Abstract

We relax the knife-edge equalities required by the FOCs in optimization to propose an approximation of rational consumer choice. We provide an axiomatically-founded measure of the extent to which the FOCs are violated, which is also interpretable in terms of a money-pump multiplier. The framework encompasses measurement errors, information unobserved to the modeler, and bounded rationality (e.g., misunderstanding prices, misperceiving utility tradeoffs, rules of thumb). We develop the testable implications of the model for demand data, and study the properties of a new index of irrationality, the FOC-Departure Index (FDI), which can be applied in all contexts for which the first-order approach is meaningful.
1. Introduction

A central idea behind rationality is the ability to optimize. In a typical differentiable setting, this entails choosing an option that equalizes the marginal benefits and marginal costs along all dimensions. If some such first-order condition were violated, then it would be possible to improve utility by adjusting the choice. Consistency with rationality leaves no room for departure from this delicate balance. One could imagine a more nuanced approach that relaxes the knife-edge equality required by the first-order condition. This would yield an approximation of rationality that is gradated by the extent of departure permitted from the first-order conditions. Such an approach naturally begs the question, how should departure from first-order conditions be measured?

In this paper, we propose such an approach in the classic setting of consumer theory, using an axiomatically founded notion of departure from first-order conditions. Formally, for $\varepsilon \in (0, 1)$ we say that demand data is $\varepsilon$-rationalizable if there exists a utility function $u$, satisfying some regularity properties\(^1\) (which ensure the second-order condition applies), such that whenever we observe the bundle $x$ being chosen at a price vector $p$, we have

$$1 - \varepsilon < \frac{\text{MRS}_{\ell \ell'}(x)}{p\ell/p\ell'} < \frac{1}{1 - \varepsilon},$$

for all $\ell \neq \ell'$. The classic first-order condition, which requires marginal rates of substitution (MRS) and opportunity costs to match, corresponds to $\varepsilon$-rationalizability for all $\varepsilon \in (0, 1)$. The larger $\varepsilon$ is, the more the MRS may depart from opportunity costs, and the more permissive $\varepsilon$-Rationality becomes. The particular way (1) measures departure from the first-order condition for a given pair of goods $\ell, \ell'$ may initially seem ad hoc. To the contrary, we show that it is essentially the unique method of measuring departures between objects such as price vectors and utility gradients, given the economic properties such objects must respect. For instance, notice that (1) is invariant to the units in which goods are measured and priced. We provide further interpretations of our measure of departure as a money-pump multiplier and as a measure of parameter misspecification for additively separable utility functions.

There may be multiple reasons demand data is inconsistent with perfect optimiza-

\(^1\)Specifically, we will require strict monotonicity, strict quasi-concavity and differentiability.
tion. There may be measurement errors when collecting price data, such as unknown promotions, sales, coupons, overhead expenses, or tax implications. Even if prices are accurately recorded, there may be factors that rationally affect choice, but which are not included in the dataset. For instance, tastes may vary a bit based on expiration dates or ripeness of produce, which are observed by the consumer but not the modeler. Bounded rationality may play a role as well. At the time of making the decision, the consumer could have imperfect understanding of the prices, or misperceive her utility tradeoffs. Instead of testing a tentative choice against all affordable alternatives, a consumer may perform only local comparisons. Such limited attention is rational when preferences are convex, and even more sensible than considering distant alternatives when contemplation is costly; but the tâtonnement process may not result in an optimal choice if the consumer’s tradeoffs are misperceived. The consumer could also be subject to framing effects, for instance, seeing another customers’ choices, interacting with a persuasive salesperson, or reacting to store displays and packaging.

Our framework accommodates such data, and encompasses all these interpretations.

The use of first-order conditions is central to many parts of economics, and naturally, the second-order condition is assumed for it to be meaningful, that is, ensure a global optimum. Without assuming quasi-concavity, any invertible demand data would be \( \varepsilon \)-rationalizable (for all \( \varepsilon \)): while the inequalities in (1) do restrict utility gradients given prices, one can always ‘trace’ curves to match these local restrictions with only finitely many observations. It may come as a surprise, therefore, that we show demand data is \( \varepsilon \)-rationalizable if and only if there is a strictly monotone and differentiable utility function \( u \) (not required to be quasi-concave!) such that for each observation \((p, x)\), the bundle \( x \) is \( u \)-maximal in the budget set associated to the price vector \( p^c \), with

\[
1 - \varepsilon < \frac{p^c_\ell / p^c_{\ell'}}{p_\ell / p'_{\ell'}} < \frac{1}{1 - \varepsilon},
\]

for all \( \ell, \ell' \). Here \( p \) represents the price vector collected by the modeler, while \( p^c \) is the price vector used by the consumer. These prices need not coincide: the consumer may use the correct prices while the modeler’s record is faulty; the consumer

\[\text{2 The topic of perception has garnered recent interest in the economic literature; see, among other works, Woodford (2012) and Steiner and Stewart (2016) on perceiving risky prospects, and Esponda (2016) on an equilibrium framework for agents who misperceive their environment.}

\[\text{3 That is, having the feature that the same bundle is never chosen for two different price vectors.}

\[\text{4 Our notion of departure thus applies here to the price vectors (the axiomatic analysis covers any combination of objects such as gradients or price vectors).} \]
may misunderstand the true prices while the modeler’s record is accurate; or some combination of these scenarios may hold. In all cases, if the presumption is that the consumer optimizes correctly, then $\varepsilon$-rationalizability admits a natural interpretation that does not rely on quasi-concavity. This is in the spirit of Afriat (1967)’s result that quasi-concavity is not testable under the rational choice model.

We provide the testable implications of $\varepsilon$-rationalizability, not just for the general class of regular utility functions, but also for the three important subclasses of additively separable, homothetic, and quasi-linear utility functions. We also develop interesting properties of the measure of choice inconsistency to which $\varepsilon$-rationalizability naturally lends itself. Namely, the FOC-Departure Index (FDI) of demand data is defined as the smallest $\varepsilon$ for which the data is $\varepsilon$-rationalizable. We show the FDI is remarkably simple to compute in the case of two-dimensional bundles, which is the setting of many well-known experiments: it is given by the largest FDI computed from pairs of observations, for which a simple, closed-form formula applies.

The FOC-Departure Index measures the smallest departure from the first-order conditions with which the data is consistent. By contrast, the classic Critical Cost Efficiency Index of Afriat (1973) considers global monetary effects, measuring the percentage of income that can be retained while eliminating revealed preference cycles. Despite these differences, we show that there is a surprising relationship between the two measures: the FDI is bounded below by the percent of income lost according to the CCEI (that is, $1 - CCEI \leq FDI$ for all datasets). Phrased differently, small departures from the first-order conditions imply only small budgetary adjustments are needed to eliminate revealed preference cycles, but not vice-versa. This would suggest that our measure is more demanding than that of Afriat, but the story is subtler once power is considered: one should take into account whether violations of rationality are likely for the budget sets observed. Using Bronars (1987)’s well-known approach, for instance, we show that there exist datasets where the FDI suggests greater rationality than does the CCEI, as well as datasets where the opposite is true.

One benefit of the FDI over Afriat’s index is that the FDI applies whenever the first-order approach is meaningful, while applying Afriat’s index requires ‘shrinking the choice set’ to be meaningful. As will be discussed further, it is unclear how to extend Afriat’s index to contexts with non-linear pricing. We easily extend $\varepsilon$-

---

5The idea is to compare the distribution of the index under the true data to the distribution arising if choices were drawn uniformly from budget frontiers.
rationalizability and the FDI to generalized demand data with non-linear pricing, as our approach leaves choice sets unchanged. As opposed to Forges and Minelli (2009), who characterize rational choices with arbitrary budget sets, the feasible set must be convex for the extended notion to be interpretable in terms of departure from global optimality. We show that one can reduce all tests and computations for demand data with non-linear pricing to tests and computations over a fictitious dataset with linear prices, using the marginal prices at observed demands.

The paper is organized as follows. We formalize the framework in Section 2, where we axiomatically characterize our notion of departure from first-order conditions. We also provide further rationales in terms of price mismatch, money pumps in the case of price misperception, and parameter misspecification in the common setting of additively separable utility. In Section 3, we characterize the testable implications for the class of regular utility functions (important subclasses are characterized in the Online Appendix). In Section 4, we develop the FOC-Departure Index and its properties. In Section 5, we consider the implications of our methodology for more specific settings where bundles correspond to lotteries\(^6\) or consumption streams. We extend beyond linear budget sets in Section 6.

2. Consumer Data and \(\varepsilon\)-Rationalizability

We observe a consumer selecting a consumption bundle at various price vectors. The demand data \(\mathcal{D}\) comprises a finite collection of pairs \((p, x)\), where \(p \in \mathbb{R}^L_{++}\) is a price vector and \(x \in \mathbb{R}^L_{++}\) is the consumption bundle demanded at \(p\). Note that a price vector, or a demanded bundle, may potentially appear in \(\mathcal{D}\) multiple times.\(^7\)

The rational benchmark posits that the consumer selects bundles through utility maximization over the budget set, whereby opportunity costs and marginal rates of substitution are equalized. Doing this correctly requires the consumer to have an accurate understanding of both prices and her utility function.

Say that a utility function is regular if it is differentiable, strictly monotone and strictly quasi-concave. In that case, the marginal rates of substitution (MRS) associ-

\(^6\)This is the setting considered in a contemporaneous and independent paper by Echenique, Imai and Saito (2018), who study departures from expected utility. Their measure of departure is defined for risk-averse expected-utility preferences, and coincides for such preferences with the general measure of departure from first-order conditions characterized here (see Section 5.1).

\(^7\)As observed in Chiappori and Rochet (1987), consistency with rationality using a differentiable utility function rules out the same bundle from being chosen at multiple price vectors. This is possible, however, under \(\varepsilon\)-rationalizability.
ated to any two goods is strictly positive and well defined by:

$$MRS_{\ell\ell'}(x) = \frac{\partial u(x) / \partial x_{\ell}}{\partial u(x) / \partial x_{\ell'}},$$

for all $\ell \neq \ell'$, at the bundle $x$. By contrast, the opportunity cost between goods $\ell$ and $\ell'$ is given by the price ratio $p_{\ell} / p_{\ell'}$.

**Definition 1 (\(\varepsilon\)-Rationalizability)** For $\varepsilon \in (0, 1]$, the demand data $D$ is $\varepsilon$-rationalizable if there exists a regular utility function $u$ such that

$$1 - \varepsilon < \frac{MRS_{\ell\ell'}(x)}{p_{\ell} / p_{\ell'}} < \frac{1}{1 - \varepsilon},$$

for each $\ell \neq \ell'$ and each $(p, x) \in D$. Then $u$ is said to $\varepsilon$-rationalize the demand data.

There are several reasons to be interested in this notion of rationalizability, as explained in the following subsections.

