$), depending only on u, such that for a pair $z_1, z_2 \in Z$: $d(z_1, z_2) < v_l \implies ||x_l^\tau(z_1) - x_l^\tau(z_2)||_{C1,1} < u$ for all entries of the functions $x^\tau \in \mathcal{X}(Z)$. Under the $||\cdot||_{C1}$ norm for $\mathcal{X}(Z)$, it follows that $||x^\tau(z_1) - x^\tau(z_2)||_{C1} < u$.[24] Finally we choose $\bar{l} = argmax_l\{||x_l^\tau(z_1) - x_l^\tau(z_2)||_{C1,1}\}$ and we fix $v = v_{\bar{l}}$. Then we can conclude that for every $u > 0$ there exists a $v > 0$ which depends only on u (since all v_l depend only on u), such that for a pair $z_1, z_2 \in Z$: $d(z_1, z_2) < v \implies ||x^\tau(z_1) - x^\tau(z_2)||_{C1} < u \ \forall \ x^\tau \in \mathcal{X}(Z)$.

To conclude, we apply the Ascoli-Arzelà theorem to the family of functions $x^\tau \in \mathcal{X}(Z)$. Since $\mathcal{X}(Z)$ is closed, uniformly bounded and equicontinuous, it is a compact subset of $\mathcal{C}^1(Z)$. $\square$

Claim 5

The following claim is an auxiliary result to be used in the sequel.

Claim 5. The map $s : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ defined as $s(x^\tau) = S^\tau$ is continuous.

Proof. First, we will prove that $D_p : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ and $D_w : \mathcal{X}(Z) \mapsto \mathcal{C}(Z, \mathbb{R}^L)$ 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 $\mathcal{X}(Z), \mathcal{M}(Z)$ are Banach spaces with the norms $||\cdot||_{C1}$ 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.

[24]Three different norms are used in this proof. The partial derivatives of x_l^τ are not required to be differentiable hence the norm in this space is the supremum norm $||\cdot||_\infty$. By contrast, $x_l^\tau$is continuously differentiable and has norm $||\cdot||_{C1,1}$. For a fixed $z = \overline{z}$, $||x^\tau(\overline{z})||_{C1,1} = max(||x^\tau(\overline{z})||_{\infty,1}, ||\nabla x^\tau(\overline{z})||_{\infty,L+1})$. Finally, the norm in $\mathcal{X}(Z)$ is $||\cdot||_{C1}$ as defined in Section 2.

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

Proof of Claim 2

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

Under the Frobenius norm, the minimization problem amounts to finding the solution to
$min_{S^r} \int_{z\in Z} Tr([S^\tau(z) - S^r(z)]'[S^\tau(z) - S^r(z)])dz$
subject to
$S^r(z) \leq 0$
$S^r(z) = S^r(z)'$
$S^r(z)p = 0$

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

The constraint set $\mathcal{M}_{\mathfrak{R}}(Z)$ is convex and closed. In fact, the cone of negative semi-definite 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 $\mathcal{M}_{\mathfrak{R}}(Z)$ 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 $\mathcal{M}_{\mathfrak{R}}(Z)$ is itself a subspace of $\mathcal{M}(Z)$.

