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

Victor H. Aguiar† and Roberto Serrano‡

This version: June 2017

Abstract

Given any observed demand behavior —by means of a demand function—, we quantify by how much it departs from rationality. The measure of the gap is the smallest Frobenius norm of the correcting matrix function that would yield a Slutsky matrix with its standard rationality properties (symmetry, singularity, and negative semidefiniteness). As a result, we are able to suggest a useful classification of departures from rationality, corresponding to three anomalies: inattentiveness to changes in purchasing power, money illusion, and violations of the compensated law of demand. Errors in comparative-statics predictions from assuming rationality are decomposed as the sum of a behavioral error (due to the agent) and a specification error (due to the modeller). Illustrations are provided using several bounded rationality models.

JEL classification numbers: C60, D10.

Keywords: consumer theory; rationality; Slutsky matrix function; bounded rationality; comparative statics; sparse-max consumer, collective model.

*This paper subsumes Aguiar and Serrano (2014). We are especially indebted to Xavier Gabaix, Michael Jerison, and Joel Sobel for suggesting many specific improvements to that earlier version of the paper. We also thank Bob Anderson, Francis Bloch, Mark Dean, Federico Echenique, Drew Fudenberg, Peter Hammond, Susanne Schennach, Larry Selden, Jesse Shapiro and the participants at numerous conferences and seminars for helpful comments and encouragement. The comments of an editor and three referees of this journal were also very useful. We thank Judith Levi for her excellent editing job.

†Department of Economics, University of Western Ontario, vaguiar@uwo.ca

‡Department of Economics, Brown University, roberto_serrano@brown.edu
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 (with infinitely divisible goods), this model makes a unique prediction in the form of a symmetric and negative semidefinite Slutsky matrix (which, as a corollary, must also be singular). 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, the aforementioned predictions (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. This paper aims to unify and systematize the implications of many of these models. In doing so, we use a well-known tool in microeconomic theory, namely, the Slutsky matrix.

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?

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 one 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. However, this approach presents serious difficulties. 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 the Slutsky matrix function of this demand 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 better computational implementability, and the derivation and interpretation of optimal solutions. Indeed, we attempt to find the smallest additive perturbation that corrects 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,\(^1\) and negative semidefiniteness). We use the Frobenius norm to measure the size of this additive factor, interpreting it as the size of the observed departure from rationality. The Frobenius norm is helpful in capturing the violations of the different regularity properties; thanks to the orthogonalities it furnishes, it provides a clean measurement of such violations.

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

(a) the norm of its antisymmetric 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

\(^1\)We abuse terminology slightly here. By “singularity with the price vector on its null space” or “singularity in \(p\),” we mean that \(p\) is a right eigenvector of the Slutsky matrix associated with a zero eigenvalue, since Walras’ law (assumed throughout the paper) implies that \(p\) is a left eigenvector of the matrix.
measures the size of the violation of negative semidefiniteness.

Our approach captures and measures anomalies in comparative statics with respect to the rational consumer. Indeed, we observe the following:

(a’) The symmetry property can be identified with the Ville axiom of revealed preference; hence, its violation amounts to “inattentiveness to changes in purchasing power” (ICPP).

(b’) The failure of singularity reveals the presence of “money illusion” (MI).

(c’) And the failure of negative semidefiniteness reveals a “violation of the compensated law of demand” (VCLD).

Our main result shows that the size of bounded rationality (BR) measured by the Slutsky matrix squared norm, can be decomposed into the sum of the squares of these three effects. In terms of the different anomalies, our main equation can be expressed as:

\[ BR^2 = ICPP^2 + MI^2 + VCLD^2. \]

In particular, following any observed behavior, we can classify the instances of bounded rationality as anomaly ICPP when symmetry fails, anomaly MI when singularity fails, and anomaly VCLD when negative semidefiniteness fails. In the case of more complex failures, by adding up the nonzero components of the norm, we can classify instances of bounded rationality as combinations of these three anomalies.

On the other hand, the decomposition offers new interesting insights. Our results reveal that there are consumers who would appear to violate all Slutsky properties, yet only some of the three anomalies explain the size of bounded rationality. For instance, in the sparse-max model of Gabaix (2014) we find that even when this boundedly rational consumer appears to present the three anomalies, only the ICPP and the MI account for the total size of bounded rationality. This seems to be an important insight for the design of public policy and for the development of marketing strategies by firms, because each anomaly may have different consequences for demand behavior. A medical analogy
may be illustrative here. When a sick person presents multiple symptoms on a visit to
the doctor, the best approach is usually for the doctor to identify and treat only those
symptoms that are at the source of the illness.

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 this with an approach in which we modify the Slutsky
matrix by a weighting matrix. For example, one can translate the norm into dollars, and
thereby provide a monetary measure; or one can instead use a budget-shares version,
which is unit-free.

The Slutsky matrix function is the key object in comparative statics analysis in con-
sumer theory. It encodes all the information about local variations in demand with respect
to small Slutsky compensated price changes. Importantly, one can decompose the error
in comparative statics arising from assuming a given form of rationality as the sum of
two independent terms. The first is the behavioral error, due to the agent (measured by
the Slutsky norm already discussed and its decomposition); and the second, due to the
modeller, is a specification error given the assumed parameterized rationality model. This
finding may help empirical analyses: for instance, we perform an empirical application
in a companion paper (Aguiar & Serrano, 2016), using a widely used experimental data
set.

Although we include a more detailed review of the literature below, the closest pre-
cursor to our work is the approximately rational consumer demand proposed by Jerison
and Jerison (1992; 1993). These papers relate the local violations of (i) negative semidef-
initeness and (ii) symmetry of the Slutsky matrix to the smallest distance between an
observed smooth demand system and a rational demand. Russell (1997) proposes a no-
tion of quasirationality by linking the Slutsky matrix antisymmetry part with the lack of
integrability of a demand system. Unlike these studies, our work is global, and thus not
necessarily tied to a small neighborhood of a price-wealth combination; moreover, it al-
lows for a generalization that treats the three kinds of violations of the Slutsky conditions
simultaneously.
The rest of this paper is organized as follows. Section 2 presents the model and showcases our results with an example of the sparse-max consumer (2014), which we revisit later as an illustration of our different results. Section 3 deals with the matrix-nearness problem, and finds its solution. Section 4 emphasizes the size of bounded rationality additive decomposition, and provides interpretations of the matrix nearness-problem in terms of the axioms of consumer theory. Section 5 presents weighted Slutsky norms. Section 6 reviews further comparative statics and the specification error. Section 7 presents additional examples and applications of the result. Section 8 contains brief guidelines for empirical applications of our methodology to verify the appropriateness of imposing shape constraints (Slutsky properties) in demand estimation. Section 9 presents a review of the literature, and Section 10 concludes. All the proofs are collected in an appendix. A separate online appendix has been prepared to expand on the empirical applications of our methodology, and its connections with other approaches.

2 The Model

2.1 Demand Functions

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

Assume that \(x(p, w)\) is continuously differentiable and satisfies Walras’ law: \(p'x(p, w) = w\) for all \((p, w) \in Z\).\(^2\) Let \(\mathcal{X}(Z) \subset \mathcal{C}^1(Z)\) denote a set of functions that satisfy these characteristics, with \(\mathcal{C}^1(Z)\) denoting the complete metric space, equipped with a norm, of vector-valued functions \(f : Z \mapsto \mathbb{R}^L\), that are continuously differentiable, uniformly bounded with compact domain \(Z \subset \mathbb{R}^{L+1}_+\).

\(^2\)The assumption of Walras’ law simplifies the presentation of our results. We do not view it as a strong assumption; in practice, it means that \(w\) need not be observed.
2.2 The Behavioral Nearness Problem

Our objective in this paper is to shed light on the size and types of bounded rationality. To that end, let \( \mathcal{R}(Z) \subset \mathcal{X}(Z) \) be the set of rational demand functions; that is, \( x^r \in \mathcal{R}(Z) \) is the solution to maximizing a complete, locally nonsatiated, and transitive preference over a linear budget constraint.

**Definition 1.** Define the distance of \( x \in \mathcal{X} \) to the set of rational demands \( \mathcal{R} \) by the least distance from an element to a set: 
\[
    d(x, \mathcal{R}) = \inf \{ d_{\mathcal{X}}(x, x^r) \mid x^r \in \mathcal{R} \}.
\]

We shall refer to this problem of trying to find the closest rational demand to a given demand as the "behavioral nearness" problem (Varian, 1990). Observe that the behavioral nearness problem at this level of generality presents several difficulties. To begin with, the constraint set \( \mathcal{R}(Z) \) (i.e., the set of rational demand functions) is not convex. In addition, the Lagrangian depends not only on \( x^r \) but also upon its partial derivatives. The typical curse of dimensionality of calculus of variations applies here with full force, in the case of a large number of commodities. Indeed, the Euler-Lagrange equations in this case do not offer much information about the problem and give rise to a large partial differential equations system. Finally, calculating analytically the solution to this program is computationally challenging.\(^3\)

2.3 Slutsky Norms and the Matrix Nearness Problem

For the reasons just cited, instead of solving directly the behavioral nearness problem, we adopt an approach that will rely on matrix spaces. Our next goal is to talk about Slutsky norms. Let \( \mathcal{M}(Z) \) be the complete metric space of \( L \times L \) matrix-valued functions, \( F : Z \mapsto \mathbb{R}^{L \times L} \), equipped with the inner product \( \langle F, G \rangle = \int_{z \in Z} \text{Tr}(F(z)G(z))dz \). Let \( \| \cdot \|_\mathcal{M} \) be the Frobenius matrix norm \( \| F(z) \|^2_\mathcal{M} = \text{Tr}(F(z)'F(z)) \) where \( \mathcal{M} = \mathbb{R}^{L \times L} \). This vector space has an (average) Frobenius norm \( \| F \|^2 = \int_{z \in Z} \| F(z) \|^2_\mathcal{M} dz \).

Next, we define the Slutsky substitution matrix function:

---

\(^3\) We refer the reader to our expanded discussion in the appendix, and also in Aguiar and Serrano (2014), which includes the use of the “almost implies near” approach from Anderson (1986), central to the inception of our ideas in this work.
Definition 2. Let $Z \subset P \times W$ be given, and denote by $z = (p, w)$ an arbitrary price-wealth pair in $Z$. Then the Slutsky matrix function $S \in \mathcal{M}(Z)$ is defined pointwise:

$$S(z) = D_p x(z) + D_w x(z) x(z)' \in \mathbb{R}^{L \times L},$$

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

The Slutsky matrix function is well-defined for all $x \in C^1(Z)$. Restricted to the set of rational behaviors, the Slutsky matrix satisfies a number of regularity conditions. Specifically, when a matrix function $S \in \mathcal{M}(Z)$ is symmetric, negative semidefinite (NSD), and singular with $p$ in its null space for all $z \in Z$ (i.e., $S(z)p = 0$), we shall say that the matrix satisfies property $\mathfrak{R}$, for short. We shall also use the short-hand “property $\sigma$” for symmetry, “property $\nu$” for NSD, and “property $\pi$” for singularity with $p$ in its null space ($p$-singularity).

Definition 3. For any Slutsky matrix function $S \in \mathcal{M}(Z)$, let its Slutsky norm be defined as follows:

$$d(S) = \min \{||E|| : S - E \in \mathcal{M}(Z) \text{ having property } \mathfrak{R} \}.$$  

The use of the minimum operator is justified. Indeed, we will prove that the set of Slutsky matrix functions satisfying $\mathfrak{R}$ is a closed 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. We shall interpret its solution as the total size of the consumer’s bounded rationality (BR), and we shall be able to decompose it into the violations of properties $\sigma$, $\nu$, and $\pi$. As explained below, such violations correspond to these three anomalous effects, respectively: inattentiveness to changes in purchasing power (ICPP), violations of the compensated law of demand (VCLD), and money illusion (MI).

2.4 Motivating Example: The Sparse-Max Consumer Model of Gabaix (2014)

Recall the sick person analogy from the introduction. Our results highlight the surprising value of our decomposition. Indeed, there are consumers who would appear to violate

---

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

---
all Slutsky properties, yet they would require that only some of them be treated. This finding might be of interest for developing public policies and marketing strategies. For instance, the following example shows that the departures from $\nu$ in the sparse-max consumer demand proposed by Gabaix are a byproduct of the departures from $\sigma$ and $\pi$.

The sparse-max boundedly rational consumer tries to maximize her continuous, monotone, strictly concave utility given by $u : X \mapsto \mathbb{R}$, subject to her budget constraint $p'x = w$. However, she misperceives prices and instead sees:

$$p^G = Mp + [I - M]p^d,$$

for true prices $p \in P$, default prices $p^d \in \mathbb{R}_+^L$, and $M \in \mathbb{R}^{L \times L}$ a diagonal matrix of attention $M = \text{diag}(\{m_l\}_{l=1}^L)$ for $m_l \in [0, 1]$. Formally, the subject solves the sparse-max problem $\text{smax}_{x|p^G} u(x)$ subject to $p'x = w$. The solution to this problem is the sparse-max demand system:

$$x^G(p, w) = x^r(p^G, w'),$$

which is equivalent to the underlying rational demand $x^r(p, w) = \text{argmax}_x u(x)$, subject to $p'x = w$ evaluated at the perceived prices $p^G$ and at a correcting level of wealth $w' > 0$. We can find $w'$ as a solution to the Walras’ law equation $p'x^G(p, w) = w$, or equivalently, $p'x^r(p^G, w') = w'$.

In this example we fix $L = 2$. Consider a Cobb-Douglas model $x^{CD}(p, w)$, with $\alpha \in [0, 1]$, such that $x^{CD}_1 = \frac{\alpha w}{p_1}$ and $x^{CD}_2 = \frac{(1-\alpha)w}{p_2}$.

Its Slutsky matrix function is:

$$S^{CD}(p, w) = \begin{bmatrix}
-\frac{\alpha(1-\alpha)w}{p_1^2} & \frac{\alpha}{p_1} & \frac{(1-\alpha)w}{p_2} \\
\frac{\alpha}{p_1} & \frac{\alpha}{p_1} & -\frac{(1-\alpha)w}{p_2^2}
\end{bmatrix}.$$
the demand system under $G$ is:

$$x_1^G = \frac{\alpha}{p_1^G} \frac{w}{\alpha \frac{p_1}{p_1^G} + (1 - \alpha) \frac{p_2}{p_2^G}}$$

and

$$x_2^G = \frac{1 - \alpha}{p_2^G} \frac{w}{\alpha \frac{p_2}{p_2^G} + (1 - \alpha) \frac{p_2}{p_2^G}}.$$

This demand system fulfills Walras’ law. This function has an additional parameter, when compared to $x^{CD}(p, w)$, i.e., the perceived price $p_i^G(m_i) = m_i p_i + (1 - m_i) p_i^d$. The vector of attention to price changes $m = (m_1, m_2)'$ weighs the actual price $p_i$ and the default price $p_i^d$.