### 2.1 How Far Apart Are the Utility Gradient and the Price Vector?

Rationality requires equality between price ratios and marginal rates of substitution, or equivalently, collinearity of price vectors and utility gradients (or $\varepsilon$-Rationalizability for all $\varepsilon \in (0, 1]$). What do we make of data lacking such knife-edge congruence? In this case, one may want to quantify the discrepancy. Consider first some possible ways to measure the distance between two vectors $x, y \in \mathbb{R}_+^2$. Some obvious candidates, such as the angular distance or the Euclidean distance between $x$ and $y$, pose an issue when the vectors are understood to be prices or marginal utilities. Indeed, given their economic interpretation, such variables are not uniquely defined. For instance, an increasing transformation of a utility function offers another representation of the same preference, but rescales utility gradients; and similarly, rescaling prices leaves budget sets unchanged. The Euclidean distance, however, is sensitive to such rescaling. Moreover, the consumer’s problem is unaffected when modifying how any given good’s quantity and price are measured (e.g., using ounces or grams, gallons or quarts, etc.). Both angular and Euclidean distance, however, are sensitive to measurement choice.

Let $\succeq$ be a weak ordering over vector pairs $(x, y) \in \mathbb{R}_+^2$: $(x, y) \succeq (x', y')$ means that “$x$ is farther from $y$ than $x'$ is from $y'$.” The following axioms capture the invariance properties discussed above.
Unit Invariance \((x, y) \succeq (x', y')\) if and only if \((\alpha x, \beta y) \succeq (\alpha' x', \beta' y')\), for all positive scalars \(\alpha, \beta, \alpha', \beta'\).

Measurement Invariance \((x, y) \succeq (x', y')\) if and only if \(((\alpha x_1, x_2), (\alpha y_1, y_2)) \succeq (\alpha' x'_1, x'_2), (\alpha' y'_1, y'_2))\), for all positive scalars \(\alpha, \alpha'\) (and similarly for good 2).

The first axiom reflects the fact that a price vector or a utility gradient is defined only up to a positive linear transformation. It permits different transformations for the different vectors \(x, x', y\) and \(y'\), but requires all dimensions of any given vector to be scaled by the same factor. The second axiom considers a different type of transformation, whereby the same dimension is scaled by the same factor in the pair of vectors compared. It captures invariance to the way in which we measure a good, and thus also how we state its price.

These two invariance properties go a long way in determining the structure of \(\succeq\). Indeed, adding only the following three regularity properties uniquely pins it down.

**Representability** \(\succeq\) is complete, transitive and continuous.

**Symmetry** \((x, y) \sim (y, x)\)

**Monotonicity** \(((\hat{\alpha}, 1), (1, 1)) \succ ((\alpha, 1), (1, 1))\) for all \(1 \leq \alpha < \hat{\alpha}\).

The first axiom ensures existence of a numerical representation. The second means that \(x\) is equally far from \(y\) as \(y\) is from \(x\). The third simply requires that increasing an \(\alpha \geq 1\) brings the vector \((\alpha, 1)\) further from \((1, 1)\).\(^8\)

**Proposition 1** There is a unique ordering \(\succeq^*\) satisfying the five axioms: \((x', y') \succeq^* (x, y)\) if and only if \(\delta(x', y') \geq \delta(x, y)\), where

\[
\delta(a, b) = \max\{\frac{a_1}{a_2}, \frac{b_1}{b_2}\},
\]

for all vectors \(a, b\) in \(\mathbb{R}^2_{++}\).

The representation in Proposition 1 compares how far apart two vectors are, in comparison to how far apart another two reference vectors are. The choice of the reference vectors simply parametrizes a bound. To fix ideas, one can measure how

\(^8\)Naturally, one would also desire the opposite relationship in the case \(\alpha \in (0, 1)\). Imposing this property would be redundant, however: the ordering we uncover in Proposition 1 satisfies it.
far $x$ is from $y$ relative to how far $(1, 1)$ is from $(1 - \varepsilon, 1)$ for $\varepsilon \in (0, 1)$, in which case:

$$\delta((1 - \varepsilon, 1), (1, 1)) = \max\{1 - \varepsilon, \frac{1}{1 - \varepsilon}\} = \frac{1}{1 - \varepsilon}.$$ 

Hence $((1 - \varepsilon, 1), (1, 1)) \succeq (x, y)$ if and only if

$$\delta(x, y) = \max\left\{\frac{x_1}{x_2}, \frac{y_1}{y_2}\right\} < \frac{1}{1 - \varepsilon},$$

which is equivalent to how Definition 2 of $\varepsilon$-rationalizability measures how far apart are $x$ and $y$, where one is the utility gradient and the other is the price vector:

$$1 - \varepsilon < \frac{x_1}{x_2} < \frac{1}{1 - \varepsilon}.$$ 

To extend beyond two goods, one can for instance consider the projections of these vectors to each pairs of goods, and take the largest distance:

$$\delta(x, y) = \max_{\ell, \ell'} \delta((x_\ell, x_{\ell'}), (y_\ell, y_{\ell'})),$$

for any two $x, y \in \mathbb{R}^L_{++}$. Our axiomatic characterization suggests this is a natural way to quantify the extent to which the consumer deviates from the optimal tradeoffs in her first-order condition, while remaining agnostic about the source of this deviation. Note also that the axioms equally apply whether we use the measure to evaluate how far apart are two price vectors, how far apart are two utility gradients, or how far apart is a price vector from a utility gradient.

### 2.2 Misperceived Prices

As we have discussed, the prices $p$ on record and the prices $p^c$ the consumer takes into account need not coincide. The consumer may use the correct prices while the modeler’s record is faulty; the consumer may misunderstand the true prices while the modeler’s record is accurate; or some combination of these scenarios may hold. As the next proposition shows, $\varepsilon$-rationalizability is equivalent to the consumer picking the optimal bundles from what she understands the budget set to be, under the requirement that her prices are not too far from the modeler’s, using the same $\delta$-criterion uncovered axiomatically in the previous subsection.
Proposition 2 \( \mathcal{D} \) is \( \varepsilon \)-rationalizable if and only if there exists a strictly monotone and differentiable utility function \( v \) such that one can associate to all \( (p,x) \in \mathcal{D} \) a price vector \( p^c \in \mathbb{R}^L_+ \) used by the consumer with

\[
x \in \arg \max_{\{p^c \cdot y \leq p^c \cdot x\}} v(y), \text{ and}
\]

\[
1 - \varepsilon < \frac{p_{\ell}/p_{\ell}^c}{p_{\ell}^c/p_{\ell}^c} < \frac{1}{1 - \varepsilon},
\]

for all \( \ell \neq \ell' \).

Necessity is rather straightforward. Indeed, by \( \varepsilon \)-rationalizability, there exists a regular utility function \( u \) whose gradient is never too far apart from the price vectors for observations in \( \mathcal{D} \). The result then follows at once by taking these gradients as the consumer’s price vectors. Sufficiency would be rather straightforward too if \( v \) was required to be regular, but notice that Proposition 2 does not require \( v \) to be quasi-concave. This makes the result more interesting, but also harder to prove (see the Appendix). In particular, it means that convexity of the consumer’s preference has no empirical content in the context of price mismatch. So, while the definition of \( \varepsilon \)-rationalizability relies on regular utility functions, the foundation for it in terms of price mismatch does not. This observation becomes relevant when comparing the FOC-Departure Index, introduce in Section 4, to Afriat’s CCEI.

Under an interpretation of price misperception by the consumer, Proposition 2 captures a notion of bounded rationality that is related to Gabaix (2014), but somewhat different in spirit. The consumer picks the optimal bundle in her perceived budget set in both cases. However, Gabaix proposes an interesting theory of how perceived prices arise as a function of an exogenous default price vector and the actual price vector, while the theory arising in Proposition 2 only requires the perceived price vector to be in the vicinity of the true price vector.

Beyond the axiomatic foundation in Section 2.1, Equation (4) admits an additional interpretation when the mismatch between prices arises from the consumer’s misperception. The discrepancy between \( p \) and \( p^c \) means the consumer can be subject to a money pump. What profit can a rational person with \( \$M \) make conducting the
following trade scheme once? He starts by using his $M$ money to buy any bundle he wants from the consumer at her perceived prices $p^c$, then trades that bundle in any way he wants on the market given the true prices $p$, and finally resells whatever goods he acquired this way back to the consumer at her perceived prices $p^c$. To be clear, this scheme is conducted before the consumer decides on her consumption plan; she is not yet maximizing her preference, only accepting trades that give her a higher perceived budget for doing so, which is always desirable. For simplicity, our tie-breaking rule is that the consumer accepts trades which leave her budget unchanged. To summarize, the third party will maximize $p^c \cdot y$ over all bundles $y$ such that $p \cdot y \leq p \cdot x$, for some bundle $x$ such that $p^c \cdot x \leq M$.

To solve this optimization problem, notice that the solution will have the bundle $x$ maximize $p \cdot x$ over the set of bundles $x$ such that $p^c \cdot x \leq M$. Indeed, making $p \cdot x$ larger increases the set of bundles $y$ such that $p \cdot y \leq p \cdot x$. Because the objective function $p \cdot x$ is linear in $x$, an optimal solution to this problem is to spend the $M$ on a good $\ell$ with the highest price ratio when comparing true prices to perceived prices: $p_\ell / p^c_\ell \geq p_k / p^c_k$ for each good $k$ (the computation is analogous to that when maximizing perfect-substitutes preferences). It remains to find a $y$ that maximizes $p^c \cdot y$ under the constraint $p \cdot y \leq (p_\ell M) / p^c_\ell$. Similar reasoning reveals that an optimal solution is to spend the $(p_\ell M) / p^c_\ell$ on a good $\ell'$ with the highest price ratio when comparing perceived prices to true prices: $p^c_\ell / p^e_\ell \geq p^c_k / p_k$ for each good $k$. The profit in that case is $(p^c_\ell / p^e_\ell)(p_\ell M) / p^c_\ell$. To summarize, the third party’s maximal profit is:

$$\max_{\ell, \ell'} \frac{p_\ell / p^e_\ell}{p^c_\ell / p^e_\ell} M = \delta(p, p^c) M.$$ 

Thus, using $\delta$ to measure how far apart the true price vector is from the perceived one also determines the money pump multiplier for a rational party conducting the simple trading scheme described above. In this view, (4) amounts to placing an upper bound $(1 - \frac{1}{\epsilon})$ on this money pump multiplier.

---

9The consumer’s endowment in goods is assumed to be large enough, or $M$ is assumed to be small enough that the consumer can provides the good that the third party wants to buy.

10If she accepts only trades that strictly increase her budget, then the rational schemer can get as close as desired from the optimal profit calculated in the next paragraph, by leaving a little bit of surplus to the consumer in both trades involved in the money pump scheme.
2.3 Misperceived Tastes

Suppose instead that the consumer is subject to errors in assessing her utility tradeoffs, or that the modeler’s data is missing contextual information that affects utility tradeoffs. As will be clear, contrary to Section 2.2, restricting attention to convex preferences is important for this interpretation of \( \varepsilon \)-Rationalizability.