Now we prove that not only the symmetric and the NSD constraints sets are closed but all $\mathcal{M}_{\mathfrak{R}}(Z)$ is closed. 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 $\Lambda^n(z) = Diag[\lambda_i^n(z)]_{i\in 1,\ldots L}$ 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 $\lambda_1^n(z) \leq \lambda_2^n(z) \leq \ldots \leq 0$ where the order is induced by the distance to the null function using the euclidean distance for scalar functions defined over Z.[25] The matrix function $Q^n(z) = [q_1^n \cdots p]$ is the orthogonal matrix with eigenvectors functions as columns. For all $D^n(z) \in \mathcal{M}_{\mathfrak{R}}(Z)$, $\lambda_L^n = 0$ is associated with the price vector $q_L^n = p$ always, to guarantee that p is in its null space. The eigenvectors are defined implicitly by the condition $D^n(z)q_i^n(z) = \lambda^n(z)q_i^n(z)$ with pointwise matrix and vector multiplication and $q_i^n \perp p$ or $\langle q_i^n, p\rangle = 0$ using the inner product for $\mathcal{C}^0(Z)$ 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 $D^n(z) \in \mathcal{M}_{\mathfrak{R}}(Z)$ for each $z \in Z$, with limit $lim_{n\to\infty} D^n(z) = D(z)$. We want to show that $D(z) \in \mathcal{M}_{\mathfrak{R}}(Z)$. It should be clear that any $D^n(z) \to D(z)$ 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 $\lambda_L^n = 0$ for all n and certainly $\lambda_L^n \to 0$ with the associated eigenvector $q_L^n = p$ for all n and $q_L^n \to p$. This condition, along with

[25] This eigenvalue functions can be labeled because the domain Z is simply connected. The order admits crossing eigenvalue functions.

symmetry, guarantees that $q^n_{i\neq L} \to q \perp p$. Finally, the set of negative scalar functions is closed. Then, $\lambda^n_{i\neq L} \to \lambda(z)_-$ with $\lambda_{i\neq L}(z)_- = min(0, \lambda_{i\neq L}(z))$. This is a negative scalar function by construction, since if $\lambda_{i\neq L}(z) > 0$ then $\lambda_{i\neq L}(z)_- = 0$. Then $\mathcal{M}_{\mathfrak{R}}(Z)$ is closed.

Let $\mathcal{M}_S(Z) \subset \mathcal{M}(Z)$ be the subset of Slutsky matrix functions, defined as the image of the Slutsky map defined over $\mathcal{X}(Z)$. Let $s : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$, be a matrix map defined by $s(x^\tau) = D_p x^\tau + D_w x^\tau [x^\tau]'$, i.e $s(\mathcal{X}(Z)) \equiv \mathcal{M}_{\mathcal{S}}(Z)$. By Claim (1), $\mathcal{X}(Z)$ is compact and by continuity of the s map (proven in Claim 5), it follows that $\mathcal{M}_{\mathcal{S}}(Z)$ is a compact set. Since $\mathcal{M}_{\mathfrak{R}}(Z) \subset \mathcal{M}_S(Z)$, and given that $\mathcal{M}_{\mathfrak{R}}(Z)$ is a closed subspace of a compact set, then $\mathcal{M}_{\mathfrak{R}}(Z)$ is also compact.

Note also that the image of the feasible set satisfying the constraints is closed (because all constraint sets images are pointwise subspaces of euclidean metric spaces of finite dimension and therefore are closed) and convex because it is the intersection of three convex sets. Under the assumption of $z \in Z$ for Z compact, then it follows that the constraint set is pointwise compact. To see why the previous statement is true observe that the set of images of $\mathcal{M}(Z)$ for a fixed z, $Im_z(\mathcal{M}(Z))$ consists of real-valued $L \times L$ matrices that forms a vector space that is isomorphic to the euclidean space $\mathbb{R}^{L^2}$. Then let $\mathcal{M}_{\mathfrak{R}}(Z)$ be the set of matrix functions that have property $\mathfrak{R}$. It follows that $Im_z(\mathcal{M}_{\mathfrak{R}}(Z)) \subset Im_z(\mathcal{M}(Z))$, is compact if and only if it is closed and bounded. Observe first that $\mathcal{M}_{\mathfrak{R}}(Z)$ is closed since it is in the intersection of three closed sets. The entries of a matrix $S \in \mathcal{R}$ are not necessarily bounded for all $(p, w) \in \mathbb{R}^{L+1}$, but $Im_z(\mathcal{M}_{\mathfrak{R}}(Z))$ is bounded. Then $\mathcal{M}_{\mathfrak{R}}(z)$ is pointwise compact.