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

$$M = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, m_1, m_2 \in [0, 1].$$

That is, the consumer may pay less than full attention to changes in $p_1$ or $p_2$, with intensities $m_1$ and $m_2$, $m_i = 1$ stands for full attention to changes in that price, and $m_i = 0$ for complete lack of attention. One of Gabaix’s (2014) elegant results relates the Slutsky matrix function of $x^G$ to the Cobb-Douglas benchmark. The behavioral Slutsky matrix evaluated at default prices (throughout this example, $p$ is approximately $p^d$, so the Slutsky matrix is calculated at $p = p^d$) yields $S^G(p, w) = S^{CD}(p, w)M$. In this case:

$$S^G(p, w) = \begin{bmatrix} -m_1 \frac{\alpha(1-\alpha)w}{p_1^d} & \frac{m_2 \alpha (1-\alpha)w}{p_2} \\ \frac{m_1 \alpha (1-\alpha)w}{p_2} & -m_2 \frac{\alpha (1-\alpha)w}{p_2^d} \end{bmatrix}.$$

This matrix violates properties $\sigma$, $\nu$, and $\pi$ for $m_1 < 1$ or $m_2 < 1$, implying that the consumer is subject to the ICPP, VCLD, and MI anomalies. The nearest Slutsky matrix that has the $\sigma$, $\pi$, and $\nu$ properties when $p = p^d$ is:

$$S^r(p, w) = \frac{m_2 p_2^2 + m_1 p_1^2}{p_1^2 + p_2^2} \begin{bmatrix} -\frac{\alpha(1-\alpha)w}{p_1^d} & \frac{\alpha (1-\alpha)w}{p_1} \\ \frac{\alpha (1-\alpha)w}{p_2} & -\frac{\alpha (1-\alpha)w}{p_2^d} \end{bmatrix}.$$
Surprisingly, $S'(p, w)$ looks like a Cobb-Douglas Slutsky matrix with a scalar perturbation, which goes to one when $m_i \rightarrow 1$.

Thus, the smallest perturbation matrix function that solves the matrix nearness problem defined before is: $E(p, w) = S(p, w) - S'(p, w)$. Here, we have:

$$E(p, w) = \frac{[m_2 - m_1]w^\alpha (1 - \alpha)}{p_1^2 + p_2^2} \begin{bmatrix} 1 & \frac{p_2}{p_1} \\ -\frac{p_1}{p_2} & -1 \end{bmatrix}.$$ 

Now, at a point $(p, w)$, we compute the pointwise Frobenius norm of the matrix function $E(p, w)$:

$$||E(p, w)||_M^2 = \langle E(p, w), E(p, w) \rangle = Tr(E(p, w)'E(p, w)) = \frac{(m_1 - m_2)^2w^2(1 - \alpha)^2\alpha^2}{p_1^2p_2^2}.$$ 

The effect of the attention parameter $m_i \in [0, 1]$ on the size of bounded rationality is transparent: $\frac{\partial}{\partial m_i}||E(p, w)||_M^2 = 2(m_i - m_j)\frac{w^2(1 - \alpha)^2\alpha^2}{p_1^2p_2^2} \leq 0$ when $m_i \leq m_j$. It is nonpositive for all prices, wealth levels, and Cobb-Douglas parameters for $m_i, m_j \in [0, 1]$, and $m_i \leq m_j$. It is positive when $m_i > m_j$. If the absolute value of the difference between $m_1$ and $m_2$ increases, the consumer appears to be at a greater distance from rationality.

We can now compare any two sparse-max consumers $x^{G_1}, x^{G_2}$ according to their relative attention to prices $|m_{1}^{G_k} - m_{2}^{G_k}|$ for $k = 1, 2$, when they have the same $\alpha$ parameter and are facing the same prices and wealth. According to the Slutsky matrix norm, it is clear that $x^{G_1}$ is more boundedly rational than $x^{G_2}$, when $|m_{1}^{G_1} - m_{2}^{G_1}| \geq |m_{1}^{G_2} - m_{2}^{G_2}|$. What matters is the relative difference in attention to the two prices.

Our main result is that the Slutsky matrix norm $||E(p, w)||_M^2$ decomposes:

$$||E(p, w)||_M^2 = ||E^\sigma(p, w)||_M^2 + ||E^\pi(p, w)||_M^2 + ||E^\nu(p, w)||_M^2,$$

where $E^i$ measures the violation of property $i$ ($i = \sigma, \nu, \pi$). Each of these terms can be calculated exactly, as shown in the next section.

It turns out that the additive decomposition for the Gabaix model yields:
\[ ||E^\sigma(p,w)||^2_M = ||E^\pi(p,w)||^2_M = \frac{1}{2}(m_1 - m_2)^2 \frac{w^2(1 - \alpha)^2\alpha^2}{p_1^2p_2^2}, \]

whereas \( ||E^\nu||^2_M = 0 \). That is, regardless of the values that \( w \) takes, the contribution of the first two kinds of violations (\( \sigma \) and \( \pi \)) are equal and amount individually to exactly half of the total distance from rationality, while the NSD component violation vanishes when the prices are evaluated at the default.

Therefore, while the Gabaix consumer violates all regularity properties, our analysis reveals that the anomaly VCLD (connected with property \( \nu \)) is a byproduct of the ICPP and MI anomalies (connected with properties \( \sigma \) and \( \pi \), respectively). The fact that both of the two nonzero violations have equal size suggests that ICPP and MI are of similar importance in this behavior.\(^5\) More formally, after imposing \( \sigma \) and \( \pi \) in this particular model, imposing \( \nu \) does not contribute anything extra for the consumer to behave rationally.

Another advantage of our methodology is that it is global. We can obtain our Slutsky norm for any compact region of prices and wealth, if we have data over such a region. For example, consider a case in which \( w \in [w, \bar{w}] \) and prices are fixed. In this case, we have a measure of the consumer’s total departure from rational behavior, calculated as follows:

\[ ||E||^2 = \frac{1}{2} \int_w^{\bar{w}} \text{Tr}(E^\sigma(w)'E^\sigma(w))dw + \frac{1}{2} \int_w^{\bar{w}} \text{Tr}(E^\pi(w)'E^\pi(w))dw = (m_1 - m_2)^2 \left( \frac{(m_1 - m_2)^2}{3} \right) \frac{(1-\alpha)^2\alpha^2}{p_1^2p_2^2}, \]

with \( p = p^d \).

To showcase the tractability of our approach in this example, consider a simple region \( Z = \{w \in [1, 2], p_1 = 1, p_2 = 1\} \). Then, \( ||E(\alpha, m)||^2 = \frac{7}{3}(m_1 - m_2)^2(1-\alpha)^2\alpha^2 \). We observe that \( \alpha \) has a nonlinear effect on \( ||E(\alpha, m)||^2 \), and the distance toward the rational matrix goes to zero either when \( \alpha \to 0 \) or \( \alpha \to 1 \) for all \( m_1, m_2 \in [0, 1] \), or when \( m_1 = m_2 \). The importance of the distortion becomes increasingly higher when \( |m_1 - m_2| \to 1 \) and \( \alpha \) gets close to \( \frac{1}{2} \). That is, the lack of attention matters more if the consumer cares about consuming both goods and as long as the consumer’s relative reaction to the two prices is significant.

\(^5\) The relative contributions to the Slutsky matrix norm \( ||E_j|| ||E|| \) for \( j \in \{\sigma, \pi, \nu\} \) have cardinal meaning. The prediction errors in demand changes due to Slutsky compensated prices attributed to both anomalies are of the same magnitude.\(^{12}\)
3 The Matrix-Nearness Problem: Its Solution

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

3.1 Preliminaries on Matrices

We begin by reviewing some definitions.

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 \( S^\sigma = S_{sym} = \frac{1}{2}[S + S'] \) is its symmetric part and \( E^\sigma = S_{skew} = \frac{1}{2}[S - S'] \) is its antisymmetric or skew-symmetric part (the orthogonal complement of the symmetric part).

Every symmetric matrix function \( S^\sigma \in \mathcal{M}(Z) \), and in particular the symmetric part of a Slutsky matrix function \( S \), can be decomposed into the sum of a singular and nonsingular part (with prices in the null space): see Claim 2 below. With our notation, the matrix function part that is singular with \( p \) in its null space will be denoted as \( S_{\sigma,\pi} \), that is, \( S_{\sigma,\pi}(z)p = 0 \). Then, we can write \( S^\sigma = S_{\sigma,\pi} + E^\pi \), where \( S_{\sigma,\pi} = PS^\sigma P \) and \( E^\pi = S^\sigma - PS^\sigma P \) is its orthogonal complement, with \( P = I - \frac{pp'}{pp} \) a projection matrix.\(^6\)

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

In general, a square matrix function may not admit diagonalization. However, thanks to Kadison (1984) we know that every symmetric matrix-valued function in the set \( \mathcal{M}(Z) \) is diagonalizable. In particular, \( S_{\sigma,\pi} \) can be diagonalized: \( S_{\sigma,\pi}(z) = Q(z)\Lambda(z)Q(z)' \). Here, the diagonal matrix of eigenvalues is \( \Lambda(z) = Diag([\lambda_l(z)])_{l=1,...,L} \), where \( \Lambda(z) \in \mathcal{M}(Z) \),

\(^6\)Because of Walras' law, when \( S \) is the Slutsky matrix of some demand function \( x \in \mathcal{X}(Z) \), then \( p'S(z) = 0 \) and \( S_{\sigma,\pi} = S^\sigma - E^\pi \) and \( E^\pi(z) = \frac{1}{pp'}[S^\sigma(z)pp' + pp'S^\sigma(z)] \).
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.\(^7\) Let \( Q(z) = [q_1(z) \ldots q_L(z)] \), where \( Q \in \mathcal{M}(Z) \) and its columns \( q_l \in C^1(Z) \) are the eigenvector functions such that for \( l = 1, \ldots, L \): \( S^{\sigma,\pi}(z)q_l(z) = \lambda_l(z)q_l(z) \) for \( z \in Z \).

Any real-valued function can be written as \( \lambda(z) = \lambda(z)_+ + \lambda(z)_- \), with \( \lambda(z)_+ = \max\{\lambda(z), 0\} \) and \( \lambda(z)_- = \min\{\lambda(z), 0\} \). This decomposition allows us to write \( S^{\sigma,\pi,\nu}(z) = S^{\sigma,\pi}(z)_- = Q(z)\Lambda(z)_-Q(z)' \) for \( \Lambda(z)_- = \text{Diag}\{\lambda_l(z)_-\}_{l=1,\ldots,L} \), with \( \lambda_l(z)_- \) the negative part of the \( \lambda_l(z) \) function. We can also write \( E^{\nu}(z) = S^{\sigma,\pi}(z)_+ = S^{\sigma,\pi}(z) - S^{\sigma,\pi}(z)_- \) for \( z \in Z \), or \( S^{\sigma,\pi}(z)_+ = Q(z)\Lambda(z)_+Q(z)' \) with \( \Lambda(z)_+ \) defined analogously to \( \Lambda(z)_- \). This is the orthogonal complement of \( S^{\sigma,\pi,\nu} \).

### 3.2 The Result

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

**Theorem 1.** Given a Slutsky matrix \( S \), it has the unique orthogonal decomposition \( S = S^s + E \), where \( S^s = S^{\sigma,\pi,\nu} \) is the solution to the matrix-nearness problem and \( E \) is the sum of the orthogonal complements of the symmetric, \( p \)-singular, and NSD parts of \( S \): \( E = E^s + E^\pi + E^\nu \). Furthermore, the Slutsky norm of the solution to the matrix-nearness problem can be decomposed as follows:

\[
\|E\|^2 = \|E^s\|^2 + \|E^\pi\|^2 + \|E^\nu\|^2.
\]

In the next section, we elaborate at length on the different components of this solution, right after outlining the proof of the theorem.

**Proof.** We first establish that the matrix nearness problem has a solution, and that it is unique. We do this in Claim 1.

**Claim 1.** The solution to the matrix nearness problem exists, and it is unique.\(^7\) The order of the eigenvalues is not essential to the results, but it is convenient for the proofs. Note that we can have many diagonal decompositions, all of which will work for our purposes.
The rest of the proof of the theorem is done in two parts. Lemma 1 gives the specific solution that imposes only the $p$-singularity and symmetry restrictions.

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

If $S^{\sigma,\pi}$ is NSD, then we would be done, since $S^{\sigma,\pi}$ would have property $\mathfrak{R}$ and minimize $||E||^2$. Otherwise, the general solution is provided in the lemma below, after we establish the pointwise orthogonality of $E^\pi$ and $E^\nu$:

**Claim 2.** The matrix $E^\pi(z)$ is pointwise orthogonal to $E^\nu(z)$. That is, $Tr(E^\pi(z)'E^\nu(z)) = 0$ for all $z \in Z$.

Using this claim, one can show next:

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

Summing up, the solution to the matrix-nearness problem is the result of applying three projections to the observed Slutsky matrix function $S$. First, we project it to the space of symmetric matrix functions and obtain $\rho_{\sigma}(S) = S^\sigma = \frac{1}{2}[S + S']$. Second, we project the output of the first projection, $S^\sigma$, to the space of matrix functions that are symmetric and singular in prices, and obtain $\rho_{\sigma,\pi}(S^\sigma) = S^{\sigma,\pi} = PS^\sigma P$. Finally, we project the output of the previous step, $S^{\sigma,\pi}$, to the space of matrix functions that are symmetric, singular in prices, and NSD, and obtain $\rho_{\sigma,\pi,\nu}(S^{\sigma,\pi}) = S^\nu = S^{\sigma,\pi,\nu} = (PS^\sigma P)_-$ the NSD part of $PS^\sigma P$. Each residual matrix is the difference between the original matrix at each step and its projection.

The orthogonalities of the relevant subspaces of matrix functions, as made evident in the steps of the proof, are responsible for the additive decomposition of the norm in the second part of the statement. The Frobenius norm and its weighted variants are essentially the only ones that correspond to an inner product on the space of matrix-valued functions.
functions. For this reason, they are the only ones that lead to the additive decomposition of the three components of the residual matrix $E$.\footnote{Higham (1988) previously found the solution of the nearest symmetric matrix to any square matrix, while Higham (1986) found the nearest NSD matrix to any square matrix. Our main Theorem generalizes these results finding first the nearest matrix that is simultaneously symmetric, singular in prices, and NSD. More importantly, to the best of our knowledge, the orthogonality and additive decomposability of the three properties $(\sigma, \pi, \nu)$, found in our main Theorem, are new to the matrix nearness literature.}

Finally, continuity of solutions is a desirable property. We close this section by stating that the solution $S^r$ to our problem is a continuous matrix-valued function:

**Claim 3.** $S^r$ is continuous.

The proof of this claim is omitted because $S^r$ 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; $S^r$ is therefore continuous by construction in all $z \in Z$.