How do consumers explore their budget sets? A rational consumer might contemplate all bundles at once, to find the best one every time she has to make a choice. Alternatively, to save on contemplation and thinking costs, she may test any given bundle by checking that there is no preferable alternative in its vicinity, thereby reaching her choice – an affordable bundle that has no better alternative in its vicinity – by tatonement. Reaching choices through a series of small adjustments might be a better description of a consumer’s thought process in some circumstances. It is unlikely that every time some prices change at the grocery store, a consumer re-assesses the entire set of affordable bundles, carefully introspects about her preference relation over those bundles, and directly selects the bundle maximizing it globally. With convex preferences, rational choices would in fact arise from thinking locally and in steps, as above. If the consumer is boundedly rational, on the other hand, she may make mistakes when assessing local utility tradeoffs by introspection, using a utility gradient that is only close to the correct one. Alternatively, she may assess her tradeoffs accurately, but these tradeoffs vary to some extent with factors unobserved by the modeler (e.g., How ripe are the fruits? How far ahead is the expiration date?).

Using the notion uncovered in Section 2.1 to bound the distance between perceived and true utility gradients, or between the actual and modeled utility gradients, again leads to \( \varepsilon \)-rationalizability.

One further justification for \( \varepsilon \)-rationalizability is relevant whether the misperceived tastes arise from the consumer’s or the modeler’s mistake. It is common in applications to restrict attention to subclasses of utility functions, such as Cobb-Douglas preferences. If the consumer’s true preference is captured by the utility function \( \prod x_t^{\alpha_t} \), then a natural way to capture misperceived tastes or parameter mis-specification on the part of the modeler is that the consumer maximizes \( \prod x_t^{\beta_t} \) where the vector of exponents \( \beta \) may vary, but cannot be too far apart from \( \alpha \). This operationalizes, in a different setting, Rubinstein and Salant (2012)’s notion of a decision maker who uses only preferences that are ‘close’ to her true one. As the next proposition shows, \( \varepsilon \)-Rationalizability is equivalent to this approach, not only for the
Cobb-Douglas model, but also for any additively separable preference. Formally, say $u$ is additively separable if $u(x) = \sum_{\ell} u_{\ell}(x_{\ell})$ for some strictly concave, differentiable and strictly monotone utility functions $u_{\ell} : \mathbb{R}_{+} \to \mathbb{R}$.

**Proposition 3**  For each $\ell = 1, \ldots, L$, let $u_{\ell}$ be a utility function over good $\ell$. Then $\mathcal{D}$ is $\varepsilon$-rationalizable with respect to the additively separable utility $u(\cdot) = \sum_{\ell=1}^{L} u_{\ell}(\cdot)$ if and only if there is $\beta : \mathcal{D} \to \mathbb{R}_{++}$ such that for all $(p, x) \in \mathcal{D}$,

\begin{equation}
(5a) \quad x = \arg \max_{\{p' y \leq p x\}} \sum_{\ell=1}^{L} \beta_{\ell}(p, x) u_{\ell}(x_{\ell}), \text{ and}
\end{equation}

\begin{equation}
(5b) \quad 1 - \varepsilon < \frac{\beta_{\ell}(p, x)}{\beta_{\ell'}(p, x)} < \frac{1}{1 - \varepsilon}, \text{ for all } \ell, \ell' \in \{1, \ldots, L\}.
\end{equation}

The inequalities in (5b) simply state that the vector $\beta$ of modified coefficients is not too far from the original unit vector of coefficients associated to $u$, using the notion uncovered in Section 2.1. In the Cobb-Douglas example, $u_{\ell}(\cdot) = \alpha_{\ell} \log(\cdot)$ for each $\ell$. For intertemporal choices, with $x_{\ell}$ representing the amount of good $\ell$ in time-period $\ell$, exponential discounting corresponds to the case $u_{\ell}(\cdot) = \delta^{\ell} \tilde{u}(\cdot)$ for some time-independent utility function $\tilde{u}$ (i.e., independent of $\ell$). In that context, Proposition 3 implies all the consumer’s errors can be attributed to misperception of the discounting function. Similarly, in a setting with risk, where each $\ell$ is a state of the world, all errors could be attributed to misperception of probabilities.

3. Testable Implications

The seminal work of Afriat (1967) shows how the generalized axiom of revealed preference (GARP) captures the empirical content of rational choice. Formally, demand data is consistent with the maximization of some strongly monotone, continuous utility function if and only it satisfies GARP. Given the widespread use of regularity in applications, subsequent work extended Afriat’s approach in that direction, showing that only a slightly stronger requirement than GARP arises.\(^{12}\)

\(^{11}\)Strict concavity may seem much stronger than our usual requirement of strict quasi-concavity. However, they are almost the same in this additive setting: in a classic result which builds on Arrow’s earlier observation, Debreu and Koopmans (1982) show that quasi-concavity of a continuous, additively separable utility function implies that all but one $u_{\ell}$’s must be concave, and the last must have features of concavity too.

\(^{12}\)Strict quasi-concavity implies a single-valued demand function. In particular, a chosen bundle is revealed strictly preferred to all other affordable alternatives, so the data satisfies SARP, not just
We now show how to build on these approaches to capture the empirical content of $\varepsilon$-Rationalizability. To do this, we apply the methodology of de Clippel and Rozen (2018). The first step is to assume the consumer follows the theory, and identify necessary restrictions that the demand data reveals about her preference. While the consumer’s preference is defined over the entire space of goods $\mathbb{R}_+^L$, it turns out that the only relevant restrictions apply to her preferences over the subset of bundles $X = \{x \in \mathbb{R}_+^L \mid (p, x) \in \mathcal{D} \text{ for some } p \in \mathbb{R}_+^L \}$ that have been observed chosen.

Suppose the DM’s preference over bundles is represented by the regular utility function $u$. For any observation $(p, x) \in \mathcal{D}$, $\varepsilon$-Rationalizability requires:

$$\frac{p_\ell}{p_{\ell'}} (1 - \varepsilon) < MRS^u_{\ell\ell'}(x) < \frac{p_\ell}{p_{\ell'}} \left(\frac{1}{1 - \varepsilon}\right), \quad \forall \ell \neq \ell'. \tag{6}$$

The linear inequalities in (6) define a convex cone $C_\varepsilon(p, x) \subseteq \mathbb{R}_+^L$ to which the gradient of $u$ at $x$ must belong. We must keep in mind, however, that the bundle $x$ is potentially demanded under multiple price vectors. Let $P(x) = \{p' \in \mathbb{R}_+^L \mid (p', x) \in \mathcal{D}\}$ be the set of price vectors at which $x$ is demanded. To capture the bounds imposed jointly by the data on the gradient of the utility function at $x$, we must define the cone:

$$C_\varepsilon(x) = \bigcap_{p' \in P(x)} C_\varepsilon(p', x). \tag{7}$$

Suppose we conjecture that the gradient of $u$ at $x$ is given by the vector $v \in C_\varepsilon(x)$. As $u$ is strictly quasi-concave, $u(x) > u(x')$ for any bundle $x' \neq x$ such that $v \cdot x' \leq v \cdot x$. In particular, $x$ is strictly preferred to any such $x'$ that was also observed chosen. Formally, $x$ is strictly preferred to all bundles in the set:

$$\Gamma(x, v) = \{x' \in X \mid x' \neq x, \ v \cdot x' \leq v \cdot x\}. \tag{8}$$

The difficulty, of course, is that we are unsure which $v \in C_\varepsilon(x)$ is the true gradient of $u$ at $x$. We only know that there exists $v \in C_\varepsilon(x)$ such that $x$ is strictly preferred to all the elements in $\Gamma(x, v)$. In the language of de Clippel and Rozen (2018), this corresponds to a lower-contour set restriction on the consumer’s preference $\succ$ over GARP. Requiring differentiability adds an ‘invertibility’ requirement that the same bundle can’t be chosen from different budget sets; see Chiappori and Rochet (1987) and Matzkin and Richter (1991).
the set $X$ of demanded bundles, as some element of \{\(\Gamma(x, v) \mid v \in C_\varepsilon(x)\)\} must be contained in the \(\succ\)-lower contour set of \(x\). There is one such lower-contour set restriction for each \(x \in X\), generating a collection of lower-contour set restrictions \(\mathcal{R}_\varepsilon(\mathcal{D})\) over the consumer’s possible preference over demanded bundles. Thus, we have shown that \(\varepsilon\)-Rationalizability of the demand data \(\mathcal{D}\) requires the collection of restrictions \(\mathcal{R}_\varepsilon(\mathcal{D})\) to be acyclically satisfiable: there must exist a (strict) acyclic relation on \(X\) that simultaneously satisfies them. For instance, the consumer’s true utility function \(u : \mathbb{R}^L_+ \rightarrow \mathbb{R}\), restricted to \(X\), defines such an acyclic relation.

We have not yet shown, however, that acyclic satisfiability of \(\mathcal{R}_\varepsilon(\mathcal{D})\) implies the existence of a regular utility function \(u : \mathbb{R}^L_+ \rightarrow \mathbb{R}\) that \(\varepsilon\)-rationalizes the data. Our first main result confirms this.

**Proposition 4** \(\mathcal{D}\) is \(\varepsilon\)-rationalizable if and only if \(\mathcal{R}_\varepsilon(\mathcal{D})\) is acyclically satisfiable.

Proposition 4 shows that testing whether demand data \(\mathcal{D}\) is \(\varepsilon\)-rationalizable amounts to checking whether the collection of restrictions \(\mathcal{R}_\varepsilon(\mathcal{D})\) is acyclically satisfiable. How does one check this? de Clippel and Rozen (2018) shows that, in the presence of lower-contour set restrictions (but usually not more generally), one checks acyclic satisfiability in the same way as SARP, simply by iteratively looking for candidate minimal elements of the unknown preference ordering induced by \(u\) on \(X\). The key step is this: having determined thus far that the elements in \(S \subset X\) can be ranked at the bottom of the ordering, are there elements of \(X \setminus S\) that qualify as a candidate-minimal element within \(X \setminus S\)? An option \(x \in X \setminus S\) qualifies as a candidate minimal element if and only if no remaining elements need to be ranked below \(x\): that is, there exists \(v \in C_\varepsilon(x)\) such that \(\Gamma(x, v) \cap (X \setminus S) = \emptyset\).\(^{13}\) Acyclic satisfiability holds if and only if we can identify a candidate minimal element in each step of the iterative procedure, thus constructing a complete ordering over the entire set of chosen bundles \(X\). Importantly, the success or failure of this procedure is path independent: it does not depend on which candidate minimal element one picks in a step when multiple alternatives qualify.

Our testing methodology for \(\varepsilon\)-rationalizability rests on the fact that the possible gradients are restricted to a set (defined by the model). The approach is portable.