In conclusion, since the Frobenius norm in $\mathcal{M}(Z)$ is a continuous and strictly convex functional and the constraint set is compact and convex the minimum is attained and it is unique. $\square$

Proof of Claim 4

Proof. First, we establish the basic properties of this new measure:

Positive: $\rho(S^\tau) \geq 0$ by construction (if $S^\tau \in \mathcal{M}(Z)$).

Unit-free: $\rho(cS^\tau) = \rho(S^\tau)$ (if $S^\tau \in \mathcal{M}(Z)$ and $c \in \mathbb{R}$).

$\rho(cS^\tau) = \frac{|c|d(S^\tau)}{|c|||S^\tau||} = \rho(S^\tau)$

Let $E^\sigma \neq 0$, $E^\pi \neq 0$ and $E^\nu \neq 0$. Next, we establish each of the componentwise bounds.

We write
$$\frac{||E^\sigma||}{||S^\tau||} = \frac{||\frac{1}{2}[S^\tau - S^{\tau'}]||}{||S^\tau||} \leq \frac{|\frac{1}{2}|2||S^\tau||}{||S^\tau||} = 1.$$

Next, since

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

By Walras' law:

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

then

$||E^\pi|| \leq \frac{1}{p'p}[||S^\sigma(z)|| \cdot ||pp'||].$

Note that $||pp'|| = p'p$ then

$||E^\pi|| \leq ||S^\sigma(z)||$

and

$\frac{||E^\pi||}{||S^\tau||} \leq \frac{||S^\sigma(z)||}{||S^\tau||} \leq 1.$

Next, since $S^{\sigma,\pi} = S^\sigma + E^\pi = \frac{1}{2}[S^\tau + S^{\tau'}] + E^\pi$

then $||S^{\sigma,\pi}|| \leq 2||S^\sigma||$

and
$||S^{\sigma,\pi}|| \leq 2||S^{\tau}||$.
It follows that:
$\frac{||E^{\nu}||}{||S^{\tau}||} \leq 2\frac{||E^{\nu}||}{||S^{\sigma,\pi}||} \leq 2$.
Finally, recall that
$\rho(S^{\tau}) = \frac{||E||}{||S^{\tau}||}$
and
$\rho(S^{\tau})^2 = \frac{||E^{\sigma}||^2+||E^{\pi}||^2+||E^{\nu}||^2}{||S^{\tau}||^2} = \frac{||E^{\sigma}||^2}{||S^{\tau}||^2} + \frac{||E^{\pi}||^2}{||S^{\tau}||^2} + \frac{||E^{\nu}||^2}{||S^{\tau}||^2}$.
Using the componentwise bounds afor established yields the overall bound of $\sqrt{6}$.

□

Lemma 3

Lemma 3. *The solution to $min_A||S^{\tau} - A||$ subject to $A(z)p = 0$ and $A(z)$ symmetric is the nearest matrix function with this characteristics for $S^{\sigma}(z)$.*

Proof. Using the symmetric, skew symmetric matrix decomposition of S^{τ} we can write:
$||S^{\tau} - A||^2 = ||S^{\sigma} + -A(z) + E^{\sigma}||^2 = ||S^{\sigma} - A||^2 + ||E^{\sigma}||^2 + 2\langle S^{\sigma} + -N, E^{\sigma}\rangle$

Since $S^{\sigma}(z) - A(z)$ is symmetric, it follows that $\langle S^{\sigma} + -N, E^{\sigma}\rangle = 0$, because $Tr([S^{\sigma}(z) + -A(z)]'E^{\sigma}(z)) = 0$, for $z \in Z$. In fact, the trace of the product of a symmetric matrix-valued function and skew symmetric valued function is zero for any $z \in Z$.

This implies that the proposed program can be written as:
$max_A||S^{\sigma} + E^{\sigma} - A(z)||^2 = ||S^{\sigma} - A(z)||^2 + ||E^{\sigma}||^2$
with $A(z) = A(z)'$ and $A(z)p = 0$.