## 4 Decomposition of the Matrix-Nearness Solution

The importance of Theorem (1) is that it provides a precise quantification of the size of the departures from rationality by a given behavior, as well as a revealing decomposition thereof.\footnote{The decomposition, a consequence of orthogonalities, can be seen as a generalized Pythagorean theorem.} Our decomposition is the unique orthogonal decomposition of $||E||^2$. The importance of the orthogonality of each of the error terms lies in its capacity to isolate the contributions of each of the Slutsky regularity conditions $\sigma$, $\pi$, and $\nu$ to the total bounded rationality error $||E||^2$. In particular, no other algorithm renders an orthogonal decomposition.

We should think of the three terms in the decomposition of $||E||^2$ as (i) the size of the violation of symmetry, (ii) the size of the violation of singularity, and (iii) the size of the violation of negative semidefiniteness of a given Slutsky matrix, respectively. This idea is illustrated in figure 1. The three terms are the antisymmetric part of the Slutsky matrix function, the correcting matrix function needed to make the symmetric part of the Slutsky matrix function $p$-singular, and the PSD part of the resulting corrected matrix function.
Figure 1: Decomposition of the Slutsky matrix norm. The size of the violation of rationality is the sum of the violations of symmetry, singularity in prices, and negative semidefiniteness.

Note that if one is considering a rational consumer, the three terms are zero. Indeed, following the steps in the proof, if $S(z)$ satisfies property $\mathcal{R}$ for all $z \in Z$, $S(z) = S^\sigma(z)$ and $E^\sigma(z) = 0$; $S^{\sigma,\pi}(z) = S^\sigma(z)$ and $E^\pi(z) = 0$; and $S^\sigma(z) = S^{\sigma,\pi,\nu}(z) = S^{\sigma,\pi}(z)$ and $E^\nu(z) = 0$. If exactly two out of the three terms are zero, the nonzero term allows us to roughly quantify violations of the Ville axiom of revealed preference (VARP) with $\|E^\sigma\|$, violations of homogeneity of degree 0 with $\|E^\pi\|$, and violations of the weak axiom of revealed preference (WARP) with $\|E^\nu\|$, respectively. These three kinds of violations correspond respectively to the ICPP, MI, and VCLD anomalies. We elaborate on these connections with the axioms of consumer theory in subsection 4.1.

4.1 Connecting with Axioms of Consumer Theory

The goal of this section is to remind the reader of the connections between the mathematical properties of the Slutsky matrix and classic axioms of consumer theory, and to emphasize the meaning of the decomposition in Theorem 1. This will further highlight the economic significance of our results.

Since Hurwicz and Uzawa (1971), we have known that for the class of continuously differentiable functions, a demand that satisfies Walras’ law can be rationalized if and only if its Slutsky matrix function is symmetric ($\sigma$) and NSD ($\nu$). A corollary of this
result is that its Slutsky matrix function is singular in prices ($\pi$) (John, 1995). The regularity conditions ($\sigma, \pi, \nu$) are linked to behavioral demand axioms. For completeness, it is useful to posit the axioms that we employ.

The first basic condition, which we assume throughout, is given by Walras’ law.

**Axiom 1.** *(Walras’ law)* \[ p'x(p, w) = w. \]

If \( x \in \mathcal{X}(Z) \) satisfies Walras’ law, then its Slutsky matrix function \( S \in \mathcal{M}(Z) \) satisfies that \( p'S(z) = 0 \) for \( z \in Z \).

Singularity in prices is given by HD0:

**Axiom 2.** *(Homogeneity of degree zero, HD0)* \[ x(\alpha z) = x(z) \text{ for all } z \in Z \text{ and } \alpha > 0. \]

If \( x \in \mathcal{X}(Z) \) satisfies Walras’ law and HD0, then \( S(z)p = 0 \) for \( z \in Z \). More generally, \( p \) is an eigenvector associated with the null eigenvalue of \( S \) if and only if \( x \) satisfies Walras’ law and HD0. The violation of HD0 results in the money illusion anomaly, MI.

The symmetry of the Slutsky matrix function is given by VARP.\(^{10}\) To state this axiom, we need to define an income path as \( w : [0, b] \rightarrow W \) and a price path as \( p : [0, b] \rightarrow P \).

Let \((w(t), p(t))\) be a piecewise continuously differentiable path in \( Z \). Jerison and Jerison (1992) define a rising real income situation whenever \((\frac{\partial w}{\partial t}(t), \frac{\partial p}{\partial t}(t))\) exist, with \( \frac{\partial w}{\partial t}(t) > \frac{\partial p}{\partial t}(t)'x(p(t), w(t)) \). A Ville cycle is a path such that: (i) \((w(0), p(0)) = (w(b), p(b))\); and (ii) \( \frac{\partial w}{\partial t}(t) > \frac{\partial p}{\partial t}(t)'x(p(t), w(t)) \) for \( t \in [0, b] \). In other words, in a Ville cycle, the demand is the same at the end of the path \( x(0) = x(b) \), even when all along the path the consumer is facing an increase in her purchasing power. This can be interpreted as an (effective) inattention to changes in the purchasing power. This terminology does not preclude cases where the source of the Ville cycle is different from a lack of attention.

**Axiom 3.** *(Ville axiom of revealed preference, VARP/Attentiveness to Changes in Purchasing Power: There are no Ville cycles.)*

Over a path along which the consumer’s purchasing power or real wealth increases, she should not make the same choice at the beginning and the end of the path. Hence, its

\(^{10}\)We present the axiomatization due to Ville as reinterpreted by Hurwicz and Richter (1979) and Jerison and Jerison (1992).
violation results in anomaly of inattentiveness to changes in purchasing power anomaly.

ICPP. Hurwicz and Richter (1979) proved that \( x \in \mathcal{X}(Z) \) satisfies VARP if and only if \( S(z) \) is symmetric and \( x \) satisfies Walras’ law.\(^{11}\)

The negative semidefiniteness condition of the Slutsky matrix is the differential expression analogous to the Compensated Law of Demand (CLD).

**Axiom 4.** (Compensated Law of Demand, CLD) \( x \in \mathcal{X}(Z) \) satisfies the CLD if for every \((p, w)\) and \((\overline{p}, \overline{w})\) such that \( \overline{w} = \overline{p}'x(p, w) \), it follows that \( (p - \overline{p})'(x(p, w) - x(\overline{p}, \overline{w})) \leq 0 \).

That is, when the consumer is compensated in the Slutsky sense, prices and demanded quantities “move in opposite directions.” The CLD is equivalent to the axiom below. Hence, its failure results in the violations of the compensated law of demand anomaly -VCLD-.

**Axiom 5.** (Wald axiom) \( x \in \mathcal{X}(Z) \) is such that for every \( w \in W \) and for all \( p \) and \( \overline{p} \),
\[
\overline{p}'x^w(p) \leq w \implies p'x^w(\overline{p}) \geq w.
\]

The singularity in prices and NSD conditions together are equivalent to the following version of WARP:

**Axiom 6.** (Weak axiom of revealed preference, WARP) \( x \in \mathcal{X}(Z) \) is such that for any \( z = (p, w), \overline{z} = (\overline{p}, \overline{w}) \): \( \overline{p}'x(p, w) \leq \overline{w} \implies p'x(\overline{p}, \overline{w}) \geq w \).

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

Finally, SARP is both necessary and sufficient for a Slutsky matrix function to have all three properties \( \sigma, \pi, \nu \):

---

\(^{11}\)A violation of the VARP is related to the possibility of intransitive preferences. When the consumer has preferences that are not transitive it is not possible to have a Slutsky matrix that has the symmetry property (Richter, 1996).
Axiom 7. (Strong axiom of revealed preference, SARP) $x \in \mathcal{X}(Z)$ is such that for any sequence $(p^1, w^1), (p^2, w^2), \cdots, (p^N, w^N)$, with $x(p^{n+1}, w^{n+1}) \neq x(p^n, w^n)$ for all $n \leq N - 1$, we have $p^N x(p^1, w^1) > w^N$ whenever $p^n x(p^{n+1}, w^{n+1}) \leq w^n$ for all $n \leq N - 1$.

In words, if $x(p^1, w^1)$ is revealed preferred to $x(p^N, w^N)$, then $x(p^N, w^N)$ cannot be directly revealed preferred to $x(p^1, w^1)$. Houthakker (1950) proved that if the demand function $x \in \mathcal{X}(Z)$ satisfies SARP if and only if it can be generated by maximizing a utility function subject to a linear budget constraint.\textsuperscript{12} A demand function satisfies VARP and WARP if and only if it satisfies SARP (Hurwicz & Richter, 1979; Hurwicz & Uzawa, 1971). If SARP fails (given the premise $x(p^N, w^N)$ is directly revealed preferred to $x(p^1, w^1)$ and WARP holds, the length of the cycles $N \geq 2$ has been connected with the degree of asymmetry of the Slutsky matrix or the intensity of the ICPP anomaly. The longer the cycle the higher the asymmetry of the Slutsky matrix (Shafer, 1977; Jerison & Jerison, 2012). We focus less on SARP than on the combination of WARP and VARP, as our main objective is to understand the decomposition of the Slutsky matrix norm. We provide an extended discussion of the relation with SARP and WARP in the Supplement to this paper.

We are now ready to establish the main result of this subsection.

**Proposition 1.** The Slutsky matrix-nearness norm $||E||^2$ will be equal to zero if and only if $x$ satisfies VARP, HD0, and the Wald Axiom (or VARP + WARP/ SARP).

Moreover

(i) If $||E^x|| > 0$, then VARP fails: the ICPP anomaly.

(ii) If $||E^\pi|| > 0$ then HD0 fails: the MI anomaly.\textsuperscript{13}

(iii) If $||E^y|| > 0$ or $||E^z|| > 0$, then WARP fails: the VCLD anomaly.

(iv) If $||E^y|| > 0$ and $||E^z|| = 0$, then the Wald axiom fails: the VCLD anomaly.

\textsuperscript{12}GARP and SARP are equivalent in our framework because demand correspondences are ruled out.

\textsuperscript{13}With Walras’ law, if HD0 fails, $||E^\pi|| \neq 0$, but it will not be the only nonzero term in the decomposition.
5 Weighted Slutsky Norms

The norm of bounded rationality that we have built so far is an absolute measure. For a specific consumer, this distance quantifies how far that individual’s behavior is from being rational. Furthermore, we can also 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, in this exercise, the setting of the decision making process is fixed in the sense that the decision problem facing each of the individuals is presented in the same way. This implies that the measure is unit-dependent, being stated in the same units (those in which the consumption goods are expressed). These same units are also used for the terms in the decomposition. Also, the Frobenius norm depends on the area of the arbitrary compact region $Z$. In this section we address these issues by proposing weighted versions of our norm.

5.1 Generalized Weighted Slutsky Norms

Let $W$ be a nonzero square and PSD matrix function of weights, $W(z) \in \mathbb{R}^{L \times L}$. This matrix is meant to encode suitable normalizations and priorities of the modeller. Let $Z \subseteq P \times W$ be sigma-measurable and have positive measure $\mu$ with a continuous, strictly positive function $\mu$, such that $S \in \mathcal{M}(Z)$ is square integrable. The weighted seminorm $\|S\|_{W,\mu}$ is defined as follows:

$$\|S\|_{W,\mu}^2 = \int_{z \in Z} \|W(z)S(z)W(z)\|_{M\mu(z)}^2 dz < \infty,$$

with $\mu$ continuous and strictly positive. Letting $\mathcal{B}(Z)$ denote the Borel set on $Z$, the space $L^2(Z, \mathcal{B}(Z); \mathcal{M}(Z))$ and its weighted counterpart $L^2(Z, \mathcal{B}(Z); \mathcal{M}(Z); \mu)$ are unitary-equivalent or unitary-isomorphic. This means that our orthogonality relations under the inner products defined by the Lebesgue measure are preserved by inner products weighted by $\mu$. Hence, suitable variants of Theorem 1 can be obtained for these weighted norms.

Importantly, observe that properties $\sigma$ and $\nu$ are unit-invariant, but property $\pi$ depends on prices and will change with a normalization choice. To solve this, we consider
a modification of property $\pi$. We say that matrix function $A(z)$ has property $\pi^*$ if and only if $A(z)$ is singular in the vector $p^* = W^{-1}(z)p$ (i.e., $A(z)p^* = 0$, and $p^*A(z) = 0$) for all $z \in Z$. Let us consider matrix function $S^*(z) = W(z)S(z)W(z)$, and observe that $S^*(z)$ has property $\pi^*$ if and only if $S(z)$ has property $\pi$. We say that a matrix function $S^*$ has property $\mathcal{R}^*$ when it has properties $\sigma, \pi^*, \nu$. It follows that a matrix function $S$ has property $\mathcal{R}$ if and only if $S^*$ has property $\mathcal{R}^*$. Indeed, we can state:

Corollary 1. The solution to the matrix-nearness problem, $S^{**} = \arg\min_{A \in \mathcal{M}(Z)} \|S - A\|_{W,\mu}$ subject to $A^* = WAW$ (the weighted version of $A$) having property $\mathcal{R}^*$, is $S^{**} = W^{-1}(WS^{*,\pi}W)_{-1}W^{-1}$, where $S^{*,\pi} = W^{-1}(P^*(WS^{*,\pi}W)P^*)W^{-1}$, $S^{*,\sigma} = S^\sigma$, $P^* = [I - \frac{p^*p^*}{p^*p^*}]$, and $p^* = W^{-1}p$. The quadratic norm decomposition is given by:

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

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

Observe that both the part concerning the NSD matrix function of violations (property $\nu$), and the part concerning the singularity of prices (property $\pi$) are affected, due to the introduction of the weighting matrix function $W$.

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

Remark 2. If $W = \text{Diag}(p)$, the weighted Slutsky matrix norm is expressed in (square) dollar amounts. The entries of this normalized matrix express changes in expenditure on each good, following a percent change in each price.

Remark 3. If $W = \frac{1}{\sqrt{\mu}}\text{Diag}(p)$, the weighted Slutsky matrix is unit-free. The entries of this normalized matrix express changes in expenditure shares on each good, following a percent change in each price.

Remark 4. When we have a degenerate $\mu$ such that $\mu(\tau) = 1$ (say, an equilibrium price and a given exogenous wealth) and $W$ is the identity, then $\|S\|_{\mu} = \|S(\tau)\|_{\mu}$ and we can

\[\text{In particular, note that } S^{**} \text{ has property } \pi \text{ because } S^{**,p} = W^{-1}(P^*(WS^{*,\pi}W)P^*)_{-1}p^* = 0, \text{ which holds by definition of the projection matrix } P^*; \text{ the same holds for } p'S^{**,r} = 0.\]
apply the local results of Jerison and Jerison (1992) for the $\sigma$ property (the only case they considered).

The standard Frobenius norm we use is chosen when one gives equal weight to each entry of the Slutsky matrix, thus preventing a specific price change on a specific commodity from receiving greater weight than any other. However, the modeller may adopt such choices at will and use arbitrary weighting matrices as a function of her interests. Finally, the integral in the norm embodies an expectation operator, which can be justified with Savage-type (1954) ideas of independence across consumption data points. In fact, we developed an axiomatic foundation of our norm, along these lines, in our working paper version.