\(^{13}\)Developing this, \(x\) qualifies if and only if there exists \(v\), representing a utility gradient, such that the corresponding MRS satisfies the bounds in (6) for all \(p\) such that \((p, x) \in \mathcal{D}\), and \(v \cdot (x' - x) > 0\) for all \(x' \in X \setminus S\) other than \(x\), which is easy to determine by linear programming (solvable in polynomial time in \(|\mathcal{D}| + L\)).
to other situations where this feature arises. For instance, suppose the modeler is unsure of the consumer’s perceived price vector $p^c$, but theorizes that for each good $\ell$, it is some mixture of a default price $p^d_\ell$ and the true price: $p^c_\ell = p^d_\ell + m_\ell(p_\ell - p^c_\ell)$, where $m_\ell \in [0, 1]$ is the extent to which the consumer shifts the perceived price of the good from the default price to the actual one. This is a weaker restriction on perception than Gabaix (2014)’s sparse-max theory, which determines the vector $m = (m_1, \ldots, m_L)$ endogenously through a sparsity-based optimization using a concave utility function.\textsuperscript{14} To test the generalization where any $m \in [0, 1]^L$ is conceivable, which leads to predictions that are less precise than Gabaix’s but robust to a variety of theories of subjective price formation, one simply replaces the cone $C_\varepsilon(x)$ with the convex set of possible gradients $C(p, p^d) = \{(p^d_\ell + m_\ell[p_\ell - p^c_\ell])_{\ell=1}^L | m \in [0, 1]^L\}$, and proceeds with the analogous lower-contour set restrictions.

4. FOC-Departure Index

Demand data is either consistent with rationality, or not. In the imperfect world of actual data, it is more useful to have ways to quantify the degree to which the data comply with the theory. $\varepsilon$-Rationalizability naturally lends itself to such measurements. The FOC-Departure Index (FDI) is the infimum over all $\varepsilon$ such that the data is $\varepsilon$-rationalizable. For instance, the FDI of demand data arising from the maximization of a regular preference is zero.

This provides an alternative measure of goodness-of-fit to Afriat (1973)’s Critical Cost Efficiency Index (CCEI), which computes the largest percentage of the consumer’s budgets that can be retained while still eliminating all revealed-preference cycles (see also Varian (1990)). To formalize Afriat’s index, define for $\sigma \in [0, 1]$ a strict revealed preference $x \succ_A,\sigma y$ if $(p, x) \in \mathcal{D}$ and $\sigma p \cdot x > p \cdot y$; and a weak revealed preference $x \succeq_A,\sigma y$ if $\mathcal{D}$ contains a sequence of observations $(p^1, x^1), \ldots, (p^n, x^n)$ where $x^1 = x$, $x^n = y$, and for each $i \in \{1, \ldots, n - 1\}$, either $x^i = x^{i+1}$ or $\sigma p^i \cdot x^i \geq p^i \cdot x^{i+1}$. Then the CCEI is the supremum of $\sigma \in [0, 1]$ such that $x \succeq_A,\sigma y$ implies not $y \succ_A,\sigma x$ for all observed choices $x, y$.\textsuperscript{15} While both the CCEI and FDI yield values between

\textsuperscript{14}His theory captures a decision maker who, for instance, realizes that spending time understanding the state of the Amazonian forest or interest rates in some distant country will be costly and yet have very little effect on her decision. The decision maker knows default values of such variables (e.g., long-run averages) and optimally allocates effort in determining what price estimate, somewhere between the default value and the true value, to use.

\textsuperscript{15}Varian (1990) generalizes Afriat’s index to allow the (proportional) budget adjustment to vary
zero and one, the directions are reversed: the CCEI measures rationality, and is 1 in that case, while the FDI measures departure from rationality, and is 0 when choices are rational. However, one can directly compare the FDI to Afriat’s inefficiency measure $1 - \text{CCEI}$, which is the minimal factor by which the consumer’s budgets must be shrunk to eliminate revealed-preference cycles.

We start in Section 4.1 by illustrating the FDI in the simpler case of two commodities (as in most experiments on the subject). In addition to providing further intuition, these results will prove useful throughout the section. As a by-product, they will provide a simple formula to compute the FDI of demand data with two commodities.\(^{16}\) Section 4.2 shows, perhaps surprisingly, that Afriat’s inefficiency measure is always smaller or equal to the FDI (independently of the number of commodities). Yet, power should be factored in when assessing how successful rationality is at explaining observed choices. We show in Section 4.3, by means of examples, that no systematic relation exists in that case: demand data may appear closer to rationality under the CCEI than under the FDI, and vice versa.

### 4.1 The Case of Two Commodities

We assume $L = 2$ throughout this subsection. As a start, consider demand data $\mathcal{D} = \{(p, x), (p', x')\}$ comprising only two observations. Whenever $p \cdot x' \leq p \cdot x$ and $p' \cdot x \leq p' \cdot x'$, there is either a violation of SARP (the case $x \neq x'$) or a violation of invertibility (the case $x = x'$). Otherwise, the data is rationalizable by a regular utility function. The next result provides simple formulas for computing the FDI of $\mathcal{D}$ in each case.

---

\(^{16}\)While we haven’t found a simple formula for the FDI with three commodities or more, it takes only $n$ applications of the enumeration procedure to identify the index with $\pm \frac{1}{2^n}$ precision. First, test for $1/2$-rationalizability; then test $1/4$-rationalizability if the previous test succeeds, and $3/4$-rationalizability otherwise; and continue this recursively $n - 2$ more times.
Figure 1: The construction of $o(x, x')$ for proving Proposition 5.

**Proposition 5**  
Take $\mathcal{D} = \{(p, x), (p', x')\}$, and assume $p'_1/p'_2 \geq p_1/p_2$ without loss of generality. Then,

$$FDI(\mathcal{D}) = \begin{cases} 
\min\{\frac{p(x-x')}{p_2(x_2-x'_2)}, \frac{p'(x'-x)}{p'_1(x'_1-x_1)}\} & \text{if } \mathcal{D} \text{ violates SARP}; \\
1 - \frac{1}{\sqrt{\delta(p, p')}} & \text{if } \mathcal{D} \text{ violates invertibility}; \\
0 & \text{otherwise}.
\end{cases}$$

The proof of Proposition 5 reveals what underlies these expressions. In particular, the FDI directly relates to how far apart each of the two price vectors is from the vector $o(x, x')$ that is orthogonal to the line passing through $x$ and $x'$ (see Figure ). Indeed, the FDI is the minimum between $f^{-1}(\delta(p, o(x, x')))$ and $f^{-1}(\delta(p', o(x, x')))$, where $f$ is the function that associates to each $\varepsilon \in (0, 1)$ the fraction $1/(1 - \varepsilon)$ appearing in the definition of $\varepsilon$-rationalizability. Of course, $o(x, x')$ is ill-defined when the issue is a lack of invertibility ($x = x'$). In that case, the FDI is simply a function of how far apart from each other are the two price vectors leading to the same choice. The problem amounts to ascertaining that $C_\varepsilon(x) \neq \emptyset$. It is not difficult to check that a vector $v$ minimizes $\max\{\delta(p, v), \delta(p', v)\}$ if and only if $v_1/v_2$ is the geometric mean of $p_1/p_2$ and $p'_1/p'_2$, in which case $\max\{\delta(p, v), \delta(p', v)\} = \sqrt{\frac{p'_1/p'_2}{p_1/p_2}} = \sqrt{\delta(p, p')}$. The FDI is then the image under $f^{-1}$ of this number.
The next result shows that, quite remarkably, the FDI of demand data with two commodities is simply the maximum of the FDI over all pairs of observations. Combined with Proposition 5, we have thus found a formula to compute the FDI.

**Proposition 6** For \( L = 2 \), the FDI of any demand data \( D \) is the maximum of the FDI’s (computed in Proposition 5) associated to each pair of observations in \( D \).

As seen from the proof of Proposition 4, demand data is \( \varepsilon \)-rationalizable if and only if for every \( x \in X \), there is \( v(x) \in C_\varepsilon(x) \) such that the auxiliary demand data \( D' = (v(x), x)_{x \in X} \) is rationalizable (in the classic sense) by a regular utility function \( u \). That, in turn, is equivalent to \( D' \) satisfying SARP and the invertibility property. One may conjecture then that Proposition 6 follows at once from Rose (1958).\(^{17}\) That would be the case if we knew that for every \( x \in X \), there is \( v(x) \in C_\varepsilon(x) \) such that for all \( y, z \in X \) the auxiliary demand data \( \{(v(y), y), (v(z), z)\} \) satisfies SARP. However, multiple SARP violations may occur when pairing a same observation \( (p, x) \) with different other observations, and one may have to consider different vectors \( v(x) \) to achieve \( \varepsilon \)-rationalizability on those pairs.\(^{18}\) This makes the proof significantly harder.

### 4.2 A Surprising Relation to the CCEI

From now on, we return to the general case with no restriction on the number of commodities. As the CCEI and FDI pertain to very different variables – adjusting incomes for the CCEI versus adjusting tradeoffs for the FDI – one might think these measures are unrelated. However, we show that there is a clear relationship between the two: a small error in perception of MRS or prices means only small budgetary adjustments are needed to eliminate revealed preference cycles.

**Proposition 7** For any demand data \( D \), \( 1 - \text{CCEI}(D) \leq \text{FDI}(D) \).

**Proof.** Fix \( \varepsilon > 0 \). Proposition 4 shows \( D \) is \( \varepsilon \)-rationalizable if and only if \( R_\varepsilon(D) \) is acyclically satisfiable. As seen from the proof of that result, acyclic satisfiability of

---

\(^{17}\)Rose observes that, with only two commodities, satisfying SARP over pairs of observations guarantees satisfying SARP over sequences of observations of any length.

\(^{18}\)For instance, suppose \( p = (1, 1) \), \( p' = (1, 2) \), \( p'' = (2, 1) \), \( x = (10, 10) \), \( x' = (1, 18) \), and \( x'' = (18, 1) \). Both \( \{(p, x), (p', x')\} \) and \( \{(p, x), (p'', x'')\} \) are \( \varepsilon \)-rationalizable with \( \varepsilon = 1/7 \). However, the \( p' \) budget line cannot be tilted to eliminate the revealed preference of \( x' \) over \( x \), while keeping \( \delta(p', v(x')) \below 1/(1-\varepsilon) \). Instead, one must make the \( p \)-budget line flatter (e.g. taking \( v(x) = (5, 6) \)) guarantees WARP while keeping \( \delta(p, v(x)) < 1/(1 - \varepsilon) \)). Thus, for \( \{(v(x), x), (v(x'), x')\} \) to satisfy WARP, \( v(x) \) must put significantly more weight on good 2 than good 1. A similar argument for \( \{(p, x), (p'', x'')\} \) implies that a different \( v(x) \), one that puts significantly more weight on good 1 than good 2, must be considered to get \( \{(v(x), x), (v(x''), x'')\} \) to satisfy WARP.
$R_\varepsilon(D)$ implies that for every $x \in X$, there is $v(x) \in C_\varepsilon(x)$ such that the auxiliary demand data $D' = (v(x), x)_{x \in X}$ is rationalizable (in the classic sense) by a regular utility function $u$. We will use $u$ to show that Afriat’s CCEI for the original data $D$ is at most $1 - \varepsilon$. This implies at once that the CCEI is lower or equal to the FDI.