With the solution $A(z)^* = S^{\sigma}(z) + E^{\pi}(z)$, with $E^{\pi}(z)$ the nearest matrix function under the Frobenius norm that makes $[S^{\sigma}(z) + E^{\pi}(z)]p = 0$.

Thus, $A(z)^* = S^{\sigma,\pi}(z)$. □

Lemma 4

Lemma 4. *The solution to the problem with only the symmetric restriction active is $S^{\sigma}(z) = \frac{1}{2}[S^{\tau}(z) + S^{\tau}(z)']$*

Proof. Consider the reduced problem with only the symmetry restriction active:
$min_{S^r} \int_{z \in Z} Tr([S^{\tau}(z) - S^{r}(z)]'[S^{\tau}(z) - S^{r}(z)])dz$
subject to
$S^{r}(z) = S^{r}(z)'$.
This is equivalent to minimize
$S^{\sigma}(z) = argmin_{S^r \in \mathcal{R}} \int_{z \in Z}[Tr(S^{\tau}(z)'S^{\tau}(z)) - Tr(S^{r}(z)S^{\tau}(z)) - Tr(S^{\tau}(z)'S^{r}(z)) + Tr(S^{r}(z)S^{r}(z))]dz$

The optimization solution when only the symmetry restriction is active is obtained using the strong Euler-Equation. The objective function expansion is obtained using the linearity of trace operator. We also use the property that $\partial Tr(AX')/\partial X = A$. Note that the Lagrangian only depends on the function $S^{r}(z)$ and not on z or its derivatives. Therefore the first order pointwise conditions give:
$-S^{\tau}(z) - S^{\tau}(z)' + 2S^{r}(z) = 0$, for $z \in Z$
which yields

$S^\sigma(z) = \frac{1}{2}[S^\tau(z) + S^\tau(z)']$.

It is useful to note that this corresponds to the symmetric decomposition of a square matrix function. □

Lemma 5

Lemma 5. *The negative semi-definite nearest matrix function to $S^\sigma(z)$ is also the NSD nearest matrix function to $S^r(z)$.*

Proof. It is a known fact that the nearest NSD matrix to $S^\sigma(z)$ is the projection of this function on the NSD cone (Higham, 1989). That is, it is the negative semidefinite part of the matrix $S^\sigma(z)$.

To obtain it, let $S^\sigma(z) = Q(z)\Lambda(z)Q(z)'$, where $\Lambda(z)$ is the diagonal matrix of eigenvalues, (with ordered entries), of $S^\sigma(z)$ and $Q(z)$ is an orthogonal matrix whose columns are the eigenvectors associated with $\Lambda(z)$. Every real, symmetric matrix has such decomposition.

Also note that one can decompose the symmetric matrix $S^\sigma(z)$ into its PSD part and NSD part: $S^\sigma(z) = S^\sigma_+(z) + S^\sigma_-(z)$. Here, $S^\sigma_+ = \sum_{\lambda_i>0} \lambda_i q_i q_i'$ and $S^\sigma_- = \sum_{\lambda_i<0} \lambda_i q_i q_i'$. We abuse notation and let $\lambda_i < 0$ denote $\lambda_i^- = min(0, \lambda_i)$ and $\lambda_i > 0$ represent $\lambda_i^+ = max(0, \lambda_i)$.

In other words, $S^\sigma(z)_- = Q(z)[\Lambda(z)_-]Q'$, with $\Lambda_-(z) = diag(min(\lambda_i, 0))_{i\in 1\cdots L}$.

We want to solve

$max_N ||S^\tau - N||$

subject to $N(z) \leq 0$, that is negative semidefinite.

Notice that, $S^\tau(z) = S^\sigma(z) + E^\sigma(z)$, can always be decomposed in the sum of its symmetric and skew-symmetric part.