5.2 Application to the Sparse-Max Example

We return to the example in subsection 2.4. We can easily modify our results there to obtain a unit-independent Slutsky matrix norm. We consider the norm $\|\cdot\|_{M, \Lambda_{p,w}}$, defined as $\|S\|_{M, \Lambda_{p,w}} = \|\Lambda_{p,w} S \Lambda_{p,w}\|_M$ with $\Lambda_{p,w} = \frac{1}{\sqrt{w}} \text{diag}(p)$. Whether goods are measured in pounds or kilos, this yields a unit-free norm, and it is expressed in terms of budget shares and elasticities. Note that $(\Lambda_{p,w} S(p, w) \Lambda_{p,w})_{ij} = s_{ij}(p, w) \frac{p_i p_j}{w} = \frac{\partial b_i(p, w)}{\partial \ln(p_j)} + \frac{\partial b_i(p, w)}{\partial \ln(w)} b_j(p, w)$ with $b_i(p, w) = p_i x_i(p, w) / w$ representing good $i$’s budget share and $\ln(\cdot)$ denoting the natural logarithm. Our modified unit-free Slutsky norm gives us:

$$\|E(p, w)\|^2_{M, \Lambda_{p,w}} = (m_1 - m_2)^2 (1 - \alpha)^2 \alpha^2.$$ 

The main intuitions about the original Slutsky norm remain true, but we now have a unit-free (percentage) index, which corresponds to errors in predictions of the budget shares and price elasticities. The dependence on prices is removed. Observe that $\Lambda_{p,w} S \Lambda_{p,w}$ has properties $\sigma, \pi^*, \text{ and } \nu$. In particular, it is singular in the unit vector $1_L = (1, \cdots, 1)' \in \mathbb{R}^L$ (i.e., $\Lambda_{p,w} S \Lambda_{p,w} 1_L = 0$ and $1_L' \Lambda_{p,w} S \Lambda_{p,w} 1_L = 0$).

A second issue is that our initial norm is an integral under a Lebesgue measure of the local Frobenius norm (i.e., for a given pair of prices and wealth). This measure may grow
proportionally with the area of $Z$. Again, consider the simple area $Z = \{w \in [\underline{w}, \overline{w}], p_1 = 1, p_2 = 1\}$. Then:

$$||E(p, w)||^2_{\Lambda_{p,w}} = \int_{z \in Z} ||E(p, w)||^2_{M,\Lambda_{p,w}} dz = (\overline{w} - \underline{w}) \cdot (m_1 - m_2)^2(1 - \alpha)^2\alpha^2.$$ 

The measure depends linearly on the area $(\overline{w} - \underline{w})$. We eliminate this pitfall by considering instead an expectation or average over the given area of prices and wealth, resulting in an area-independent index. We use the new normalized (and unit-free) modified Slutsky matrix norm $||E||^2_{\Lambda_{p,w,\mu}} = \int_{z \in Z} ||\Lambda_{p,w}E(z)\Lambda_{p,w}||^2_{M}\mu(z) dz$. For instance, assume that $\mu$ defined over $Z$ is a uniform distribution over $[\underline{w}, \overline{w}]$; we obtain then our unit-free and area-free Slutsky Matrix norm:

$$||E||^2_{\Lambda_{p,w,\mu}} = (m_1 - m_2)^2(1 - \alpha)^2\alpha^2.$$ 

This new index provides an average error in the considered area, which is closed-form, unit-free, area-free, and comparable among different consumers. For instance, letting $\alpha = \frac{1}{2}$, one can now visualize $||E(m_1, m_2)|| = \frac{1}{4}|m_1 - m_2|$ in the $m_1, m_2$ space (figure 2a). We can see that the 45-degree line is equivalent to rationality. This includes the full rationality case $m_1 = m_2 = 1$, as well as the “as if” rational behavior of $m_1 = m_2 < 1$. When $|m_1 - m_2| < 0.10$, the Slutsky matrix norm is close to 2% which can be considered very close to rationality. At the other extreme, when $|m_1 - m_2| = 1$, the Slutsky matrix norm is 25% in the given region $Z$, suggesting that even in that case the consumer does not distort her demand elasticities beyond that amount. If our index hypothetically reached 1 (if $||S||^2_{\Lambda_{p,w,\mu}} = 1$), this would mean that the residual matrix is equal to the behavioral Slutsky matrix, $E = S$. In this case, the budget share elasticities are distorted by up to 100% of their rational value, and the consumer’s behavior is the completely opposite of rationality, namely the compensated demand will have positive slope, be subject to money illusion, and also suffer from the ICPP anomaly.
6 Connecting with Demand: Model Misspecification

We have solved the matrix nearness problem on the basis of the Slutsky regularity conditions. Now, in order to connect back to demand, the exercise is one of integrating from the first-order derivatives of the Slutsky matrix terms. In such an integration step, a constant of integration shows up, which we interpret as a “specification error.” That is, starting from our observed Slutsky matrix function $S(x)$, and making use of Theorem 1 the nearest matrix function $S^r$ satisfying all the regularity properties. One difficulty, however, is that in general there may not exist a demand function $x^r$ generating $S^r$.

Now suppose that we specify $x_1$ and $x_2$ as two demand functions with which we wish to explain the consumer’s behavior. For example, $x_1$ might be a Cobb-Douglas demand properly estimated, whereas $x_2$ might be a different estimate, perhaps allowing the entire class of CES demands.
As we will show, the error in comparative statics analysis from using $x_1$ is $||S(x) - S(x_1)||^2 = ||S^r - S(x_1)||^2 + 2\langle E^r, S(x_1) - S^r \rangle + ||E||^2$, while the error from using $x_2$ is $||S(x) - S(x_2)||^2 = ||S^r - S(x_2)||^2 + 2\langle E^r, S(x_2) - S^r \rangle + ||E||^2$. This means that the behavioral error found in Theorem 1, namely, $||E||^2 = ||S(x) - S^r||^2$, does not change with $x_1$ or $x_2$. That is, the behavioral error is invariant to the choice of the parametrized rational approximation. The other term corresponds to the (rational) model specification error, as a function of the parametrized class considered. We establish this useful separation in our next result.

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

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

Our next result follows:

**Proposition 2.** For any solution $x^{*r} \in \arg\min_{x^r \in \mathbb{R}^r(Z)} ||S(x) - S(x^r)||^2$, it follows that $||S(x) - S(x^{*r})||^2 = ||S^r - S(x^{*r})||^2 + 2\langle E^r, S^r - S(x^{*r}) \rangle + ||E||^2$, where $0 \leq \langle E^r, S^r - S(x^{*r}) \rangle \leq ||S^r - S(x^{*r})|| \cdot ||E^r||$. Here, $||E||^2$ is the Slutsky matrix norm –the behavioral error term– and the other two terms are a generalized specification error, consisting of $||S^r - S(x^{*r})||^2$ –the proper specification error– and the last term is a nonnegative and bounded correlation term or penalty.

**Remark 5.** The correlation term shows up because the error capturing the failure of NSD is not orthogonal to the specification error, while the errors associated with the other failures of the Slutsky properties are, as the proof of the proposition reveals.

**Remark 6.** The non uniqueness of the solution to Proposition 2 is a consequence of the properties of the set of rational demands $\mathcal{R}(Z)$. The set $\mathcal{R}(Z)$ is not convex (Grodal, 1974). Also, the objective function $||S(x) - S(x^r)||^2$ need not be convex on $x^r \in \overline{\mathbb{R}}^r(Z)$. The reason is that $S(x + y)$ for $x, y \in \overline{\mathbb{R}}^r(Z)$ may not be equal to $S(x) + S(y)$, given the
nonlinear wealth effects in the Slutsky matrix function. This result is compatible with cases where two demand functions have the same Slutsky matrix function.

6.1 Returning Once Again to the Sparse-Max Consumer Example

We now return to the example in subsection 2.4. The example shows that there exists a rational demand function that is behaviorally closer to the sparse-max consumer demand proposed by Gabaix than to the “underlying rational” model of his framework.

To simplify the computations, we take $m_1 = 1$, $m_2 = 0$, which is the case of a consumer who is totally attentive to $p_1$ and totally inattentive to $p_2$, and perform a local analysis: fix $(p, w) = (p^d, w)$ as the reference point. First, using the same parameter $\alpha$, we compare the distance between $S^G$ and $S^{CD}$. If we use $x^{CD}$ as an approximation of $x^G$ at the default prices, we obtain the following local comparative statics error:

$$||S^G(p^d, w) - S^{CD}(p^d, w)||^2_M = \frac{(\beta^d_1)^2 + (\beta^d_2)^2}{(\beta^d_1)^2} \frac{(1-\alpha)^2\sigma^2}{(\beta^d_1)^2}.$$ By our result earlier in this section, we can decompose this quantity as the sum of a specification error and $||E(p^d, w)||^2_M$. In fact, it follows that $||S^{CD}(p^d, w) - S^G(p^d, w)||^2_M = ||S^{CD}(p^d, w) - S^r(p^d, w)||^2_M + ||E(p^d, w)||^2_M$, with the specification error $||S^{CD}(p^d, w) - S^r(p^d, w)||^2_M = \frac{(1-\alpha)^2\sigma^2}{(\beta^d_1)^2}$ and the behavioral error $||E(p^d, w)||^2_M = \frac{(1-\alpha)^2\sigma^2}{(\beta^d_1)^2}$.

We know also that, even if we could improve the specification error, the Slutsky matrix error norm $||E(p^d, w)||^2_M$ will not change at all. To illustrate this, we find the Cobb-Douglas demand that minimizes the total error quantified in Proposition 2 at the reference price-wealth pair $(p^d, w)$. We can write this problem parametrically, through characterizing the Cobb-Douglas family by a parameter $\beta \in [0, 1]$: $x^{CD2}(p, w, \beta) = \left(\frac{\beta w}{p_1} + \frac{(1-\beta)w}{p_2}\right)'$, where $\beta$ in general will differ from the given $\alpha$. We solve the problem $\beta^* \in \text{argmin}_{\beta \in [0, 1]} ||S^{CD2}(p^d, w, \beta) - S^G(p^d, w)||^2_M$. There are two solutions $\beta^* = \frac{1}{2} \pm \sqrt{\frac{\beta^d_2 \beta^d_1^2 + (\beta^d_2)^2}{2\beta^d_1 \beta^d_2} [1 - 2\alpha^2]}$, which depend on the given parameters and the fixed default price.

Because the solution to the matrix-nearness problem yields a matrix associated with a Cobb-Douglas utility function, the specification error is zero, that is, $||S^{CD2}(p^d, w, \beta^*) -$


\[ S^r(p^d, \overline{w})\|_M^2 = 0. \] Hence, in this case, the total error in comparative statics reduces to the behavioral error captured by the Slutsky norm: \[ ||S^{CD}2(p^d, \overline{w}, \beta) - S^D(p^d, \overline{w})||_M^2 = ||E(p^d, \overline{w})||_M^2, \]

where \[ ||E(p^d, \overline{w})||_M^2 = \frac{\nu^2 \alpha^2 (1 - \alpha)^2}{|p_d^1|^2 |p_d^2|^2}. \]

Observe that the error does not depend on \( \beta^* \) and remains unchanged with respect to the case of \( S^{CD} \). Of course, even if we used the best rational model to minimize errors in the comparative statics analysis, locally, we would not reduce at all the Slutsky matrix error norm that stems from lack of rationality, as captured by properties \( \sigma, \pi, \nu \). This means that generally a rational consumer cannot approximate well a sparse-max consumer when \( m \) is sufficiently close to zero.

7 Further Applications: Comparative Statics of Bounded Rationality

We offer four applications in this section. First, we continue to study the limited attention model of the sparse-max consumer in its general form. Second, we consider the multiple-selves (or household) decision making problem. Third, we study the mental accounting model. Finally, we analyze the quasi-hyperbolic discounting model. Here we focus on how the intensity of the anomalies varies with the bounded rationality parameters in each of the models. In the same vein, we want to know when we can say whether a consumer is more boundedly rational than another, within a given class.

There is a growing number of bounded rationality models, offering distinct explanations of deviations from rationality. The Slutsky matrix norm decomposition could be used as a unifying alternative, possibly a more agnostic approach, to bounded rationality based on quantifying departures from the axioms of revealed preference. We hope the methodology proposed herein could provide a guide for the (empirical) researcher on which properties she should mainly focus.\(^{15}\)

\(^{15}\)We thank an anonymous referee for suggesting this interpretation of our approach.
7.1 General Sparse-Max Consumer: Further Effects of Limited Attention on Bounded Rationality

We move to an $L$-good, non-parametric version of our motivating example. Recall that the source of bounded rationality for the sparse-max consumer model is the (possibly) misperceived price:

$$p^G = Mp + [I - M]p^d,$$

for true prices $p \in P$, default prices $p^d \in \mathbb{R}^L_{++}$, and $M \in \mathbb{R}^{L\times L}$ a diagonal matrix of attention $M = \text{diag}(\{m_l\}_{l=1}^L)$ for $m_l \in [0,1]$. Formally, the consumer solves the sparse-max problem $\text{smax}_{x|p^G} u(x)$ subject to $p'x = w$. The solution to this problem is the sparse-max demand system $x^G(p, w) = x^r(p^G, w')$, where the underlying rational demand is $x^r(p, w) = \text{argmax}_{x \in \mathcal{X}} u(x)$, subject to $p'x = w$ (and $w'$ is a correcting wealth that we defined in the motivating example).

Gabaix (2014) proves that, when prices are evaluated at the default price $p = p^d$, the sparse-max consumer Slutsky matrix is related to the underlying rational Slutsky matrix $S_{x^r}(p, w) = D_p x^r + D_w x^r x^{r'}$ with the following equation:

$$S^G(p, w) = S_{x^r}(p, w) M.$$

This relation is a multiplicative perturbation of a rational Slutsky matrix function $S_{x^r}(p, w)$. We are now interested in finding the effect of changes in the behavioral parameter $M$ on the different anomalies of behavior. Our analysis is done locally with respect to prices at the reference price $p = p^d$.

- The errors corresponding to the inattentiveness to changes in purchasing power (ICPP) are given by:

$$||E^\sigma(p, w)||_M^2 = \sum_{i,j} (s_{i,j}^r(p, w))^2 \frac{(m_j - m_i)^2}{2}.$$

Surprisingly, the ICPP actually depends not only on the absolute value of the attention parameter $m_i$, but mainly on its relative value across goods. The Slutsky matrix norm
help us conclude that when $|m_i - m_j| < \epsilon$ for $\epsilon > 0$ arbitrarily small for all goods $i, j \in \{1, \cdots, L\}$, the sparse-max consumer behaves locally as if she were a rational consumer. This implies that one necessary condition for the inattention parameters to meaningfully affect the ICPP anomaly is that they be heterogeneous across goods.