Take any $(p, x) \in D$, and consider the indifference curve of $u$ passing through $x$. As illustrated in Figure 2(a), the budget set may not be tangent to the indifference curve at $x$, but the line determined by the vector $v(x)$ is tangent, given the rationalizability of $D'$. We claim that $\frac{p \cdot y}{p \cdot x} \geq 1 - \varepsilon$ for any bundle $y$ above this tangent line, i.e., any $y$ such that $v(x) \cdot y \geq v(x) \cdot x$. This holds trivially if $p = v(x)$, so suppose they are different and consider the optimization problem

$$\min_{\{y \mid v(x) \cdot y \geq v(x) \cdot x\}} p \cdot y.$$  

The constraint must bind at the optimum, else the objective could be further reduced. Also, as seen in Figure 2(a), linearity of the objective and constraint imply the solution must occur at a bundle $y$ with only one positive component: that is, there is $\ell$ such that $y_\ell = \frac{v(x) \cdot x}{v(x)}$ and $y_i = 0$ for all $i \neq \ell$. Using the fact that $v(x) \in C_\varepsilon(x)$, the minimal expenditure satisfies:

$$p \cdot y = p \cdot y_\ell = p_\ell \sum_{i=1}^L \frac{v_i(x)}{v(x)} x_i \geq (1 - \varepsilon) p_\ell \sum_{i=1}^L \frac{p_i}{p_\ell} x_i = (1 - \varepsilon) p \cdot x.$$  

By quasi-concavity of $u$, any bundle $z$ with $u(z) \geq u(x)$ must satisfy $v(x) \cdot z \geq v(x) \cdot x$. Hence, the above inequality shows that if $(1 - \varepsilon)$-percent of income is retained, the choice $x$ from the original budget set is strictly preferred under $u$ to all bundles in the remaining budget set. To finish the proof that the CCEI is at most $1 - \varepsilon$, observe that any cycle in $\succeq_{A, 1-\varepsilon}$ would imply a cycle in the corresponding utilities generated by $u$, which is impossible. Q.E.D.

As a corollary, demand data that is $\varepsilon$-rationalizable satisfies Halevy et al (2018)’s $v$-GARP, where $v = (1 - \varepsilon, \ldots, 1 - \varepsilon)$, ensuring that the existence of a locally non-satiated (even concave) utility function respecting revealed preference comparisons that occur even when income is scaled down by a factor $1 - \varepsilon$.

We now show through examples that the inequality in Proposition 7 can hold strictly for some datasets, and with equality for others.
Example 1  Consider demand data with two observations: the bundle $x = (a, b)$ is chosen at the price vector $p = (1, \pi)$, and the bundle $x' = (b, a)$ is chosen at the price vector $p' = (\pi, 1)$, where $\pi > 1$. This situation is illustrated in Figure 2(b). By Proposition 5, the FDI is $1 - \frac{1}{\pi}$ independently of the values $a, b$. Indeed, for any $x$ and $x'$ which are oppositely placed at the frontier of their respective budget sets, these choices correspond to the same sizable error in local perception of MRS. By contrast, Afriat’s CCEI becomes arbitrarily close to one (as thus the inefficiency is arbitrarily close to zero) as $x, x'$ get arbitrarily close to the intersection of the two budget sets. There also exists demand data where the two measures agree. To see this, shift $x$ to the vertical axis and $x'$ to the horizontal axis. Since $p \cdot x = p' \cdot x'$ and $x' = (p' \cdot x'/\pi, 0)$, $1 - \text{CCEI} = 1 - \frac{p' \cdot x'}{p \cdot x} = 1 - \frac{p' \cdot x'/\pi}{p \cdot x} = 1 - \frac{1}{\pi}$.

4.3 No Relation Once Taking Power Into Account

Whether consistency with Rationality is remarkable clearly depends, at least to some extent, on the combination of budget sets being tested. For instance, Rationality is impossible to refute when all budget sets are related by inclusion. By contrast, a pair of budget sets like in Figure 1(b) opens the possibility of a WARP violation. In that sense, the specific value of the CCEI or the FDI derived from a consumer’s actual choice is not that informative without being contrasted against the distribution of those indices arising under some alternative behavioral hypothesis.
While different criteria have been proposed over the years to capture power, we focus here on the approach that is most often applied in experimental papers. This method, suggested by Bronars (1987) and inspired by Becker (1962), proposes to use as a reference point the distribution of CCEI’s arising from a random collection of choices, under the assumption that each bundle on the frontier of a budget set is equally likely. Most experimental papers argue that the rational choice model captures observed choices rather well, because the distribution of CCEIs arising from the data is a significant FOSD shift towards lower values of Afriat’s index.

Of course, the approach can be replicated using the FDI instead. Interestingly, while we know from Proposition 7 that $1 - \text{CCEI} \leq \text{FDI}$, this does not mean that subjects will necessarily appear as less rational when applying Bronars’ methodology to the FDI. Indeed, the distribution of FDI’s for the randomly generated demand data used as the reference point will itself shift towards higher values. Here is a simple theoretical example to illustrate this point.

**Example 2** We start by revisiting the example depicted in Figure 2(b). Remember that the FDI of $\{(p, x), (p', x')\}$ is $1 - \frac{1}{\pi}$ (see Example 1). We now prove that this is the largest FDI that can be reached with pairs of choices on the boundary of these two budget sets. To see this, let $y, y'$ be two distinct bundles such that $y$ is on the budget line associated to $p$, $y'$ is on the budget line associated to $p'$, and there is a SARP violation: $p \cdot y' \leq p \cdot y = p \cdot x$ and $p' \cdot y \leq p' \cdot y' = p' \cdot x'$. By Proposition 5, the FDI is

$$1 - \frac{1}{\pi} \max\{X, \frac{1}{X}\}, \text{ with } X = \frac{y_1' - y_1}{y_2 - y_2'}.$$

Hence indeed the largest FDI is reached with $X = 1$. Notice that this occurs with probability zero when drawing uniformly from bundles on the boundary of the two budget sets. Hence, the probability that the FDI of a randomly drawn demand data is larger or equal to that of the demand data in Example 1 is zero. By contrast, there is a strictly positive probability that the Afriat inefficiency of a randomly drawn demand data will be bigger than that of the demand data from the example.

For an example where the comparison is opposite, consider the demand data $D = \{(p, x), (p', z)\}$ where $p, p', x$ are as in Example 1 and $z$ amounts to spending the budget under $p'$ entirely on good 1, i.e., $z = (\frac{p' \cdot x'}{p_1'}, 0)$. Notice that the Afriat inefficiency of

\[19\] So far we have assumed that the demand data contains only strictly positive bundles. So instead one can assume that the consumer spends a positive, but very small amount on good 1 in $z$. Since
demand data picked on the budget lines associated to $p$ and $p'$ is smaller or equal to that of $D$ if and only if the bundle picked on the $p$-budget line is strictly to the left of $x$ and the bundle picked on the $p'$-budget line is strictly to the right of $x'$. It is easy to check that any such demand data also has a FDI larger than that of $D$, but also that there is a positive mass of other bundle combinations leading to a larger FDI than that of $D$. Thus, this time, the probability that the FDI of a randomly drawn demand data is larger than that of $D$, is larger than the probability that the Afriat inefficiency of a randomly drawn demand data is larger than that of $D$.

5. Further Restrictions on Preferences

In many contexts, one is willing to impose further properties on preferences. Classic examples in consumer theory include quasi-linearity, homotheticity and additive separability. Expected utility is a usual assumption for choice under risk, as is exponential or hyperbolic discounting for time preferences.

Let $\mathcal{P}$ be a set of properties imposed on utility functions. Demand data is $\varepsilon$-Rationalizable for $\mathcal{P}$ if it satisfies Definition 1 with the added requirement that $u$ satisfies the properties listed in $\mathcal{P}$. Of course, the resulting FDI for $\mathcal{P}$ weakly increases as one expands the list of properties. In the Online Appendix, we provide the testable implications of $\varepsilon$-Rationalizability when imposing (a) quasi-linearity, (b) homotheticity, or (c) additive separability in addition to regularity. For these classes, we did not find characterization results in terms of acyclic satisfiability. This is consistent with the fact that (as far as we are aware) there are no SARP-like characterizations of standard rationalizability within these classes of preferences. Instead, classic tests are typically described by the existence of a solution to certain sets of inequalities, an approach we extend to the case of $\varepsilon$-Rationalizability. Though less insightful and somewhat harder to interpret, these tests are still practically useful to determine consistency with the theory. To illustrate, we present below variants and special cases of these results, which are destined to be applied to recent experimental data pertaining to choices from budget sets involving risk or delayed payments.

the computations that follow are continuous, the strict comparison we derive between the probability of having a smaller CCEI or FDI will remain correct.

20For instance, for all $\hat{z}$ halfway between $x'$ and $z$ on the $p'$-budget line, there is a positive mass of bundles $\hat{x}$ to the right of $x$ on the $p$-budget line for which the FDI of $\{(p, \hat{x}), (p', \hat{z})\}$ is strictly larger than that of $D$.
5.1 Risk

In a contemporaneous and independent paper, Echenique, Imai and Saito (2018) define and study what we would call \( \varepsilon \)-Rationalizability for a risk-averse expected utility maximizer. Their characterization of the empirical content is based on a weakening of Echenique and Saito (2015)’s SAROEU.\(^{21}\) Our result for additively separable preferences, now provided in the Online Appendix, solves a similar question by adapting Varian (1983)’s test of rationality for additively separable preferences.\(^{22}\)

To illustrate, we present our methodology in the simpler setting of recent lab experiments on risky portfolio choices. Given two equally-likely states, subjects face multiple linear budget sets and decide how to allocate money across states given these constraints. There are thus two commodities: money in state 1 and in state 2 (like Arrow securities). Demand data \( \mathcal{D} \) is \( \varepsilon \)-Rationalizable for a strictly risk-averse expected utility maximizer if there exists a strictly concave, differentiable, and strictly monotone Bernoulli function \( v : \mathbb{R}_+ \rightarrow \mathbb{R} \) such that

\[
1 - \varepsilon < \frac{v'(x_1)/v'(x_2)}{p_1/p_2} < \frac{1}{1 - \varepsilon},
\]

for each \((p, x) \in \mathcal{D}\). This is indeed the same equation as (2) when \( u(x) = \frac{v(x_1) + v(x_2)}{2} \) is the expected utility of the lottery that gives \( x_1 \) or \( x_2 \) with equal probability.