Then the objective functional can be written as

$||S^\tau - N|| = ||S^\sigma + E^\sigma - N||$

Then:

$||S^\tau - N||^2 = ||S^\sigma + -N + E^\sigma||^2 = ||S^\sigma - N||^2 + ||E^\sigma||^2$

Because $S^\sigma(z) - N(z)$ is symmetric it follows that $\langle S^\sigma(z) + -N(z), E^\sigma(z)\rangle = 0$, since $Tr([S^\sigma(z) + -N(z)]'E^\sigma(z)) = 0$, for any z. That is the trace of the product of a symmetric matrix-valued function and skew symmetric valued function is zero for any $z \in Z$.

This implies $N^*(z) = S^\sigma(z)_-$. The solution is the negative semidefinite part of $S^\sigma(z)$. □

Claim 6

Claim 6. 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 $\mathcal{A}(z)$ of symmetric singular matrix functions with the property that $p'A(z)p = 0$ as follows: $\mathcal{A}(z) = \mathcal{P}(z) \oplus \mathcal{N}(z)$ for all $z \in Z$, where

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

and

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

To see that this is a direct sum decomposition, first observe that $\mathcal{P}(z) \cap \mathcal{N}(z) = \{0\}$, with 0 denoting the zero matrix function, by construction. Furthermore, any $A(z) \in \mathcal{A}(z)$ 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$\mathcal{N}(z)$ is a subspace in the space of symmetric matrix-valued functions, it follows that $\mathcal{P}(z)$ 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 $z \in Z$. □

Claim 7

Claim 7. The map $a : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ defined element-wise as $a(x^{\tau}) = S^{r} - S^{\tau}$ is continuous.

Proof. The continuity of the map a follows directly from the continuity of the Slutsky map s and the continuity of the projections maps that generates S^{r}. By Claim 5, we know that $s : \mathcal{X}(Z) \mapsto \mathcal{M}(Z)$ is continuous. It remains to be shown that the projection maps are indeed continuous. For this we need that the range of the projection map is a closed subspace under the metric induced by the norm of $\mathcal{M}(Z)$. The first projection is $p_1 : s(\mathcal{X}(Z)) \mapsto Sym(\mathcal{M}(Z))$ with range equal to the closed subspace of symmetric matrix-valued functions on $\mathcal{M}(Z)$, therefore p_1 is continuous. The second projection is $p_2 : Sym(\mathcal{M}(Z)) \mapsto \mathcal{P}(Z)$ with $\mathcal{P}(Z)$ defined as in Claim 6. The closedness of $\mathcal{P}(Z)$ is not trivial and is proved now. Take a sequence of matrices $\{E^{\pi,n}\}_{n\in\mathbb{N}}$ in $\mathcal{P}(Z)$, now consider the definition of this space and it must be the case that $Tr(E^{\pi,n}pp') = 0$ with $E^{\pi,n}(z)p \neq 0$ for $E^{\pi,n}(z) \neq 0$. Taking limits in the three conditions (one equality and two inequalities), it follows that $lim_{n\to\infty}E^{\pi,n}(z) = E^{\pi} \in \mathcal{P}(Z)$, therefore $\mathcal{P}(Z)$ is closed and p_2 is a continuous map. Finally, the third projection $p_3 : Sym(\mathcal{M}(Z)) \oplus \mathcal{P}(Z) \mapsto Sym(\mathcal{M}(Z))_{-}$ that is with range equal to the closed cone of negative semidefinite matrix-valued functions is also a continuous map. It follows that a is continuous. □

Proof of Remark

If Walras' law hold then $|q'E^{\pi}(p,\overline{w})r| = 0$ for any $q, r \in B_p$.