- The errors corresponding to money illusion (MI) are given by:

$$||E^\pi(p, w)||^2_{\text{map}} = \frac{1}{4(p'p)^2} \sum_{i,j} p_j \sum_{k=1}^{L} P_{ik}^x(p, w) p_k m_k + p_i \sum_{k=1}^{L} p_k S_{kj}^x(p, w) m_k]^2.$$ 

We know first that $\sum_{k=1}^{L} P_{ik}^x(p, w) p_k = 0$ because of HD0 of the underlying rational demand; we also know that $\sum_{k=1}^{L} p_k S_{kj}^x(p, w) = 0$ because of Walras’ law. If $|m_i - m_j| < \epsilon$ for $\epsilon > 0$ small enough for all goods $i, j$, then it is easy to see that $\sum_{k=1}^{L} P_{ik}^x(p, w) p_k m_k \simeq 0$ and $\sum_{k=1}^{L} p_k S_{kj}^x(p, w) m_k \simeq 0$. Thus, the relative magnitude of the attention parameters across goods once again is what matters.

- The errors corresponding to the violations of the compensated law of demand (the VCLD anomaly) related to the Slutsky matrix property $\nu$ behave in a fundamentally different way than those of the ICPP and MI anomalies. Even for cases where there is full heterogeneity of attention across goods $m_i \neq m_j$ for all goods $i, j$, this component of the decomposition may be zero. The error matrix is $E^\nu = \frac{1}{2}(P[S^{x^\nu} M + MS^{x^\nu}]P)_+$, and hence,

$$||E^\nu||^2 = \sum_i \lambda^2_i(S^{\sigma,\pi}),$$

where $\lambda_i(S^{\sigma,\pi}) = max\{0, \lambda_i(S^{\sigma,\pi})\}$ is the positive part of the eigen-values of $S^{\sigma,\pi}$.

- The error matrix $E^\nu$ is zero if and only if the following inequality holds for all price change vectors $v \in \mathbb{R}^L$:

$$v'S^{x^\nu} M v \leq v'S^{x^\nu} M p \frac{(v'p)}{p'p).$$
The reader ought to think of \( v \) as a small price change. Notice that the LHS of the previous inequality is equal to the inner product of price changes and sparse-demand changes: \( v' S^x M v = dp'dx^s \). We know that if \( S^G = S^x M \) is NSD, then the compensated law of demand holds: \( dp'dx^s \leq 0 \). This also implies that HD0 holds, so \( S^G(p, w)p = S^x M p = 0 \). These two facts, taken together, imply that both terms \( E^x \) and \( E^v \) are zero.

If \( S^G = S^x M \) is not NSD or if there is a VCLD anomaly for some \( v \in \mathbb{R}^L \), then \( v' S^x M v \geq 0 \). If \( v' S^x M p(\frac{v_p}{p_p}) = v' S^x M v_p \geq 0 \) is also nonnegative, where \( v_p = p(\frac{v_p}{p_p}) \) is the projection of the (small) price change \( v \) on the (positive) price vector \( p \), then \( v' S^\pi_x v \leq v' S^G v \); that is, the symmetric and singular approximation is closer to satisfying NSD than the behavioral Slutsky matrix. In fact, for some parametric specifications like the sparse-max Cobb-Douglas motivating example, this correcting effect is so strong that we have \( E^v = 0 \). In general, we conclude that when we correct for MI by imposing \( \pi \), in the sparse-max model we may also be correcting for the anomaly of VCLD at least in some direction of price changes. This is an analogous observation to the motivating example, where the CLD is not violated massively by the sparse-max consumer.

- Again, we see that if \( |m_i - m_j| < \epsilon \) for a small \( \epsilon > 0 \), then \( M p \simeq m p \) for some constant \( m \); this implies that \( v' S^x M p(\frac{v_p}{p_p}) \simeq 0 \) for all price changes \( v \in \mathbb{R}^L \). Then we know that \( E^v \) is arbitrarily close to zero when the attention parameters are homogeneous among the different goods.

We summarize our observations in the following remark:

**Remark 7.** Let us consider two sparse-max consumers \( x^{G_i}, i = 1, 2 \), such that \( x^{G_i}(p, w) = x^x(\{m^i_l p_l + (1-m^i_l) p'_l\}_{l=1}^L, w') \), i.e., they only differ in the attention matrix \( M^i \neq M^j \). We say that \( x^{G_i} \) is more boundedly rational than \( x^{G_j} \) if and only if \( ||E^{G_i}|| \geq ||E^{G_j}|| \). Then we have established that \( x^{G_i} \) is more boundedly rational than \( x^{G_j} \) when \( |m^i_l - m^i_k| \geq |m^j_l - m^j_k| \) for all \( l, k \in \{1, \cdots, L\} \), i.e., when the relative inattention across goods is uniformly greater for \( i \) in comparison with \( j \).

---

16 The reason is that the symmetric and singular part matrix, multiplied by the price changes, is such that \( v' S^\pi_x v = \frac{1}{2} [v' S^x M p v + v' P M S^x v] = v' S^x M v - v' S^x M p(\frac{v_p}{p_p}) \), which we now know is closer to having the NSD property than the original behavioral Slutsky matrix \( S^G \) by the subtraction of the nonnegative term \( v' S^x M p v \geq 0 \).
7.2 The Multiple-Selves Consumer Problem: The Effects of Preference Aggregation on Bounded Rationality

We consider a demand function that is generated by a consumer unit composed of two members (dual-selves). The most well-known version of this model was first introduced by Browning and Chiappori (1998) in order to explain a household’s decision making. Here we are interested in the multiple-selves interpretation, the same individual consumer has different utilities, and the individual attempts to solve her consumer maximization as a collective problem. This framework relaxes the implicit assumption, under the individual rationality framework, of a single self. In the multiple-selves model, an individual consumer bounded rationality arises from the (potential) inability to hold a unitary criterion to evaluate an alternative. In order to make decisions, the consumer must aggregate the multiple selves in some way. The multiple-selves or multiple-rationales framework, as a way to model bounded rationality, has attracted the attention of decision theorists (Ambrus & Rozen, 2015), behavioral economists (Kalai et al., 2002; Köszegi & Sztchi, 2012), and psychologists (Lachmann, 1996). Here, we consider the case where the consumer unit tries to maximize the aggregate welfare:

\[
\max_{q^A, q^B, q^H} \mu(p, w)u^A(q^A, q^H) + (1 - \mu(p, w))u^B(q^B, q^H)
\]

\[\text{s.t. } p'[q^A + q^B + q^H] = w,\]

where \(u^j\) for \(j \in \{A, B\}\) is the utility of member \(j\), and \(q^j\) is the consumption level of \(j = \{A, B, H\}\) where \(H\) denotes goods that are “enjoyed” collectively. The smooth function \(\mu(p, w) \geq 0\) is a Pareto weight that is usually assumed to be HD0. We assume that the objective function is smooth and that the solution to the problem is unique. The collective demand of the consumer unit is:

\[
x^H(p, w) = q^A(p, w) + q^B(p, w) + q^H(p, w),
\]
where $(q^A(p, w), q^B(p, w), q^H(p, w))'$ is the solution to the consumer unit problem. Note that $x^H(p, w) = f(p, w, \mu(p, w)) \in \mathbb{R}^L$ is a function $f$ of the Pareto weight. The collective demand $x^H$ satisfies HD0 and Walras’ law. This implies that $E^\pi = 0$. We explore now the rest of the Slutsky matrix errors, for $\sigma$ and $\nu$.

Browning and Chiappori (1998) prove that the Slutsky matrix of the consumer unit is equal to the sum of a matrix function $S^{H^r}$ that has properties $\sigma, \nu$, and $\pi$ and an additive perturbation matrix function $U$:

$$S^H = S^{H^r} + U,$$

where $U = vu'$ with $u_i = \frac{\partial f_i}{\partial \mu}$ for the $f$ such that $x^H(p, w) = f(p, w, \mu(p, w))$ and $v_j = \frac{\partial \mu}{\partial p_j} + \frac{\partial \mu}{\partial w} x^C_j$ for $i = \{1, \cdots, L\}$.

First, note that due to the HD0 assumption on the Pareto weight, there is no MI anomaly and $S^H$ has property $\pi$. The multiple-selves consumer is not fooled by money illusion.

- The errors corresponding to the inattentiveness to changes in purchasing power (ICPP) are given by:

$$||E^\sigma|| = \frac{1}{4} \sum_{i,j} (u_j v_i - u_i v_j)^2,$$

which depends entirely on the level of asymmetry of the perturbation $U$, a function of the Pareto weight. If $\mu(p, w) = 0$, then we have $U = 0$ and thus $E^\sigma = 0$. Moreover, if $|u_j v_i - u_i v_j| = 0$, then $E^\sigma = 0$. This means that there may be cases with a nonzero Pareto weight function that still have property $\sigma$. It is easy to understand why this situation could happen, even if some consumption unit members react in one direction to an increase in the purchasing power, the other member may react in the opposite direction, offsetting the effect on the collective demand. If the term $|u_j v_i - u_i v_j|$ grows, then the ICPP anomaly is found.\(^{18}\)

\(^{17}\)Since $S^{H^r}$ has property $\pi$ (by John (1995)), this implies that $U$ also has property $\pi$. The closest rational Slutsky matrix is $S^r = [S^{H^r} + \frac{1}{2} [U + U']]'$.

\(^{18}\) The important case of a constant Pareto weight $\mu(p, w) = \mu$ implies that $E_{ij}^\sigma = 0$. In the case where $\frac{\partial f_i}{\partial \mu} = \kappa$ for all $j$ (i.e., when the effect of the Pareto weight on the collective demand is homogeneous
• The errors corresponding to the VCLD are given by $E^\nu = \sum_i \lambda^2_i(S^\sigma)$ or the positive part of the eigenvalues of $S^\sigma = [S^{H^\nu} + \frac{1}{2}[U + U']]$, such that $\lambda^+_{i}(S^\sigma) = \max\{0, \lambda_i(S^\sigma)\}$. (Note that $S^\sigma = S^\sigma^{\pi}$ in this case.) Again, the intensity of the VCLD anomaly is given by $U$; in this case, by the term $\frac{1}{2}[U + U']$. Notice that, if we consider small price changes $r \in \mathbb{R}^L$, the inner product of the price changes and consumption unit changes $dp'dx^H = r'S^{H^\nu}r$ can be written as:

$$r'S^{H^\nu}r = r'S^{H^\nu}r + r'Ur.$$  

Thus, if we observe a VCLD anomaly $dp'dx^H \geq 0$, we must have $r'Ur \geq 0$. More importantly, the intensity of this violation is given only by $U$, since having $r'S^{\sigma}r \geq 0$ means that $r'Ur \geq 0$ is necessarily nonnegative. Again, if $\mu(p,w) = 0$, then $U = 0$ and $E^\nu = 0$. The simple case of a constant Pareto weight implies that $r'vu'r = 0$ for all $r \in \mathbb{R}^L$, which in turn implies that $E^\nu = 0$.

**Remark 8.** A general conclusion in the case of the collective demand $x^H$ is that if the Pareto weights compensated cross-price effects are homogeneous across goods, then for practical modelling purposes, the multiple-selves consumer will behave similarly to a rational consumer. Formally, if the Pareto weight cross price effects term is small enough $|u_jv_i - u_iv_j| \leq \epsilon$ for $\epsilon > 0$ for all goods, then the collective demand $x^H$ appears rational. This underlines again the importance of the relative effects of the bounded rationality parameter $\mu$, resembling the sparse-max example (where we require the attention parameter to be heterogeneous across goods for bounded rationality to matter). In general, the VCLD anomaly is controlled by the term $r'Ur \geq 0$. This roughly means that there is a positive correlation between the collective-demand elasticity of the Pareto weight and the (compensated) price elasticity of the Pareto weight. If both go in the same direction, then the VCLD is more important, and so is $E^\nu$. The intuition behind this result is that the compensated law of demand is violated as a result of a change in relative power (i.e., a change in $\mu(p, w)$) that depends on prices that makes the quantities increase as across goods), then the ICPP anomaly is controlled by the difference between the Slutsky-compensated price changes of the Pareto weight: $\left[\frac{\partial \mu_i}{\partial p_i} + \frac{\partial \mu_i}{\partial w_i}x_i^C\right] - \left[\frac{\partial \mu_j}{\partial p_j} + \frac{\partial \mu_j}{\partial w_j}x_j^C\right]$. If this term is homogeneous across goods as well, $E^\sigma$ will be close to zero.

19By construction, we know that $r'S^{H^\nu}r \leq 0$.  

34
prices rise due to the higher relative importance of one of the consumer unit members.
On the other hand, the ICPP anomaly depends entirely on the level of asymmetry of the
perturbation $U$, a function of the Pareto weight.