Suppose \( \mathcal{D} \) is \( \varepsilon \)-Rationalizable in this sense. By Proposition 3, for each \((p, x) \in \mathcal{D}\), one can find two weights \((\alpha_1(p, x), \alpha_2(p, x))\) such that

\[
x \in \arg \max_{\{y | p \cdot y \leq p \cdot x\}} \alpha_1(p, x)v(y_1) + \alpha_2(p, x)v(y_2),
\]

and

\[
1 - \varepsilon < \frac{\alpha_1(p, x)}{\alpha_2(p, x)} < \frac{1}{1 - \varepsilon}.
\]

Hence there exists a Lagrange multiplier \( \gamma(p, x) > 0 \) such that

\[
v'(x_\ell) = \frac{\gamma(p, x)p_\ell}{\alpha_\ell(p, x)}, \text{ for both } \ell = 1, 2.
\]

\(^{21}\)SAROEU stands for the Strong Axiom for Revealed Objective Expected Utility.
\(^{22}\)There is a small difference, though, as additive separability alone leaves the possibility of state-dependent Bernoulli functions.
For both $\ell = 1, 2$, $v$’s strict concavity implies that $v(m) < v(x_\ell) + v'(x_\ell)(m - x_\ell)$, for all $m \neq x_\ell$. By (11), this is equivalent to

$$v(m) < v(x_\ell) + w_\ell(p, x)p_\ell(m - x_\ell), \text{ for all } m \neq x_\ell,$$

with $w_\ell(p, x) = \gamma(p, x). Let $M = \{x_1 \mid x \in X\} \cup \{x_2 \mid x \in X\}$ be the set of dollar amounts in chosen bundles. To conclude, one can find a Bernoulli utility $\hat{v}(m) \in \mathbb{R}$, for each $m \in M$, and rescaled weights $w(p, x) \in \mathbb{R}^2_{++}$, for each $(p, x) \in \mathcal{D}$, such that

$$1 - \varepsilon < \frac{w_1(p, x)}{w_2(p, x)} < \frac{1}{1 - \varepsilon},$$

for each $(p, x) \in \mathcal{D}$ (as $\beta$ inherits this property from $\alpha$), and

$$\begin{cases} 
\hat{v}(x'') < \hat{v}(x') + w_\ell(p, x)p_\ell(x'' - x_\ell) & \text{if } x_\ell \neq x''_\ell \\
w_\ell(p, x)p_\ell = w_\ell(p', x')p''_\ell & \text{if } x_\ell = x''_\ell.
\end{cases}$$

for all $\ell, \ell' \in \{1, 2\}$ and all $(p, x), (p', x') \in \mathcal{D}$. The top line in (14) follows from (12) applied to $m = x''_\ell$, while the bottom line follows from (11). As it turns out, the necessary condition we found is also sufficient.

**Proposition 8** \(\mathcal{D}\) is $\varepsilon$-Rationalizable for a strictly risk-averse expected utility maximizer if, and only if, (13) and (14) hold for some $\hat{v} : M \to \mathbb{R}$ and $w : \mathcal{D} \to \mathbb{R}^2_{++}$.

Despite the strict inequalities, we show in the Online Appendix that the condition in Proposition 8 can be checked by linear programming.

Using this new notion of approximate expected utility, Echenique, Imai and Saito (2018) revisit experimental data from Choi, Kariv, Müller, and Silverman (2014), Carvalho, Meier, and Wang (2016), and Carvalho and Silverman (2017). Interestingly, they find that “younger subjects, those who have high cognitive abilities, and those who are working, are closer to EU behavior than older, low ability, or passive, subjects. For some of the three experiments, we also find that highly educated, high-income subjects, and males, are closer to EU.”

By permitting preferences beyond expected utility, our work suggests a way to improve our understanding of these correlations. One may wonder which channel(s) is most relevant in determining them. Could it be that the better compliance with expected utility in a sub-population is actually due to better compliance with ratio-
nality? To formalize this, observe that

\[ FDI_{EU}(D) = FDI(D) + [FDI_{EU}(D) - FDI(D)], \]

where \( FDI_{EU} \) denotes the FOC-Departure Index for a strictly risk-averse expected utility maximizer. Using the theory here, one can study whether a correlation of \( FDI_{EU} \) with some socio-economic variable (as established by Echenique et al.) is mostly due to a correlation of the FDI with those variables, or of the residual term instead. We are in the process of checking this in the data.

5.2 Time

Consider consumer data from convex time budget (CTB) data experiments (as introduced by Andreoni and Sprenger (2012)). Assume for simplicity that all observations pertain to two fixed times.\(^{23}\) There are thus two commodities: money at the earlier date (good 1), and money at the later date (good 2). Maximizing a discounted sum of utilities provides a standard refinement of rationality in such settings. For instance, with time-independent utility, the consumer might maximize \( u(x) = v(x_1) + \delta v(x_2) \) for some \( 0 \leq \delta \leq 1 \) and some strictly concave, differentiable, and strictly monotone function \( v : \mathbb{R}_+ \rightarrow \mathbb{R} \). Allowing for time-dependent utility is equivalent to requiring additive separability: the consumer might maximize \( u(x) = v(x_1) + w(x_2) \) for some some strictly concave, differentiable, and strictly monotone functions \( v \) and \( w \) (having the \( \delta \) is redundant in this case since it can be included in the definition of \( w \)).

As indicated earlier, conditions for \( \varepsilon \)-Rationalizability for additively separable preferences are provided in the Online Appendix. A straightforward variant of Proposition 8 covers the case of discounted utility with state-independent utility. Suppose that one wants to then gauge the goodness of fit of these nested models across populations. This time, we would look at the following decomposition:

\[ FDI_{DTI}(D) = FDI(D) + [FDI_{AS}(D) - FDI(D)] + [FDI_{DTI}(D) - FDI_{AS}(D)], \]

where \( FDI_{AS} \) denotes the FOC-Departure Index for additively separable preferences.

\(^{23}\) The analysis can be extended to data involving CTB’s with different times, e.g. part of the data pertaining to choices between \( t = 0 \) and \( t = 1 \), and part of the data pertaining to choices between \( t = 1 \) and \( t = 2 \).
FDTI refers to the more stringent case of discounting with time-independent utility. We are examining the correlation of these terms with socioeconomic data.

6. Beyond Linear Budget Sets

In an empirical study of nonlinear pricing in electricity markets, Ito (2014) points out that while optimization requires understanding marginal prices, “nonlinear pricing and taxation complicate economic decisions by creating multiple marginal prices for the same good.” Indeed, the empirical literature on non-linear pricing, with applications to labor and taxation (see the survey of Saez, Slemrod and Giertz (2012)) and utilities markets (see Ito (2014) and references therein) provides evidence that consumers respond to nonlinear pricing in a manner inconsistent with classic theories. These studies illustrate the importance of relaxing the assumption of perfect collinearity between marginal prices and the utility gradient.

The analysis developed so far lends itself very well to such an endeavor. We begin by extending our earlier ideas to quantify how far perceived prices would have to be from marginal prices to recover collinearity. Generalized demand data $\mathcal{D}$ comprises a finite collection of pairs $(P, x)$, where $P = (P_1, \ldots, P_L)$ is a collection of strictly increasing and differentiable price functions $P_\ell$, and $x \in \mathbb{R}^L_{++}$ is the consumption bundle demanded for the budget set

$$B(P, x) = \{ y \in \mathbb{R}^L_+ | \sum_\ell P_\ell(y_\ell) \leq \sum_\ell P_\ell(x_\ell) \}.$$ 

Indeed, $P_\ell(q)$ represents the total price to pay for buying a quantity $q \geq 0$ of good $\ell$. Naturally, $P_\ell(0) = 0$ for each $\ell$. We presume that $B(P, x)$ is convex (as would be true if $P_\ell$ is convex for all $\ell$), in which case the first-order conditions are not only necessary but sufficient for an optimum. Differentiability is imposed to guarantee that marginal prices are well-defined, in which case the boundary of $B(P, x)$ is smooth. This can easily be relaxed at the cost of more cumbersome notation, by considering the sub-differentials.

For each $(P, x) \in \mathcal{D}$, the Marginal Price Ratio associated to any two goods $\ell, \ell'$ is the following strictly positive number:

$$\text{MPR}_{\ell, \ell'}^P(x) = \frac{P_\ell'(x_\ell)}{P_{\ell'}'(x_{\ell'})}.$$
**Definition 2** (ε-Rationalizability) For ε ∈ (0, 1), the generalized demand data $\mathcal{D}$ is ε-rationalizable if there exists a regular utility function $u$ such that

$$1 - \varepsilon < \frac{MRS_{x\ell}(x)}{MP_{P\ell}(x)} < \frac{1}{1 - \varepsilon},$$

for each $\ell \neq \ell'$ and each $(P, x) \in \mathcal{D}$. Then $u$ is said to ε-rationalize the generalized demand data.

A couple of remarks are in order. First, we retain the name ε-rationalizability; this cannot create any confusion, since Definition 2 boils down to Definition 1 with linear prices. Second, recalling the discussion from Section 2.1, the inequalities in (15) ensure the gradient of $u$ and the marginal price vector $MP(x) := (P_1'(x_1), \ldots, P_L'(x_L))$ are not too far apart at chosen bundles:

$$\delta(\nabla u(x), MP(x)) < \frac{1}{1 - \varepsilon},$$

for all $(P, x) \in \mathcal{D}$.

The following observation follows at once from Definitions 1 and 2. It is powerful nevertheless, as it shows that all the results derived in Sections 3 and 4 apply to ε-rationalizability of generalized datasets.

**Proposition 9** Consider a generalized dataset $\mathcal{D}$, and the fictitious dataset $\hat{\mathcal{D}}$ defined as the set of observations $(p, x)$ with $p = MP(x)$, for all $(P, x) \in \mathcal{D}$. Then $\mathcal{D}$ is ε-rationalizable if, and only if, $\hat{\mathcal{D}}$ is ε-Rationalizable.

Thus one simply tests ε-rationalizability of generalized demand data using our earlier tests for standard demand data.\(^{24}\) In particular, we have thus found an easily computable and intuitive notion of goodness-of-fit for consumer data involving nonlinear prices.

The FDI easily extends to the nonlinear setting: it is the infimum among all $\varepsilon$ such that the generalized data is ε-rationalizable. Here there is an elucidating contrast with Afriat’s CCEI. Our approach offers a valid measure of rationality whenever the

\(^{24}\)Moreover, as we discussed at the end of Section 3, the testing approach is portable to some specific theories of misperception. For instance, in the context of nonlinear pricing, one might be interested in testing the theory of a consumer who maximizes a regular utility function $u$ given a perceived price vector $p^c$ that is a convex combination of average and marginal prices, and restricted to belong to $C(P, x) = \{[\alpha P_\ell'(x_\ell) + (1 - \alpha) \frac{P(x_\ell)}{x_\ell}]|_{\ell \in L} | \alpha \in [0, 1]\}$.}

26
first-order approach applies. The CCEI, on the other hand, only applies whenever it is clear what it means to ‘shrink’ a budget set. There are different quantifiable ways to ‘shrink’ a nonlinear budget set; these may lead to different CCEI’s, and it is unclear which interpretation is correct. Our approach does not face this issue, as it leaves the budget set unaffected.

REFERENCES


Carvalho, L., and D. Silverman (2016), Complexity and Sophistication, *mimeo*.