In particular for a fixed $\overline{w}$:

$|q'E^{\pi}(p,\overline{w})r| = |-q'\frac{1}{p'p}[S^{\sigma}(p,\overline{w})pp' + pp'S^{\sigma}(p,\overline{w}) - \frac{[S^{\sigma}(p,\overline{w})p]'p}{p'p}pp']r| = 0$

The proof is separated in three parts:

(i) First component:

$E^{\pi,1}(p) = -q'\frac{1}{p'p}S^{\sigma}(p,\overline{w})pp'r = q'[D_p x^{s,w}(p) - \frac{1}{w}D_p x^{s,w}(p)px^{s,w}(p)']p\frac{p'r}{p'p}$

Since $x^{s,w}(p)'p = w$ then $E^{\pi,1} = [q'D_p x^{s,w}(p)p - q'D_p x^{s,w}(p)p]\frac{p'r}{p'p} = 0$.

(ii) Second component:

$E^{\pi,2}(p) = q'\frac{1}{p'p}pp'S^{\sigma}(p,\overline{w})r = \frac{q'p}{p'p}p'[D_p x^{s,w}(p) - \frac{1}{w}D_p x^{s,w}(p)px^{s,w}(p)']r$

Deriving $p'x^{s,w}(p,\overline{w}) = \overline{w}$ with respect to prices, it follows that $p'D_p x^{s,w}(p) = -x^{s,w}(p)'$.

$E^{\pi,2}(p) = \frac{q'p}{p'p}[-x^{s,w}(p)'r + x^{s,w}(p)'r] = 0$

(iii) Third component:

$E^{\pi,3}(p) = q'\frac{1}{p'p}\frac{[S^\sigma(p,\overline{w})p]'p}{p'p}pp'r = \frac{1}{p'p}[q'[S^\sigma(p,\overline{w})p]'pp]\frac{p'r}{p'p}$

By definition of the conditional Slutsky matrix function for a fixed $\overline{w}$:

$E^{\pi,3}(p) = \frac{1}{p'p}[q'[p'D_p x^{s,w}(p)'p - \frac{1}{w}p'x^{s,w}(p)p'D_p x^{s,w}(p)'p]p]\frac{p'r}{p'p}$

$E^{\pi,3}(p) = \frac{1}{p'p}[q'[w - \frac{1}{w}p'x^{s,w}(p)w]p]\frac{p'r}{p'p}$

$E^{\pi,3}(p) = \frac{q'p}{p'p}\frac{p'r}{p'p}[w - p'x^{s,w}(p)] = 0$ when Walras law hold.

Proof of Remark 6

$max_A||S^\tau - A||_\$$ with A satisfying $\mathfrak{R}$.

Then notice that $||S^\tau - A||_\$ = ||\overline{S^\tau - A}||$

$||\overline{S^\tau - A}||^2 = ||\Lambda_p[S^\tau - A]\Lambda_p||^2 = ||\Lambda_p[S^\sigma - A + E^\sigma]\Lambda_p||^2$

$= ||\Lambda_p[S^\sigma - A]\Lambda_p||^2 + ||\Lambda_p E^\sigma \Lambda_p||^2$.

Observe that if A is symmetric then $\Lambda_p A \Lambda_p$ is symmetric, if A is skew-symmetric then $\Lambda_p B \Lambda_p$ is also skew-symmetric. This means that $\langle \Lambda_p E^\sigma \Lambda_p, \Lambda_p[S^\sigma - A]\Lambda_p\rangle = 0$.

Similarly, for the general case:

$||\Lambda_p[S^\tau - A]\Lambda_p||^2 = ||\Lambda_p[S^{\sigma,\pi} - A + E^\sigma - E^\pi]\Lambda_p||^2 = ||\Lambda_p E^\sigma \Lambda_p||^2 + ||\Lambda_p[-E^\pi + S^{\sigma,\pi} - A]\Lambda_p||^2$

Where the inner decomposition of S^τ comes from the following Lagrangian:

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

Equivalently,

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

FOC:

$S^{\sigma,\pi}(z) = \frac{1}{2}[S^\tau(z) + S^\tau(z)'] + \Lambda_p^{-2}[\lambda p' - U + U']\Lambda_p^{-2}$;