## 7.3 Mental Accounting

The Mental Accounting model (Thaler, 1985; Hastings & Shapiro, 2013) variant that
we consider here was formulated by Farhi & Gabaix (2015). In this case, we assume
the consumer has $K \geq 1$ fixed rigid mental accounts, indexed $k = 1, \cdots, K$, with men-
tal account function $\omega^k(p, w)$. We assume that $\omega^k(\cdot)$ is homogeneous of degree 1 (i.e.,
$\omega^k(\alpha p, \alpha w) = \alpha \omega^k(p, w)$ for all $\alpha > 0$). Also we define the vector function of all men-
tal accounts $\omega(p, w) = (\omega^1, \cdots, \omega^K)$. The mental accounting function must satisfy that
$\sum_{k=1}^K \omega^k(p, w) = w$. The consumer still tries to maximize a utility function $u$ defined on
$\mathbb{R}^L$ subject to $p^k x^k(p, w) = \omega^k(p, w)$ where $x^k$ is the vector of commodities associated
with account $k$ such that $x = (x^1, \cdots, x^K)' \in \mathbb{R}^L$. Formally, the boundedly rational consumer solves:

$$x(p, w) = x^r(p, \omega(p, w)) = \arg \max_q u(q) \quad s.t. \quad p^k q^k = \omega^k(p, w) \forall k = 1, \cdots, K.$$

The Slutsky matrix function of this consumer is given by:

$$S_{i,j}(p, w) = \frac{\partial x^r_i(p, \omega(p, w))}{\partial p_j} + [\nabla \omega^r x^r(p, \omega(p, w)) x^r_j(p, \omega(p, w))].$$

$^{20}$A sufficient condition for $E_{i,j}(p, w) = 0$ is that the marginal utility of income is equalized
across the mental accounts. In general, $\|E(p, w)\|^2$ is given by the expression:

$^{20}$Farhi & Gabaix (2015) prove that the modified $k$-compensated Slutsky matrix is symmetric, NSD
and homogeneous of degree zero. The $k$–compensated Slutsky matrix is: $S^k_{i,j}(p, w) = \frac{\partial x^r_i(p, \omega(p, w))}{\partial p_j} + \frac{\partial x^r_j(p, \omega(p, w))}{\partial \omega^k} x^r_i(p, \omega(p, w)).$
\[ \|E(p, w)\|_M^2 = \sum_{i,j=1}^{L} \left( \frac{\partial x_i^r(p, \omega(p,w))}{\partial p_j} - \frac{\partial x_j^r(p, \omega(p,w))}{\partial p_i} + \sum_{k=1}^{K} \left( \frac{\partial x_i^r(p, \omega(p,w))}{\partial \omega^k} x_j^r(p, \omega(p,w)) - \frac{\partial x_j^r(p, \omega(p,w))}{\partial \omega^k} x_i^r(p, \omega(p,w)) \right) \frac{\partial \omega^k(p, w)}{\partial w} \right)^2. \]

**Remark 9.** The main conclusion from this analysis is that for any two consumers that have a demand function consistent with the mental accounting model, that differ only in its mental accounts \( \omega^{1,k}(p, w) \neq \omega^{2,k}(p, w) \) for some \( k = 1, \cdots, K \): We can observe, that a point \( z = (p, w) \in Z \) such that \( x^{r,1}(p, \omega(p,w)) = x^{r,2}(p, \omega(p,w)) \), \( \frac{\partial x_i^{r,1}(p, \omega(p,w))}{\partial p_j} = \frac{\partial x_i^{r,2}(p, \omega(p,w))}{\partial p_j} \), \( \frac{\partial \omega^{1,k}(p, w)}{\partial w} = \frac{\partial \omega^{2,k}(p, w)}{\partial w} \); consumer 1 is more boundedly rational than consumer 2, according to the Slutsky matrix norm, if the differences between the marginality utility of income across the mental accounts is bigger for the consumer 1 that for consumer 2.\(^{21}\)

### 7.4 Demand Function Generated by Unobserved Sequential Choice:

**Naive Quasi-Hyperbolic Discounting**

In this example, we study the case of a demand function that is generated by a sequential decision-making procedure, that is latent or hidden for the modeller. In particular, the modeller observes a (static) demand function when the actual decision-making takes place in a three-period environment. We assume that the consumer cannot make her consumption decision for all goods at once. Instead, the consumer receives an amount of wealth at the initial time period. Then she must first decide how much to buy from good 1, then how much to purchase from good 2, and the final day, she spends the remaining money in the third good. Our aim is to measure the violations of property \( \mathcal{R} \) by a naive and sophisticated quasi-hyperbolic discounters. The optimization problem for a consumer that can pre-commit is: \( \max_{\{x_i^p\}_{i=1,2,3}} u(x_1^p) + \beta u(x_2^p) + \beta^2 u(x_3^p) \) subject to

\(^{21}\)Formally, when \( |\frac{\partial x_i^{r,1}(p, \omega(p,w))}{\partial \omega^k} - \frac{\partial x_i^{r,2}(p, \omega(p,w))}{\partial \omega^k}| \geq |\frac{\partial x_i^{r,1}(p, \omega(p,w))}{\partial \omega^{k'}} - \frac{\partial x_i^{r,2}(p, \omega(p,w))}{\partial \omega^{k'}}| \) for all \( k, k' = 1, \cdots, K \) mental accounts and for all \( i, j = 1, \cdots, L \) goods.
to the budget constraint \( \sum_{i=1}^{3} p_i x_i^p = w. \)

This is called the exponential discounting solution and it produces a demand system that is consistent with rationality. In contrast, the naive quasi-hyperbolic discounter will have the following demand system: In the first period, the consumer assumes she will stick to her commitment in the second period, and consumes the same amount as in the pre-commitment case (i.e. \( x_1^h = x_1^p \)). However, when period 2 arrives, she re-optimizes taking as given the remaining wealth and chooses:

\[
x_2^h = \frac{w - p_1 x_1^h}{p_2 + p_3 (\beta \theta \frac{p_2}{p_3})^{1/\sigma}}, \quad x_3^h = [\beta \theta \frac{p_2}{p_3}]^{1/\sigma} x_2^h.
\]

The analytical result for the matrix nearness norm has a nice expression (for simplicity we fix \( p_j = 1, w = 1 \) for \( j = \{1, 2, 3\} \)):

\[
||E^h(p, w)||_M^2 = \frac{3}{2} \frac{2}{\sigma^2} \left( \frac{\sigma - 1}{\sigma^2} \left( (\beta \theta)^{1/\sigma} - (\beta \theta^2)^{1/\sigma} \right)^2 \right)
\]

which readily gives us that: \( ||E^h(p, w)||_M^2 = 0 \) when either (i) \( \sigma = 1, \) and \( \delta = 0 \) (i.e., the demand is rational); (ii) or \( \beta = 1, \) and finally (iii) when \( \beta \to 0, \theta \to 0. \) 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 \( x_1^h \to \frac{w}{p_1} \) and \( x_2^h, x_3^h \to 0. \) In case (i), the logarithmic utility case, the hyperbolic discounter manages to keep her commitment and therefore her consumption is time consistent and \( ||E^h(p, w)||_M^2 = 0. \)

Remark 10. For any two consumers consistent with the naive quasi-hyperbolic behavior, who have the same CRRA utility, but different present-bias parameters \( \beta_2 > \beta_1, \) one can conclude that consumer 1 is more boundedly rational than consumer 2 (when \( \sigma > 1, \) and \( \theta \in (0, 1) \)). This means that the higher the present bias, the higher the Slutsky matrix norm. Another observation that we can draw from this example is that only the ICPP anomaly appears (i.e., \( ||E||^2 = ||E^x||^2 \)). This seems to make sense, as increases in purchasing power may be offset by the sequential nature of the consumer choice and the temporal inconsistencies it introduces.

\[
22\text{The solution of this problem gives the demand system: (i) } x_1^p = [p_1 + p_2 (\beta \theta \frac{p_2}{p_1})^{1/\sigma} + p_3 (\beta \theta^2 \frac{p_2}{p_1})^{1/\sigma}]^{-1} w.
(ii) \ x_2^p = [\beta \theta \frac{p_2}{p_1}]^{1/\sigma} x_1^p \ (iii) \ x_3^p = [\beta \theta^2 \frac{p_2}{p_1}]^{1/\sigma} x_1^p.
\]
7.5 Demand Function Generated by Unobserved Sequential Choice: Sophisticated Quasi-Hyperbolic Discounting

One would expect that the sophisticated quasi-hyperbolic discounter is less boundedly rational than the naive case. However, the Slutsky matrix norm helps to appreciate some of the subtleties and assess which conditions of rationality are fulfilled by this type of consumer. In fact, we establish that the sophisticated quasi-hyperbolic discounter can generate demand functions that are more boundedly rational in the sense of the Slutsky matrix norm than the naive case. In other words, actions taken by the sophisticated consumer, understanding her time inconsistency, may take her further away from the rationality paradigm in terms of small demand reactions to price changes.\(^{23}\) In the sophisticated 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:

\[
\max_{x_{sh}^1} u(x_{sh}^1) + \beta \theta u(x_{h}^2) + \beta \theta^2 u(x_{h}^3),
\]

where \(x_{h}^2, x_{h}^3\) (conditional on \(x_{h}^1\)) are known to her in \(t = 1\). However, she can control only how much she consumes in the first period.\(^{24}\)

The Slutsky matrix norm for this consumer is given by the quantity (for simplicity we fix \(p_j = 1, w = 1\) for \(j = \{1, 2, 3\}\)):

\[
||E^{sh}(p, w)||^2_M =
\]

\[^{23}\text{In fact, when trying to predict the local behavior (demand changes due to small price changes), a modeller may make bigger errors with a sophisticated hyperbolic consumer than with a naive one, when imposing rationality.}\]

\[^{24}\text{Then, the first period consumption under sophisticated hyperbolic discounting is: } x_{1}^{sh} = \left[\frac{p_1 \beta \theta + p_1 \beta \theta^2 \epsilon_{p_1}}{p_2 + p_3 \beta \theta \epsilon_{p_2}^{\frac{1}{2}}} \right]^{\frac{2}{3}} + p_1 \right]^{-1} w.\]
\[ \frac{3}{2} \sigma^4 \left( (\beta \theta)^{\frac{1}{2}} + 1 \right)^2 \left( (\beta \theta)^{\frac{1}{2}} + \beta \right)^2 \left( \theta \left( (\beta \theta)^{\frac{1}{2}} + 1 \right)^{-1} \left( (\beta \theta)^{\frac{1}{2}} + \beta \right)^{\frac{1}{2}} + 1 \right)^4. \]

This implies that: (i) when \( \sigma = 1, \delta = 0 \) for any \( Z \); (ii) when \( \beta = 1, \delta = 0 \); and (iii) when \( \beta = 0, \delta = 0 \). Also, it follows that \( ||E||^2 = ||E^\sigma||^2 \), which means that only the ICCP anomaly appears in this case. Now, we compare this quantity with the case of the naive hyperbolic discounter.

We consider the ratio \( r = \frac{||E^{sh}(p,w)||^2_M}{||E^{nh}(p,w)||^2_M} \). When \( r < 1 \), the sophisticated hyperbolic consumer has a lower Slutsky matrix norm for any parameter configuration. The ratio \( r \) depends crucially on the parameter \( \sigma \). For \( \sigma = 1 \), both \( E^{sh}, E^{nh} \) are equal to zero: this is a knife-edge case, in which the marginal rates of substitution yield optimal consumption equal to the commitment baseline. For \( \sigma < 1 \), the sophisticated hyperbolic discounter has a uniformly lower \( ||E||^2_M \) and thus \( r < 1 \). However, for \( \sigma > 1 \), the naive hyperbolic discounter has a uniformly lower \( ||E||^2_M \), hence \( r > 1 \). Although this may seem counterintuitive, it tells us that the closest rational consumer (which need not be the commitment baseline) is closer to the naive than to the sophisticated consumer. To see this, note that the rational behavior fails due to two possible sources: (i) The lack of self-control in the consumption of period 2 (overspending); and (ii) in the case of the sophisticated type, an additional effect due to the strategic change in period 1 consumption (overspending or underspending). Hence, we find a higher failure for the sophisticated case than in the naive case for \( \sigma > 1 \), because for an increase in prices, for the naive case, the change in period 1 consumption equals that in the rational case; and only the second and third period demand changes are different. In contrast, for the sophisticated case, all three good demand changes differ from the rational benchmark. In particular, the divergence is due to a higher residual wealth at the end of period 1, leading to a higher overspending in period 2, for the sophisticated case.
7.6 Other Models

Our framework can be applied to any behavioral consumer model as long as it can be written in the classic consumer theory environment (namely, if the model produces a continuously differentiable demand function, as in the previous two applications). We note, though, that not every behavioral model can be analyzed using our approach. One important example is the model of “Salience in Consumer Choice” by Bordalo et al. (2013), as it stands. Nonetheless, although the general principles that the salience model proposes are cast in a discrete choice set framework, it could be translated into the classic consumer theory environment that we consider here. This should be an interesting avenue for future research.

8 Empirical Application: Shape Constraints Classification in Demand Estimation with Heterogeneity and Measurement Error

Here we show how our theoretical results can be applied to verify and quantify the appropriateness of imposing the Slutsky regularity conditions as shape constraints, in the presence of heterogeneity (and measurement error). Demand estimation modellers often impose shape constrains, or in the language of this paper, they impose the \( \sigma, \pi, \) and \( \nu \) properties on their demand estimators, under the assumption that the population of consumers is consistent with rationality. However, there is a growing interest in testing the appropriateness of imposing the shape constrains in the demand estimation exercise (see Haag et al. 2009; Hoderlein 2011, and Dette et al. (2016)). This section provides a framework that allows the use of our Slutsky decomposition result in Theorem 1 to quantify the appropriateness of each individual shape constrain in a simple and clear way. We are able to do this when we restrict heterogeneity among consumers in the data set in a way that is similar to Blundell et al. (2014).

Our theoretical exercise has maintained the assumption that the modeller can observe
the demand function of the individual consumer. In reality, the modeller may only have access to a cross-section or a panel data set of consumers. The typical data set contains information about joint distribution of the individual consumers’ budget shares for \( L \) goods, the prices they face, and their total expenditure level (or wealth). We show that, even when the individual demand is not directly observable, it can be estimated from the population data.

Formally, consider the case of \( L - 1 \) dimensional budget share function, such that \((p, w, \epsilon) \mapsto b(p, w, \epsilon) \in [0, 1]^{L-1} \) where \( z = (p, w) \in Z \) and \( \epsilon \) is the realization of a random variable \( \xi \) that represents, following Blundell et al. (2014), a (time-invariant) individual specific heterogeneity term that reflects unobserved heterogeneity in preferences. We can make assumptions on the function \( b \) such that we are able to recover \( b(p, w, \epsilon) \) from observing the joint distribution of budget shares, prices, and wealth, \((B', P', W')\)' (where capital letters denote random variables and lower-cases denote realizations or deterministic objects).

8.1 Heterogeneity

We follow Blundell et al. (2014) to restrict the heterogeneity in consumption. Our main restriction are a monotonicity requirement on the budget share function, and a normalization about the distribution of the heterogeneity in a nonparametric environment.

Assumption 1. (Uniform Heterogeneity) We assume that \( B_l = b_l(P, W, \xi) \) for \( \xi \sim U[0, 1] \) for all \( l = 1 \cdots, L - 1 \).

We must emphasize that the heterogeneity term may affect the budget share vector nonlinearly through the nonparametric deterministic function \( b_l \). The distributional restriction of uniformity is a normalization, given that we have a nonparametric \( b \) (Blundell et al., 2014).

The second restriction controls how the good-specific budget share is affected by the heterogeneity term. We impose that the heterogeneity term affects the budget share in a monotone way. This is a common restriction in the literature (Blundell et al., 2014; Cunha et al., 2010).
Assumption 2. (Invertibility) The function $b_l(p, w, \cdot)$ is monotone almost surely for all $l = 1 \cdots, L - 1$.

This assumption can be interpreted easily, when the heterogeneity term corresponds to an individual specific taste shock, and the good $L$ is an outside alternative.

We finally consider an exclusion restriction condition that is also standard in the econometric literature interested in estimating the individual demand information from the population.

Assumption 3. (Nonsystematic Mistakes/Heterogeneity) We let $F_{\xi|P,W} = F_{\xi}$, where $F_{\xi|P,W}$ is the conditional C.D.F. of $\xi$ on $P,W$, and $F_{\xi}$ is the marginal C.D.F. of $\xi$.

The budget share system is given by $b(p, w, \epsilon)$ and by the budget share of good $L$, denoted as $b_L(p, w, \epsilon) = 1 - \sum_{l=1}^{L-1} b_l(p, w, \epsilon)$. (This is consistent with Walras’ law, maintained throughout the paper.) We define $Pr(B \leq b|P,W) = G(b|P,W)$ and we observe that if we consider the quantile $\tau \in [0,1]$ we can identify $b(p, w, \epsilon)$ in a similar fashion to (Blundell et al., 2014).