---

25 Proportionally pulling in the budget frontier has some appeal, as it retains the nonlinear shape. Shrinking monetary income as in the original definition may result in a budget set with an entirely different shape. Notice that both approaches give the same result in linear budget sets. Of course, in non-monetary settings, a consumer’s income is given by the value of their endowment. With linear prices, one can halve the budget by halving the endowment. With non-linear prices, halving the endowment need not halve monetary income (and again, may change the shape of the budget set.)


Saez, E., J. Slemrod and S. Giertz (2012), The Elasticity of Taxable Income


Appendix

Proof of Proposition 1 (Ordering satisfying the axioms from Section 2.1)

Let $\alpha = 1/x_2$, $\beta = 1/y_2$, $\alpha' = 1/x_2'$ and $\beta' = 1/y_2'$. Then, Unit Invariance tells us that

$$(x, y) \succeq (x', y') \iff ((x_1/x_2, 1), (y_1/y_2, 1)) \succeq ((x_1'/x_2', 1), (y_1'/y_2', 1)).$$

Using Measurement Invariance with $\alpha = y_2/y_1$ and $\alpha' = y_2'/y_1'$, this means

$$(x, y) \succeq (x', y') \iff \left( \left( \frac{x_1}{y_1} \right), (1, 1) \right) \succeq \left( \left( \frac{x_1'}{y_1'} \right), (1, 1) \right).$$

Regularity implies the existence of a function $f : \mathbb{R}^2_{++} \times \mathbb{R}^2_{++} \to \mathbb{R}$ such that $(x, y) \succeq (x', y')$ if and only if $f(x, y) \geq f(x', y')$. Defining $g(\alpha) := f((\alpha, 1), (1, 1))$, we have $(x, y) \succeq (x', y')$ if and only if $g(x_1/y_1) \geq g(x_1'/y_1')$. By Monotonicity, $g$ strictly increases for $\alpha \geq 1$. By symmetry, it must be that $g(\alpha) = g(1/\alpha)$ and thus $g(\alpha) = g(\max\{\alpha, 1/\alpha\})$. As $\max\{\alpha, 1/\alpha\} \geq 1$ for all $\alpha$, the result follows from inverting $g$.

Q.E.D.

Lemma 1 Let $A \subset \mathbb{R}^L_{++}$ be a finite set of bundles. The following statements hold:

(a) There exists a price vector $q \in \mathbb{R}^L_{++}$ and a bundle $x \in A$ such that $q \cdot x > q \cdot y$ for all $y \in A \setminus \{x\}$

(b) There exists a price vector $q \in \mathbb{R}^L_{++}$ such that $q \cdot x \neq q \cdot y$ for all $x, y \in A$.

Proof. (a) Let $x$ be the unique maximal element of $A$ according to the following lexicographic order: $x$ has the largest quantity of good 1 among all bundles in $A$, it has the largest quantity of good 2 among all the bundles singled out so far, etc. Let $q = (1, \eta, \eta^2, \ldots, \eta^{L-1})$. We claim that (a) is satisfied using $x$ and the price vector $q$, provided that $\eta > 0$ is small enough. Indeed, for each alternative $y \in A$, let $\ell$ be the first good for which $y_\ell \neq x_\ell$. Then, by definition of $x$, $y_\ell < x_\ell$. Then $q \cdot (x - y) = \eta^{\ell-1}[(x_\ell - y_\ell) + \sum_{k=\ell+1}^{L} \eta^{k-\ell}(x_k - y_k)]$. Clearly, there exists a threshold $\eta^*$ such that this expression is strictly positive for all $\eta < \eta^*$, which proves (a).

(b) Pick $x$ and $q$ as in (a). Call them $x^1$ and $q^1$. If $q^1 \cdot y \neq q^1 \cdot z$, for all pairs $y, z$ of distinct bundles in $A \setminus \{x^1\}$, then we are done. Otherwise, we show that there exists another price vector $q^2$ that preserves expenditure comparisons under $q^1$ when they are strict, but breaks some ties: (i) $q^1 \cdot y > q^1 \cdot z \Rightarrow q^2 \cdot y > q^2 \cdot z$ for all $y, z \in A$, and
We begin with necessity. Given regularity, note that condition (5a) holds for which means that a contradiction. Thus the existence of such a sequence 
exists a smooth, concave and strictly monotone utility function rationalizing claimed. By Chiappori and Rochet (1987) and Matzkin and Richter (1991), there that 
that (ii) in the definition of \( \hat{\varepsilon} \)
\[ \leq \]
for all \( 1 \)
\[ \leq \]
such that \( x, y \) \( \in \) \( X \) \{ \( x \in \mathbb{R}^L_+ \mid (p, x) \in \mathcal{D} \) for some \( p \in \mathbb{R}^L_+ \) \} (see (b) in the Lemma).
Consider new price vectors \( \tilde{p}(p, x) = \eta q + (1 - \eta)p(p, x) \), where \( \eta > 0 \) is small enough that (i) \( \delta(\tilde{p}(p, x), p) < \varepsilon \), for all \( (p, x) \in \mathcal{D} \), and (ii) \( \tilde{p}(p, x) \cdot x' > \tilde{p}(p, x) \cdot x \), for all \( (p, x) \) and \( (p', x') \) in \( \mathcal{D} \) with \( p(p, x) \cdot x' \geq p(p, x) \cdot x \). Clearly the demand data \( \mathcal{D}' \) inherits the invertibility property from \( \mathcal{D}' \). We now check that \( \mathcal{D}' \) satisfies SARP. Suppose there is a sequence \( \{(p^k, x^k)\mid 1 \leq k \leq K \} \) in \( \mathcal{D} \) such that \( x^{k+1} \neq x^k \) and \( \tilde{p}(p^k, x^k) \cdot x^k \geq \tilde{p}(p^k, x^k) \cdot x^{k+1} \) for all \( 1 \leq k \leq K \) (with the convention \( K+1 = 1 \)). First, observe \( p^k(p^k, x^k) \cdot x^k \geq p^k(p^k, x^k) \cdot x^{k+1} \) for all \( 1 \leq k \leq K \), thanks to (ii) in the definition of \( \tilde{p} \). Now, suppose that \( p^k(p^k, x^k) \cdot x^k = p^k(p^k, x^k) \cdot x^{k+1} \) for all \( 1 \leq k \leq K \). Then \( q \cdot x^k \geq q \cdot x^{k+1} \), for all \( 1 \leq k \leq K \). Furthermore, the inequalities must be strict, by definition of \( q \). We get \( q \cdot x^1 > q \cdot x^2 > \cdots > q \cdot x^K > q \cdot x^1 \), a contradiction. Thus the existence of such a sequence \( \{(p^k, x^k)\mid 1 \leq k \leq K \} \) means that \( \mathcal{D}' \) violates GARP, a contradiction. It must be that \( \mathcal{D}' \) satisfies SARP, as claimed. By Chiappori and Rochet (1987) and Matzkin and Richter (1991), there exists a smooth, concave and strictly monotone utility function rationalizing \( \mathcal{D}' \), which means that \( \mathcal{D} \) is \( \varepsilon \)-rationalizable.

**Proof of Proposition 3** (\( \varepsilon \)-Rationalizability and additively separable utility)

We begin with necessity. Given regularity, note that condition (5a) holds for \( x \in \mathbb{R}^L_+ \).
if and only if the first-order conditions for optimality hold. Thus optimality for the additive utility function in (5a) holds by setting $\alpha_\ell(p, x)/\alpha_{\ell'}(p, x) := (p_\ell/p_{\ell'})MRS_{\ell,\ell'}^u(x)$.

The bounds (5b) are implied by $\varepsilon$-Rationalizability using $u$.

For sufficiency, (5a) implies that for each $(p, x) \in D$, the optimality condition

$$(\alpha_\ell(p, x)/\alpha_{\ell'}(p, x))MRS_{\ell,\ell'}^u(x) = p_\ell/p_{\ell'}$$

holds for all $\ell, \ell'$. Using (5b), it is easy to see that $u(\cdot) = \sum_{\ell=1}^L u_\ell(\cdot) \varepsilon$-rationalizes $D$.

**Proof of Proposition 4** ($\varepsilon$-rationalizability and acyclic satisfiability)

It remains to prove sufficiency. By acyclic satisfiability, there is a strict ordering $\succ$ over $X$ with the feature that for all $x \in X$, there is $v(x) \in C_\varepsilon(x)$ such that $x \succ x'$ for all $x' \in \Gamma(x, v)$. We now construct an auxiliary demand data $D' = (v(x), x)_{x \in X}$ where the vector $v(x)$, which is strictly positive, is taken to be the price vector when $x$ is chosen. This data satisfies SARP, since any cycle in the revealed preferences from $D'$ would imply a cycle in $\succ$, a contradiction. Moreover, all the bundles $x \in X$ are unique by construction. Thus the auxiliary data $D'$ satisfies Matzkin and Richter (1991)'s Theorem 2\textsuperscript{\infty}(a), and there exists a regular (in fact even infinitely differentiable and strictly concave) utility function $u : \mathbb{R}_+^L \to \mathbb{R}$ that rationalizes $D'$ in the classic sense. In particular, for any $x \in X$, optimality of the demands requires $MRS_{\ell,\ell'}^u(x) = v_\ell(x)/v_{\ell'}(x)$. Finally, since $v(x) \in C_\varepsilon(x)$, we know by construction that for every $p \in P(x)$,

$$\frac{p_\ell}{p_{\ell'}} (1 - \varepsilon) < MRS_{\ell,\ell'}^u(x) = \frac{v_\ell(x)}{v_{\ell'}(x)} < \frac{p_\ell}{p_{\ell'}} \left( \frac{1}{1 - \varepsilon} \right).$$

Hence the original demand data $D$ is $\varepsilon$-rationalizable.  

**Q.E.D.**

**Proof of Proposition 5** (FDI with Two Commodities and Two Observations)

*Proof.* Suppose that $x \neq x'$, and that $\delta(p, o(x, x')) \leq \delta(p', o(x, x'))$ (a similar argument applies if the opposite inequality holds). Following the reasoning from Section 3, it is easy to check that the demand data is $\varepsilon$-rationalizable if and only if there exists $y \in \{x, x'\}$ and $v \in C_\varepsilon(y)$ such that $v \cdot y' > v \cdot y$, where $y'$ is the element in $\{x, x'\}$ distinct from $y$. (In words, one of the two budget lines can be adjusted within the limits imposed by $\varepsilon$ in a way that eliminates the SARP violation.) First, we check that the demand data is $\varepsilon$-rationalizable for all $\varepsilon > f^{-1}(\delta(p, o(x, x')))$, where $f(x) := 1/(1 - x)$ for all $x$. Notice that $o(x, x') \cdot x = o(x, x') \cdot x'$, by construction, and hence one can find a vector $v(x)$ such that $v(x) \cdot x' > v(x) \cdot x$ and $v(x)$ close
enough to $o(x, x')$ that $\varepsilon > f^{-1}(\delta(p, v(x)))$, or $\frac{1}{1-\varepsilon} > \delta(p, v(x))$. Second, consider an $\varepsilon \leq f^{-1}(\delta(p, o(x, x')))$. It is not difficult then to check that $v \cdot x' \leq v \cdot x$ and $v' \cdot x \leq v' \cdot x'$, for all $v, v' \in \mathbb{R}_+^2$ such that $\delta(v, p) \leq \frac{1}{1-\varepsilon}$ and $\delta(v', p') \leq \frac{1}{1-\varepsilon}$. Hence $D$ is not $\varepsilon$-rationalizable.