$S^{\sigma,\pi}(z)p = 0$

$S^{\sigma,\pi}(z) = S^{\sigma,\pi}(z)'$

Postulate that $\Lambda_p^{-2}[\lambda p' - U + U']\Lambda_p^{-2} = E^\pi$. This proposed solution satisfies the FOC an by uniqueness of the solution due to the nature of the objective functional one has that $S^{\sigma,\pi}(z) = S^\sigma(z) + E^\pi$ under the dollar norm and $||S^\tau - S^{\sigma,\pi}||_\$ = ||E^\sigma - E^\pi||_\$$. Therefore the general matrix nearness problem with the dollar-norm is equivalent to computing:

$S^{\overline{\tau}} = argmin_{A\leq 0}||\Lambda_p[S^{\sigma,\pi} - A]\Lambda_p||^2 - 2\langle \Lambda_p E^\pi \Lambda_p, \Lambda_p[S^{\sigma,\pi} - A]\Lambda_p\rangle$, where $S^{\overline{\tau}}$ has property $\mathfrak{R}$. This problem has a closed-form solution.

A Simplified Algorithm for the Plug-In Estimator of the Distance to Rationality

We can write $\gamma = \int_Z \frac{g(z)}{h(z)}h(z)dz$ or $\gamma = \mathbb{E}[\frac{g(P,W)}{h(P,W)}]$ where h is the probability density function (pdf) of (P,W) that is assumed to be bounded away from zero and whose support is Z. This suggests a simple semiparametric estimation procedure. The functional statistic of interest is given implicitly by the moment $\mathbb{E}[m(z,\gamma,g)] = 0$, where $m(z,\gamma,g) = \frac{g(z)}{h(z)} - \gamma$. In this setting, we use the analogy principle to estimate γ in two parts. First, a nonparametric step to compute $\hat{g}, \hat{h}$ and, second, a parametric step in which we can use the sample average to estimate γ:

$$\hat{\gamma} = \frac{1}{N}\sum_{i=1}^{n}\frac{1}{\hat{h}(z^i)}\hat{g}(z^i).$$

This approach has a computational advantage in the estimation of g, but does not provide an estimator for the matrix function E.

The construction of $\hat{g}$ takes several nonparametric steps that nevertheless are simple to obtain due to our analytical approach. In short, we need a nonparametric consistent estimator $\hat{\alpha}^{\tau}$ and its partial derivatives (that converge uniformly to the true functions). We assume that we have them. Then, we can build an estimator of $\hat{S}^{\tau}(p^i, w^i)$ for $i \in \{1, \cdots, N\}$.[26] Finally, using our analytical results pointwise, we can get an estimator of $\hat{E}(p^i, w^i)$. Then, we use this object to build $\hat{g}(p^i, w^i)$ also pointwise, and this helps us to obtain $\hat{\gamma}$.

The construction of $\hat{E}^{\sigma}(p^i, w^i)$ and $\hat{E}^{\pi}(p^i, w^i)$ is computationally trivial since they only require $\hat{S}^{\tau}(p^i, w^i)$. An important computational step in the estimation of g is the pointwise projection of $\hat{S}^{\sigma,\pi}(p^i, w^i)$ in the NSD space of matrix functions to compute $\hat{E}^{\nu}(p^i, w^i)$. Following our results, this problem amounts to computing an eigen-value decomposition of the numeric matrix $\hat{S}^{\sigma,\pi}(p^i, w^i)$ at each i. The steps are summarized below:

- Input: Consistent nonparametric estimates of α^{τ} and its partial derivatives, and h.
- Result: The semiparametric estimator$\hat{\gamma}$.
- First, build $\hat{S}^{\tau}(p^i, w^i)$ for every pair (p^i, w^i), construct $\hat{S}^{\sigma,\pi}(p^i, w^i)$ and $\hat{E}^{\sigma}(p^i, w^i)$, $\hat{E}^{\pi}(p^i, w^i)$.
- Second, obtain $\hat{E}^{\nu}(p^i, w^i) = \hat{S}^{\sigma,\pi}(p^i, w^i)_+$ as the numeric matrix projection on the NSD matrix cone.
- Third, obtain $\overline{g}^i = Tr(\hat{E}(p^i, w^i)'\hat{E}(p^i, w^i))$ for all $i \in \{1, \cdots, n\}$.
- Finally, Compute the sample average $\hat{\gamma} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\hat{h}(p^i,w^i)}\overline{g}^i$.

The semiparametric estimator $\hat{\gamma} \xrightarrow{p} \gamma$ as $n \to \infty$ if the estimators of α^{τ} and its derivatives converge in probability to their true value.

We assume we have a local kernel polynomial estimator of α^{τ} and its partial derivatives such that if $n \to \infty$, then $\hat{s}^{\tau}(p^i, w^i) \xrightarrow{p} s^{\tau}(p^i, w^i)$, where $\hat{s}^{\tau}(p^i, w^i)_{l,k} = \widehat{\partial_{p_k}\alpha_l^{\tau}} + \widehat{\partial_w \alpha_l}(p^i, w^i)\hat{\alpha_k}(p^i, w^i)$.[27] Then, we have an entrywise convergent estimator of the Slutsky matrix $\hat{S}^{\tau}(p^i, w^i) \xrightarrow{p} S^{\tau}(p^i, w^i)$. If this holds, we can get a pointwise convergent estimator of $g(p^i, w^i)$ (i.e., the image of g at (p^i, w^i)) for all i by applying the continuous mapping theorem). In fact, we know that $S^{\sigma}, S^{\sigma,\pi}$ and $S^{\sigma,\pi,\nu}$ are obtained as projections on closed subspaces, and therefore, each of the projections is a continuous map. Fix (p^i, w^i), if $\hat{S}^{\tau}(p^i, w^i) \xrightarrow{p} S^{\tau}(p^i, w^i)$, first observe that $\hat{E}^j(p^i, w^i) = [\hat{S}^{\tau}(p^i, w^i) - \mathcal{P}_j(\hat{S}^{\tau,i}(p^i, w^i))]$ for $j \in \{\sigma, \pi, \nu\}$, where $\mathcal{P}_j$ is the projection in the closed subspace of matrices with property j (i.e. $\mathcal{P}_j$ is a continuous map). Then, by the continuous mapping theorem, it follows that $\hat{E}^j(p^i, w^i) \xrightarrow{p} E^j(p^i, w^i)$. Then, for (p^i, w^i), the $\overline{g}^{i,j} = Tr(\hat{E}^j(p^i, w^i)'\hat{E}^j(p^i, w^i))$ also converges in probability to $g^j(p^i, w^i)$ for all j since the trace operator is continuous. Moreover, it is smooth. Finally, because of the additive separability of the components of g, we have $\overline{g}^i = \overline{g}^{\sigma,i} + \overline{g}^{\pi,i} + \overline{g}^{\nu,i}$. We conclude that $\overline{g}^i \xrightarrow{p} g(p^i, w^i)$.

In the case of the semiparametric estimator, consistency is guaranteed because $\hat{\gamma}$ is the sample average of $\frac{\overline{g}^i}{\hat{h}(p^i,w^i)} \xrightarrow{p} \frac{g(p^i,w^i)}{h(p^i,w^i)}$ under the assumptions on h and g.

[26] We can construct $\hat{S}^{\tau}$ using a local polynomial estimator of the Slutsky matrix function as proposed by (Haag et al., 2009).

[27] Here, $\widehat{\partial_{p_k}\alpha_l^{\tau}}$ is the estimator of $\frac{\partial \alpha_l(p,w)}{\partial p_k}|_{(p^i,w^i)}$and $\widehat{\partial_w \alpha_l}$ is defined analogously.