Lemma 3. If Assumptions 1, 2, and 3 hold, we can identify $b(p, w, \epsilon)$ from the joint distribution of $(B', P', W')$, i.e., it must be that $b_l(p, w, \tau) = G_l^{-1}(\tau|P = p, W = w)$ for all $\tau \in [0,1]$.

The small difference of our statement with Blundell et al. (2014) is that we consider an environment where variation in prices is infinite as well as variation in wealth and budget shares, while Blundell et al. (2014) consider finite variation in prices. We think our environment is realistic for consumer scanner data (e.g., Nielsen Scanner data), or high frequency consumer data (where we have many measures in a short interval of time), while the type of data set in Blundell et al. (2014) is more common in survey data. With the previous result in hand, we can state the main result of this section, which allows the modeller to quantify the appropriateness of imposing shape constraints $\sigma, \pi$, and $\nu$ in demand quantiles. This is important, because it allows the modeller to let the data speak.

---

25Extension to consider endogeneity of prices and the use of instrumental variables to correct for it, have been explored in (Blundell et al., 2014), and for concreteness, are ignored in this section.
and reveal which of the three shape constraints implied by rationality is more appropriate. Seen in this light, our exercise goes beyond testing and is closer to learning the patterns in consumer behavior. Before stating our main result we observe that the Slutsky matrix function \( S(p, w, \epsilon) \) associated to the budget share function \( b(p, w, \epsilon) \) is also identified, as it can be computed from \( b(p, w, \tau) \).

**Proposition 3.** If Assumptions 1, 2, and 3 hold, the Slutsky matrix decomposition of \( S(p, w, \epsilon) \), associated with \( b(p, w, \epsilon) \), is identified, i.e., it must be that for every \( \tau \in [0, 1] \) and for every \( z = (p, w) \in \mathcal{Z} \), the Slutsky matrix function \( S(p, w, \tau) \) associated with \( b_l(p, w, \tau) = G^{-1}_l(\tau | P = p, W = w) \) has the unique orthogonal decomposition \( S(p, w, \tau) = S^e(p, w, \tau) + E(p, w, \tau) \), where \( S^e(p, w, \tau) = S^{\sigma, \pi, \nu}(p, w, \tau) \) is the solution to the matrix-nearness problem and \( E(p, w, \tau) \) is the sum of the orthogonal complements of the symmetric, \( p \)-singular, and NSD parts of \( S(p, w, \tau) \): \( E(p, w, \tau) = E^e(p, w, \tau) + E^s(p, w, \tau) + E^\nu(p, w, \tau) \). Furthermore, the Slutsky norm of the solution to the matrix-nearness problem can be decomposed as follows:

\[
||E(p, w, \tau)||^2_M = ||E^e(p, w, \tau)||^2_M + ||E^s(p, w, \tau)||^2_M + ||E^\nu(p, w, \tau)||^2_M.
\]

The proof of the previous result is omitted, as it follows trivially from Theorem 1 and Lemma 3. Proposition 3 represents a novel way to quantify the appropriateness of the shape constraints \( \sigma, \pi, \) and \( \nu \). Notice that the modeller can now see which of the three shape constraints is more difficult to maintain for every quantile \( \tau \in [0, 1] \). For instance, it may be the case that some quantile \( \tau \in [0, 1] \) (associated with \( E(p, w, \tau) \)) has a low value of \( \nu \) and high values for the errors associated with \( \sigma, \pi \), while \( \tau \in [0, 1] \) has a low value of \( \sigma, \pi \) and a high value for \( \nu \), in this case, our methodology allows the modeller to identify the heterogeneity in departures from rationality as a function of the quantile.

This framework has as a special case the random utility model. We say a random budget share \( b(\cdot, \cdot, \xi) \) is consistent with a random utility maximizer subject to a linear constraint almost surely if and only if \( \text{Prob}[\xi = \epsilon : x(\cdot, \cdot, \epsilon) = \arg \max_{q:p = w} u(q, \epsilon)] = 1 \) where \( b_l(\cdot, \cdot, \epsilon) = p_l x_l(\cdot, \cdot, \epsilon)/w_l \) and \( u(\cdot, \epsilon) \) is a realization of a random utility device.
Corollary 2. If Assumptions 1, 2, and 3 hold, and \( b(\cdot, \cdot, \xi) \) is generated by a random utility maximizer subject to a linear constraint almost surely, it follows that \( T(p, w) = \int_0^1 ||E(p, w, \tau)||_M^2 d\tau = 0 \) for all \( z = (p, w) \in Z \).

The amount \( T(p, w) \) is a test statistic for random utility for every \( z = (p, w) \in Z \). Evidently, it also follows that the average test statistic also tests for random utility \( T = \int_0^1 \int_{z \in Z} ||E(p, w, \tau)||_M^2 \mu(z) dz d\tau = 0 \). This simple observation is helpful, as averaging among observations can be used to obtain semiparametric rates of convergence as in Newey (1994). We leave this as a future avenue of research, taking us beyond the scope of this theoretical work.

8.2 Measurement Error

The type of environment that we have considered may be prone to measurement error. Formally, we need to obtain \( G_l(b|P, W) = Pr(b \leq B^*_l|P, W) \), when we observe two perturbed or imperfect measures of the budget-share random variable \( B^*_l = B^*_l + \eta^*_l \) for \( t = 1, 2 \), where \( B^*_l = b_l(P, W, \xi) \) is the perfectly measured budget-share random variable of the previous subsection.

This can be done following Schennach (2016), when we assume that \( \eta^*_l \perp B^*_l|P, W \) for all \( l \) and \( t \). We also need to assume that \( \eta^1 \) and \( \eta^2 \) are mutually independent, \( E[\eta^*_l|P, W] = 0 \), and \( E[e^{itB^*_l}] \neq 0 \); then we can use Kotralski’s lemma to find the distribution of \( B^*|P, W \) (using Li (2002) extension of the lemma for multivariate settings). In this way, repeated measurements can provide an estimate of \( G_l(b, P, W) \), in the presence of measurement error. Then we can apply the results in the heterogeneity subsection. This exercise can be greatly simplified if the modeller knows the distribution of the measurement error. In this case we do not need repeated measures and we can use the deconvolution theorem to achieve our objective (Schennach, 2016). In specific applications, we can think of the repeated measures as coming from panel consumer data or from validation data sets.

In particular, using Roy’s lemma we have that \( x(p, w, \epsilon) = -\frac{\Sigma_{w}v(p, w, \epsilon)}{\partial v(p, w, \epsilon)/\partial w} \), where \( v(p, w, \epsilon) = max_{q, p' = w} u(q, \epsilon) \). We say that that \( u(\cdot, \epsilon) \) is deterministic since all randomness is given by the taste shock \( \xi \).
8.3 Finite Data Sets and the Slutsky Matrix Norm

The Slutsky matrix norm approach can be adapted and used in a finite data set environment. The central idea is that a finite collection of observations of demand choices and prices can be used to obtain demand function interpolators or extensions. Using this idea, we can build the minimal Slutsky matrix norm associated with an extension of a finite data set. This approach can be used to learn and classify patterns of boundedly rational consumer behavior in experimental data sets. We provide a formal discussion of these ideas and an empirical application to the experimental data set of Ahn et al. (2014) in the Supplement to this paper. We find that the violations of property $\sigma$ are more prevalent than the violations of $\nu$ in the data set of individual consumers (the violations of $\pi$ cannot happen in the considered experiment by design).

9 Related Literature

The treatment closest to our work is the approximately rational consumer demand proposed by Jerison and Jerison (1992; 1993); see also Russell (1997). Jerison and Jerison (1992; 1993) are able to relate local violations of negative semidefiniteness and symmetry of the Slutsky matrix to the smallest distance between an observed smooth demand system and a rational demand. Russell (1997) proposes a notion of quasirationality. Russell’s argument links the Slutsky matrix antisymmetry part with the lack of integrability of a demand system. Our work takes a global approach to the problem of measuring the size of bounded rationality and generalizes the results to the case of violations of singularity of the Slutsky matrix. Our new approach allows us to treat the three kinds of violations of the Slutsky conditions simultaneously.

Using limited data sets—as opposed to entire demand functions, as we do—the canonical treatment of measuring deviations from rational consumer behavior was established 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 (1985; 1990) refines Afriat’s measure
by focusing on the minimum adjustment of the budget constraint needed to eliminate violations of GARP. Houtman and Maks (1985) measure deviations from GARP through identifying the largest subset of choices that is consistent with maximizing behavior. More recently, Echenique, Lee, and Shum (2011) propose a new measure of violations of revealed preference behavior called the “money pump index.” Also, Jerison and Jerison (2012) propose a way to bound Afriat’s index of cost-efficiency, using an index of violations of the symmetry and negative semidefiniteness Slutsky conditions. See also Dean & Martin (2015), Apesteguia & Ballester (2015), among others. It would be interesting to compare our Slutsky matrix norm with these other approaches and in fact, in Aguiar and Serrano (2016), we begin that line of work. The revealed preference tradition offers a number of welfare or normative approaches, which implicitly assume that the consumer wants to optimize, but fails to reach her goal. In contrast, our approach is positive. We are concerned with failures in comparative statics analysis for any given consumer behavior when we compare it to rationality.

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

10 Conclusion

By using the Slutsky matrix—a cornerstone of microeconomic theory—we provide a tractable and unifying framework to measure and classify the different kinds of bounded rationality that a consumer choosing over linear budget sets may exhibit. The size of bounded rationality can be decomposed as the sum of three effects, each of which cap-
tures a different anomaly: inattentiveness to changes in purchasing power (ICPP), money illusion (MI), and violations of the compensated law of demand (VCLD). 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 better computational implementability and provides a tractable approach with a closed-form solution. Our approach is also suited to measuring the specification error that arises from assuming a given form of rationality when the consumer’s behavior is actually not rational.

References


Appendix

This appendix provides the proofs of our results, and includes additional material relevant to our approach.

10.1 Difficulties of the Behavioral Nearness Problem

As mentioned in Subsection 2.2, the behavioral nearness problem $d(x, R) = \inf \{d_X(x, x^\tau)|x^\tau \in R\}$ presents some difficulties in terms of existence, uniqueness, tractability, as well as being subject to the curse of dimensionality of the calculus of variations. More importantly, it does not lend itself to a decomposition in axioms of revealed demand, unlike the matrix-nearness problem. We illustrate some of these difficulties here. To enhance comparability with the matrix-nearness problem, we choose the quadratic or euclidean norm for $X$, $d_X(x, x^\tau) = ||x - x^\tau||_X$ where $||x - x^\tau||_X = \int_{z \in Z} [x(z) - x^\tau(z)]' [x(z) - x^\tau(z)] dz$. The behavioral nearness problem can be then expressed as $\inf_{x^\tau \in R(Z)} ||x - x^\tau||_X$. Thanks to Grodal (1974) and Mas-Colell (1974), we know that the set of rational demands $R(Z)$ is ill-behaved. It is neither closed nor open, which affects existence. It is not convex either, which affects uniqueness. In Section 6 we have made the behavioral nearness problem operational, though: we can guarantee existence by limiting our search for the nearest rational demand in a closed and bounded subset of rational demand functions in $X(Z)$, $\overline{R}(Z) \subset R(Z)$. For our purposes, this is not a very strong restriction, allowing us to write the problem as $\min_{x^\tau \in R(Z)} ||x - x^\tau||_X$. In contrast, imposing convexity can result in an important loss of generality. Convexity of $\overline{R}(Z)$ means that the result of taking any $x, y \in \overline{R}(Z)$ and aggregating them with some weight $\alpha \in (0, 1)$, results in a rational demand $\hat{x} = \alpha x + (1 - \alpha)y \in \overline{R}(Z)$. In general, this is equivalent to saying that the Slutsky matrix function of the aggregate demand $S(\hat{x})$ has the property $R$. But the aggregation of rational demands is known to be very restrictive, in particular, only small subsets of rational models such as those with an indirect utility that admits a Gorman form are known to aggregate (i.e., meaning that $\hat{x}$ will be rational). The lack of convexity of the set $\overline{R}(Z)$ will prevent us, in general, to obtain a uniqueness result.
Even if nonuniqueness is overlooked, there may be problems with tractability, related to the curse of dimensionality of the calculus of variations. To illustrate this point we will consider an extremely simplified version of the behavioral nearness problem.

we only impose property $\sigma$ on the Slutsky matrix associated with $x \in \mathcal{X}$. Imposing additional constraints associated with $\nu$ and $\pi$ will exacerbate the problem. With this in hand, we can write the simplified behavioral nearness problem as:

$$\min_{x \in \mathcal{X}} ||x - x'||$$

subject to $D_p x'(z) + D_w x'(z)x'(z) - D_p x'(z) + x'(z)D_w x'(z)'$ for all $z \in Z$. Existence is guaranteed by choosing a subset $\mathcal{X}(Z) \subset X(Z)$ that is closed, and such that the set of demands that satisfy the restriction is closed as well, and the objective is continuous in $x'$.

This is a variational problem that can be solved using a Euler-Lagrange equation. For now, we are ignoring Walras’ law, that can be easily imposed by considering the demand sub-system $x_{-L} = (x_1, \cdots, x_{L-1})'$.

The Lagrangian associated with the problem is:

$$\mathcal{L} = \int_{z \in Z} [x(z) - x'(z)]'[x(z) - x'(z)]dz +$$

$$\int_{z \in Z} \lambda_{L^2 \times 1} vec[D_p x'(z) + D_w x'(z)x'(z) - D_p x'(z) + x'(z)D_w x'(z)']dz.$$ 

The Euler-Lagrange equations are:

$$\frac{\partial \mathcal{L}}{\partial x_l} - \sum_{a \in \{p_1, p_2, \cdots, p_L, w\}} \frac{\partial \mathcal{L}}{\partial a} \frac{\partial x_l}{\partial x_l} = 0 \quad \forall l \in \{1, \cdots, L\}$$

$$D_p x'(z) + D_w x'(z)x'(z) - D_p x'(z) + x'(z)D_w x'(z)' = 0.$$ 

This helps to illustrate the curse of dimensionality: the Euler-Lagrange equations consists of a system of $L$ nonlinear partial differential equations, in addition to $L^2$ partial
differential equations coming from the restriction.

The lack of tractability can be seen when we simplify the Euler-Lagrange equations further:

\[ x_l(z) - x'_l(z) + \sum_{k=1}^{L} \lambda_{\rho(k)} \frac{\partial x_k(z)}{\partial w} = 0 \quad \forall l \in \{1, \ldots, L\} \]

\[ S(x^r) = S(x^r)', \]

where \( \rho(k) \) maps \( k \) to the appropriate index of the multiplier vector (this is needed due to the vectorization of the constraint). An analytical solution for this problem is hard (Kawamura et al., 2012).

Finally, the behavioral nearness problem cannot take advantage of the orthogonalities, furnished by the matrix function space and which do not translate to the function space.\(^{27}\)

\section*{10.2 Proofs}

\textbf{Proof of Claim 1}

\textit{Proof.} The problem is \( \min_{S^r} ||S - S^r|| \) 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

\[ \min_{S^r} \int_{z \in Z} Tr([S(z) - S^r(z)]'[S(z) - S^r(z)])dz \]

subject to the stated constraints.