Suppose now that $x = x'$. Clearly, the demand data is $\varepsilon$-rationalizable if and only if $C_{\varepsilon}(x)$ is nonempty, which means that there exists $v \in \mathbb{R}_+^2$ such that both $\delta(p, v)$ and $\delta(p', v)$ are strictly inferior to $\frac{1}{1-\varepsilon}$. Let’s look for $v$’s that minimize the maximum of these two numbers. Remember the convention that $p_1' \geq p_1 \geq p_2'$, and hence $v_1/v_2$ will fall between these two ratios. With that in mind, it is easy to check that

$$\max\{\delta(p, v), \delta(p', v)\} = \max\{\frac{v_1/v_2}{p_1/p_2}, \frac{p_1'/p_2'}{v_1/v_2}\},$$

which will be reached when $v_1/v_2$ equals the geometric mean of $p_1/p_2$ and $p_1'/p_2'$, in which case the value of the max is $\sqrt{\frac{p_1'/p_2'}{p_1/p_2}}$. This number is strictly inferior to $\frac{1}{1-\varepsilon}$ if and only if $\varepsilon > 1 - \sqrt{\frac{p_1/p_2}{p_1'/p_2'}}$, hence the result.

**Proof of Proposition 6 (FDI with Two Commodities)**

*Proof.* Clearly, if a demand data is $\varepsilon$-rationalizable, then so is any pair of observations in it. Hence, the FDI of any demand data if larger or equal than the FDI associated to any pair of observations. As for the converse, we show that the demand data is $\varepsilon$-rationalizable, for any $\varepsilon$ strictly larger than the FDI’s associated to all pairs of observations.

By Proposition 4, it is sufficient to prove that $\mathcal{R}_\varepsilon(D)$ is acyclically satisfiable. We do that by induction: for each subset $X'$ of remaining elements in $X$, there is an element $x \in X'$ and $v \in C_{\varepsilon}(x)$ such that $\Gamma(x', v) \cap X' = \emptyset$.

First we establish that $C_{\varepsilon}(x) \neq \emptyset$ even for those $x$ where invertibility fails. Let $p'$ be the price vector achieving the maximal ratio $q_1/q_2$ over all price vectors $q$ such that $(q, x) \in D$. Let $p$ be the price vector achieving the minimal ratio $q_1/q_2$ over all price vectors $q$ such that $(q, x) \in D$. We have to show that there exists $v \in \mathbb{R}_+^2$ such that $\delta(q, v) < \frac{1}{1-\varepsilon}$ for all $q$ such that $(q, x) \in D$. Notice that $\delta(q, v) = \max\{\frac{q_1/q_2}{v_1/v_2}, \frac{v_1/v_2}{q_1/q_2}\} \leq \max\{\frac{p_1/p_2}{v_1/v_2}, \frac{v_1/v_2}{p_1/p_2}\} \leq \max\{\delta(p, v), \delta(p', v)\}$. The proof of Proposition 5 shows that $\max\{\delta(p, v), \delta(p', v)\} = \sqrt{\delta(p, p')}$ and thus the conclusion follows from the statement of Proposition 5 and the hypothesis that each pair of points in $D$ is $\varepsilon$-rationalizable.

Fix now $X' \subseteq X$ a set of remaining elements in $X$. Let $X''$ be the subset of
elements \( y \) in \( X' \) such that there is no alternative \( z \) in the convex hull of \( X' \) with \( z \leq y \). If \( X'' \) is a singleton, say \( X'' = \{y\} \), then \( z \geq y \) for all \( z \in X' \). Then \( \Gamma(y, v) \cap X' = \emptyset \), for all \( v \in \mathbb{R}_{++}^2 \), and in particular for any \( v \in C_\varepsilon(y) \). Next, suppose that \( X'' \) contains two or more elements. Let’s enumerate \( X'' \): \( X'' = \{x^1, x^2, \ldots, x^K\} \) with the property that \( x^j_1 < x^k_1 \) for all \( j < k \) (in which case we also have \( x^j_2 > x^k_2 \)). If there is \( x \in X'' \) such that \( p \cdot x < p \cdot y \) for all \( y \in X' \) and all \( p \) such that \( (p, x) \in D \), then \( \Gamma(x', v) \cap X' = \emptyset \) for all \( v \in C_\varepsilon(x) \), and we are done. So let’s assume instead that for all \( k \) there exists \( p \) and \( j \neq k \) such that \( (p, x^k) \in D \) and \( p \cdot x^j \leq p \cdot x^k \).

By the structure of \( X'' \), given any such \( p \), either all \( j \)'s with this property are larger than \( k \), or all such \( j \)'s are smaller than \( k \). Of course, when varying the \( p \)'s such that \( (p, x^k) \in D \), it may be possible to find larger-than-\( k \) \( j \)'s associated to some \( p \)'s, and smaller-than-\( k \) \( j \)'s associated to others. For each \( k \), let \( f(k) \) be the letter \( L \) (resp. \( R \)) if we find only smaller-than-\( k \) (resp. larger-than-\( k \)) \( j \)'s when considering the various \( p \)'s such that \( (p, x) \in D \). Thus \( f \) indicates the direction in which Samuelson revealed preferences can occur. Finally, \( f(k) = LR \) if there exists both \( p \) and \( p' \) such that \( (p, x^k) \in D \), \( (p', x^k) \in D \), \( p \cdot x^j \leq p \cdot x^k \) for \( j < k \), and \( p' \cdot x^j \leq p' \cdot x^k \) for \( k < j \).

Suppose there exists \( j \) such that \( f(x^j) = R \) and \( f(x^{j+1}) = L \). Let \( p^j \) be the price vector with the smallest price ratio \( p^j_1/p^j_2 \) among those such that \( (p^j, x^j) \in D \). Let \( p^{j+1} \) be the price vector with the largest price ratio \( p^{j+1}_1/p^{j+1}_2 \) among those such that \( (p^{j+1}, x^{j+1}) \in D \). Since the FDI of \( \{(p^j, x^j), (p^{j+1}, x^{j+1})\} \) is strictly less than \( \varepsilon \), it must be that \( \delta(o(x^j, x^{j+1}), p^j) < \frac{1}{1-\varepsilon} \) or \( \delta(o(x^j, x^{j+1}), p^{j+1}) < \frac{1}{1-\varepsilon} \). Suppose the latter holds (a similar reasoning holds in the other case), then consider a vector \( v \) that is very close to \( o(x^j, x^{j+1}) \) but with a slightly smaller first component (making the line orthogonal to it slightly flatter). By construction, \( \Gamma(x^{j+1}, v) = \emptyset \) and \( v \in C_\varepsilon(x^{j+1}) \), as \( \delta(p^{j+1}, v) \) will remain strictly smaller than \( \frac{1}{1-\varepsilon} \) when \( v \) is close enough to \( o(x^j, x^{j+1}) \).

Now notice that \( f(x^j) = R \) and \( f(x^K) = L \). If there is no \( j \) such that \( f(x^j) = R \) and \( f(x^{j+1}) = L \), then there must be a sequence of indices \( \{j, j+1, \ldots, k\} \) such that \( f(j) = R \), \( f(k) = L \) and \( f(i) = LR \) for all \( j < i < k \). The reasoning from the previous paragraph can be applied to \( j \) and \( j+1 \). The only issue that may arise is that \( o(x^j, x^{j+1}) \) need not belong to \( C_\varepsilon(x^{j+1}) \). That will happen if \( \delta(o(x^j, x^{j+1}), q) > \frac{1}{1-\varepsilon} \), where \( q \) is the price vector with the smallest price ratio \( q_1/q_2 \) among those such that \( (q, x^{j+1}) \in D \). Let \( q'_1/q'_2 \) be the price ratio in between \( q_1/q_2 \) and \( p^{j+1}_1/p^{j+1}_2 \), such that \( \delta(q, q') = \frac{1}{1-\varepsilon} \). If \( q' \cdot x^{j+1} < q' \cdot x^{j+1} \) for \( i > j+1 \), then we are done because \( \Gamma(x^{j+1}, v) = \emptyset \) and \( v \in C_\varepsilon(x^{j+1}) \), for some \( v \) close enough to \( q' \) (obtained by decreasing a little bit
the first component).

Assume instead that \( q' \cdot x^{j+1} \leq q' \cdot x^i \) for \( i > j + 1 \). Let \( p^{j+2} \) be the price vector with the largest price ratio \( p^{j+2}_1/p^{j+2}_2 \) among those such that \( (p^{j+2}, x^{j+2}) \in \mathcal{D} \). Since the FDI of \( \{(q, x^{j+1}), (p^{j+2}, x^{j+2})\} \) is strictly less than \( \varepsilon \) and \( \delta(o(x^{j+1}, x^{j+2}), q) \geq \frac{1}{1-\varepsilon} \), it must be that \( \delta(o(x^{j+1}, x^{j+2}), p^{j+2}) < \frac{1}{1-\varepsilon} \). Now the same reasoning as in the above paragraph applies. We would be done if the new \( q' \) that arise from it has the property that \( q' \cdot x^{j+2} < q' \cdot x^i \) for \( i > j + 2 \). If not, then we iterate the construction. Remember though that \( f(x^K) = L \), and thus, even though the \( q' \) arising in each step may have failed the property until now, it will have to pass it in the last step. \( Q.E.D. \)

**Proof of Proposition 8** (Testable implications of \( \varepsilon \)-rationalizability for strictly risk-averse expected utility maximizer)

Necessity was proved in the text. As for sufficiency inequalities on the top line of (14) implies that there exists a strictly concave, infinitely differentiable and strictly monotone utility function \( v : \mathbb{R}_+ \to \mathbb{R} \) such that \( v(x_\ell) = \hat{v}(x_\ell) \) and \( v'(x_\ell) = w_\ell(p, x)p_\ell \) for each \( \ell = 1, 2 \) and each \( (p, x) \in \mathcal{D} \) (well-defined given the bottom line of (14)). Writing \( \alpha_\ell(p, x) \) for \( 1/w_\ell(p, x) \), this implies that for each \( (p, x) \in \mathcal{D} \), the bundle \( x \) maximizes \( \alpha_1(p, x)v(y_1) + \alpha_2(p, x)v(y_2) \) for all \( y \in \mathbb{R}^2_+ \) such that \( p \cdot y \leq p \cdot x \). Clearly, the \( \alpha \)'s inherit property (13) from the \( \beta \)'s, and we conclude from Proposition 3 that \( \mathcal{D} \) is \( \varepsilon \)-rationalizable by \( u(x) = \frac{v(x_1) + v(x_2)}{2} \). \( Q.E.D. \)