The objective function is strictly convex, because of the use of the Frobenius norm.

This norm is also a continuous functional.

\footnote{Alternative approaches include Blundell et al. (2008), which imposes SARP as restrictions, work numerically for cases with finite variation in prices. In our environment, demand changes with prices continuously, and prices can take any values in \( P \subseteq \mathbb{R}_+^L \). More realistically, there may be situations where prices have a high variation such as in our empirical application. In those cases, it is not practical to impose SARP as a restriction, since it involves ruling out a possibly infinite number of revealed preferences cycles.}
nite 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 $\mathcal{M}_{\mathcal{R}}(Z)$ 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 $\mathcal{M}_{\mathcal{R}}(Z)$ is itself a subspace of $\mathcal{M}(Z)$.

Now we prove that not only the symmetric and the NSD constraints sets are closed but so is the set of all $\mathcal{M}_{\mathcal{R}}(Z)$. Any matrix function in the constraint set is a symmetric NSD matrix with $p$ in its null space. Therefore, every sequence of matrix functions in the constraint set has the form $D^n(z) = Q^n(z)\Lambda^n(z)Q^n(z)'$, where $\Lambda^n(z) = \text{Diag}[\lambda^n_i(z)]_{i=1,...,L}$ with ascending ordered eigenvalues functions. It follows that the eigenvalue function in position $L, L$ is the null eigenvalue $\lambda_L = 0$, or the null scalar function. That is, imposing an increasing order the position $1, 1$ is then held by $\lambda^n_1(z) \leq \lambda^n_2(z) \leq \ldots \leq 0$ where the order is induced by the distance to the null function using the Euclidean distance for scalar functions defined over $Z$. The matrix function $Q^n(z) = [q^n_1 \ldots p]$ is the orthogonal matrix with eigenvectors functions as columns. For all $D^n(z) \in \mathcal{M}_{\mathcal{R}}(Z)$, $\lambda^n_L = 0$ is associated with the price vector $q^n_L = p$ always, to guarantee that $p$ is in its null space. The eigenvectors are defined implicitly by the condition $D^n(z)q^n_i(z) = \lambda^n(z)q^n_i(z)$ with pointwise matrix and vector multiplication and $q^n_i \perp p$ or $\langle q^n_i, p \rangle = 0$ using the inner product for $C^0(Z)$ for $i = 1, \ldots, L - 1$ and for all $n \in \mathbb{N}$. Take any sequence of $\{D^n(z)\}_{n \in \mathbb{N}}$ with $D^n(z) \in \mathcal{M}_{\mathcal{R}}(Z)$ for each $z \in Z$, with limit $\lim_{n \to \infty} D^n(z) = D(z)$. We want to show that $D(z) \in \mathcal{M}_{\mathcal{R}}(Z)$. 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 $\lambda^n_L = 0$ for all $n$ and certainly $\lambda^n_L \to 0$ with the associated eigenvector $q^n_L = p$ for all $n$ and $q^n_L \to p$. This condition, along with 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 $\lambda_{i \neq L}(z)_- = \min(0, \lambda_{i \neq L}(z))$. 

55
This is a negative scalar function by construction, since if \( \lambda_i \neq L(z) \) then \( \lambda_i = L(z) - 0 \). Then \( M_{\mathcal{R}}(Z) \) is closed. Observe that \( M_{\mathcal{R}}(Z) \) is closed since it is in the intersection of three closed sets. In conclusion, since the Frobenius norm in \( M(Z) \) is a continuous and strictly convex functional and the constraint set is closed and convex, the minimum is attained and it is unique.

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

**Proof of Lemma 1**

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

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

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

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

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

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

Solve for \( A(z) \):

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

By the symmetry restriction,

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

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

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

We obtain the expression:

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

Now we observe that we can write,
\[ A(z) = S^\sigma(z) + W^\sigma(z) \text{ with } W(z) = \lambda(z)p' \text{ and } W^\sigma = \frac{1}{2}[p\lambda(z)' + \lambda(z)p'], \text{ the symmetric part of } W(z). \]

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

To complete the proof we let \( r(z) = -S^\sigma(z)p \) be the feedback error (of non-singularity in prices). Then we can write the matrix function \( W^\sigma(z) = r(z)p' + pp'S^\sigma(z)p' \) that is the symmetric part of the outer product of the feedback error and the price vector.

Replacing \( r(z) \) by its definition, we get that \( W^\sigma = -E^\pi; \)

\[ E^\pi(z) = \frac{1}{p'p}[S^\sigma(z)pp' + pp'S^\sigma(z) - \frac{1}{p'p}pp'S^\sigma(z)pp'], \text{ then } E^\pi = S^\sigma - PS^\sigma P \text{ with } P = I - \frac{pp'}{p'p} \text{ and by Walras' law it simplifies to:} \]

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

this expression along with the implied multipliers \( \lambda \) and \( U \), satisfies all the first-order conditions of the problem. Since we can use arguments identical to those in Claim 1 –only not imposing 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.

\[ \square \]

**Proof of Claim 2**

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

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

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

and

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

To see that this is a direct sum decomposition, first observe that \( \mathcal{P}(z) \cap \mathcal{N}(z) = \{0\} \), with
0 denoting the zero matrix function, by construction. Furthermore, any \( A(z) \in \mathcal{A}(z) \) can be written as a sum of \( A(z) = E^\sigma(z) + N(z) \) since \( A(z)p = 0 \) or (exclusive) \( A(z)p \neq 0 \), for \( A(z) \neq 0 \). Furthermore, \( p'A(z)p = p'E^\sigma(z)p + p'N(z)p = 0 \) for any \( E^\sigma(z), N(z) \). Then the decomposition is precisely \( A(z) = E^\sigma(z) \) 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 \( E^\nu(z)p = 0 \) and \( Tr(E^\nu(z)pp') = 0 \), it follows that \( Tr(E^\pi(z)E^\nu(z)) = 0 \), for \( z \in Z \).

**Proof of Lemma 2**

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

Now, the best NSD matrix approximation of the symmetric valued function \( S^\sigma,\pi \) is \( S^r = S^\sigma,\pi,I \). 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 \( S^r(z) = S^\sigma,\pi,I(z) = Q(z)\Lambda(z)Q(z)' \). 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 eigen-
value function associated with the $q_L(z) = p$ eigenvector. Then we have that $\lambda_L(z)_- = 0$
is also associated to the eigenvector $p$, and we can conclude that $S^r(z)p = 0$.

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

Proof of Proposition 1

By Hurwicz & Uzawa (1971) we know that the demand $x$, satisfies Walras’ law (in our
case, has a compact domain $Z$), satisfies WARP, and has an Slutsky matrix function
$S$ that satisfies symmetry if and only if there is a continuous utility function such that
$x(p, w) = \argmax_q u(q) \text{ s.t. } p'q = w$.

We want to show that the Slutsky matrix nearness norm $||E||^2$ will be equal to zero
if and only if $x$ satisfies VARP, HD0, and the Wald Axiom (VARP and the WARP).

We first prove that if VARP, HD0 and the Wald Axiom (VARP and the WARP) holds
then $||E|| = 0$ by proving the moreover part of the statement.

(i) If the Ville Axiom holds by Hurwicz & Richter (1979) we have that $S$ is symmetric,
or $E^\sigma = 0$. This implies that $||E^\sigma||^2 = 0$. The contrapositive implies that $||E^\sigma||^2 > 0$
then the Ville axiom fails.

(ii) If the HD0 holds then we have that $S(p, w)p = 0$, by Walras’ law we have that
$p'S(p, w) = 0$. This means that $S^{\sigma,\pi} = S^\sigma$, because $P$ is a projection matrix to the
space of matrix functions such that $p'S(p, w)p = 0$ and $S(p, w)p = 0$. This implies that
$PS^\sigma P = S^\sigma$. Then $||E^\pi|| = 0$. The contrapositive establishes that $||E^\pi|| > 0$ then HD0
fails (as Walras’ law is assumed to hold for $x$).

(iii) If WARP holds then we have that HD0 holds (John, 1995) because we are as-
suming Walras’ law. If WARP holds by Kihlstrom et al. (1976) we know that $S$ is NSD.
This implies that $S^{\sigma,\pi,\nu} = S^\sigma$, by HD0 (and (ii)) we have that $S^{\sigma,\pi} = S^\sigma$, and by the
fact that $S$ is NSD then $S^\sigma$ is also NSD, note that $v'Sv \leq 0$ for all $v \in \mathbb{R}^L$, implies that
$v'S^\sigma v \leq 0$ because $v'\left[\frac{1}{2}[S + S']\right]v = v'Sv \leq 0$. This implies that $S^\sigma_+ = 0$. This implies
that $||E^\pi|| = 0$ and $||E^\nu|| = 0$. The contrapositive of this result implies that if either
$||E^\pi|| > 0$ or $||E^\nu|| > 0$ then WARP fails.
(iv) If the Wald Axiom holds then \( S \) is NSD (John, 1995). If \( E^\pi = 0 \), this means that \( S(p, w)p = 0 \) and \( p'(s, w) = 0 \) and \( S^\pi = S^\sigma \). Also by the arguments in (iii) we have that \( S^\sigma \) is NSD, thus \( S^\sigma \) is symmetric, then by Hurwicz & Uzawa (1971) we know that for demand \( x \) is symmetric, then by Hurwicz & Uzawa (1971) we know that for demand \( S \) is symmetric, then by Hurwicz & Uzawa (1971) we know that for demand \( x \) is symmetric. Then WARP holds by Kihlstrom et al. (1976) (in their results WARP is called Weak WARP). If \( NSD \) then WARP holds by Kihlstrom et al. (1976) (in their results WARP is called Weak WARP). If \( ||E|| = 0 \), then ||\( E^\sigma || = 0 \), this implies that \( S \) is NSD as \( S^\sigma \) is NSD. Then WARP holds by Kihlstrom et al. (1976) (in their results WARP is called Weak WARP). If \( ||E|| = 0 \), then \( ||E^\sigma || = 0 \), this implies that \( S \) associated with the demand \( x \) is symmetric, then by Hurwicz & Uzawa (1971) we know that for demand \( x \), there is a continuous utility function such that \( x(p, w) = argmax_{q} u(q) \) s.t. \( p'q = w \). We prove that this implies that there are no Ville cycles by the way of contradiction. If there is a Ville cycle or an ICPP anomaly, for instance, for a path \( (w(t), p(t)) \) that is piecewise continuously differentiable path in \( Z \). we have a rising real income situation whenever \((\partial w / \partial t)(t), (\partial p / \partial t)(t)) \) exist, with \( \partial w / \partial t(t) > \partial p / \partial t(t)'x(p(t), w(t))) \). A Ville cycle requires that: (i) \( (w(0), p(0)) = (w(b), p(b)) \); and (ii) \( \partial w / \partial t(t) > \partial p / \partial t(t)'x(p(t), w(t))) \) for \( t \in [0, b) \). By differentiation and the Walras’ law we have that this holds if and only if \( p(t)'\partial x / \partial t > 0 \) and \( x(0) = x(b) \). This implies that \( \int_0^b p(t)'\partial x / \partial t dt > 0 \) and by utility maximization we have that there is a Lagrange multiplier \( \lambda \) such that: \( \partial u / \partial t(t) = \lambda(t)'p(t)'\partial x / \partial t \), with \( \lambda(t)p(t) = \nabla_x u(x(t)), thus \( \int_0^b \partial u / \partial t dt > 0 \) but the existence of a Ville cycle implies that \( u(b) > u(0) \), which contradicts the fact that \( x(0) = x(b) \). This is a contradiction. Thus VARP must hold. Finally, by Walras’ law if WARP holds then we have that HD0 and the Wald axiom holds (John, 1995).

**Proof of Proposition 2**

*Proof.* The proof has three steps: First, we use the decomposition of a matrix into its symmetric and antisymmetric parts: \( ||S(x) - S(x^r)||^2 = ||S - S(x^r)||^2 = ||S^\sigma + E^\sigma - S(x^r)||^2 = ||S^\sigma - S(x^r)||^2 + 2\langle E^\sigma, S^\sigma - S(x^r) \rangle + ||E^\sigma||^2, \)
where \( S^\sigma - S(x^r) \) is a symmetric matrix function and \( E^\sigma \) is a skew-symmetric matrix function. Then \( \langle E^\sigma, S^\sigma - S(x^r) \rangle = 0 \). Then \( \| S(x) - S(x^r) \|^2 = \| S^\sigma - S(x^r) \|^2 + \| E^\sigma \|^2 \).

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

\[
\langle E^\pi, S^{\sigma,\pi} - S(x^r) \rangle = \langle E^\pi, S^{\sigma,\pi} \rangle - \langle E^\pi, S(x^r) \rangle = 0.
\]

Therefore, \( \| S(x) - S(x^r) \|^2 = \| S^{\sigma,\pi} - S(x^r) \|^2 + \| E^\sigma \|^2 + \| E^\pi \|^2 \).

Finally, we deal with the NSD part: \( \| S^{\sigma,\pi} - S(x^r) \|^2 = \| S^{\sigma,\pi,\nu} + E^\nu - S(x^r) \|^2 = \| S^{\sigma,\pi,\nu} - S(x^r) \|^2 + 2 \langle E^\nu, S^{\sigma,\pi,\nu} - S(x^r) \rangle + \| E^\nu \|^2 \). It follows that:

\[
\langle E^\nu, S^{\sigma,\pi,\nu} - S(x^r) \rangle = \langle E^\nu, -S(x^r) \rangle \leq \| S^r - S(x^r) \| \cdot \| E^\nu \|.
\]

For the lower bound, \( \langle E^\nu, -S(x^r) \rangle \geq 0 \) because both \( E^\nu \) and \(-S(x^r)\) are PSD matrices.

This implies that \( \| S(x^{\ast r}) - S(x) \|^2 = \| S(x^{\ast r}) - S^r \|^2 + 2 \langle E^\nu, S(x^{\ast r}) - S^r \rangle + \| E^\nu \|^2 \), with the bounds as asserted. \( \square \)

**Proof of Lemma 3**

*Proof.* Using Assumptions 1, and 3, and recalling that \( Pr(b \leq B_l|P, W) = G_l(b|P, W) \), we notice that \( G_l(b_l(p, w, \epsilon)|P = p, W = w) = \epsilon \). Under 1, we know that \( \xi \sim U[0, 1] \) and thanks to 3 we also know that \( P, W \perp \xi \), then it follows that \( B_l = G_l^{-1}(\xi|P, W) \) produces a random variable with the same distribution as the budget share of good \( l \) for all \( l = 1, \cdots, L \). We let \( b_l(p, w, \tau) = G_l^{-1}(\tau|P = p, W = w) \), where \( G_l^{-1}(\cdot|P, W) \) is the conditional quantile budget share of good \( l \). \( \square \)