7-2024

Coarse revealed preference

Gaoji HU

Jiangtao LI
Singapore Management University, jtli@smu.edu.sg

John K-H QUAH

Rui TANG

Coarse Revealed Preference*

Gaoji Hu Jiangtao Li John K.-H. Quah Rui Tang

July 6, 2024

Abstract

We propose a novel concept of rationalization, called *coarse rationalization*, tailored for the analysis of datasets where an agent's choices are imperfectly observed. We characterize those datasets which are rationalizable in this sense and present an efficient algorithm to verify the characterizing condition. We then demonstrate how our results can be applied through a duality approach to test the rationalizability of datasets with perfectly observed choices but imprecisely observed linear budget sets. For datasets that consist of both perfectly observed feasible sets and choices but are inconsistent with perfect rationality, our results could be used to measure the extent to which choices or prices have to be perturbed to recover rationality.

KEYWORDS: Coarse dataset, rationalization, revealed preference, Afriat's Theorem, perturbation index, price misperception index

* We thank Mark Dean, Eddie Dekel, David Dillenberger, Faruk Gul, Bart Lipman, Pietro Ortoleva, Wolfgang Pesendorfer, Satoru Takahashi, Jingni Yang, and numerous participants at various seminars and conferences for helpful discussions. **Hu**: Shanghai University of Finance and Economics, *hu.gaoji@gmail.com*. **Li**: Singapore Management University, *jtli@smu.edu.sg*. **Quah**: National University of Singapore, *ecsqkhj@nus.edu.sg*. **Tang**: Hong Kong University of Science and Technology, *ruitang.economics@gmail.com*.

1 Introduction

The seminal results in revealed preference analysis begin with a dataset collected from a consumer and find conditions that are necessary and sufficient for it to be consistent with rationality. To be specific, suppose that the consumer chooses bundles from the consumption space $\mathbb{R}^n_+$. The dataset is a finite set of observations $\mathcal{O} = \{(A^t, B^t)\}_{t \in T}$ indexed by t, where A^t and B^t are nonempty subsets of $\mathbb{R}^n_+$ with $A^t \subseteq B^t$. With the interpretation that A^t is the set of bundles which the consumer has been observed to choose from the *budget set* B^t, there are two natural concepts of rationalization.

The **first concept** requires a preference $\succsim$, i.e., a complete and transitive binary relation on the consumption space, with its strict part denoted by $\succ$, such that the optimal choices within B^t consist precisely of A^t, i.e.,

for all $x \in A^t$*, we have* $x \succsim y$ *for all* $y \in B^t$ *and* $x \succ y$ *for all* $y \in B^t \setminus A^t$.

Richter's Theorem (Richter, 1966) characterizes those datasets $\mathcal{O}$ which are rationalizable in this sense. The **second concept** requires a preference $\succsim$ such that every bundle in A^t is optimal (with respect to $\succsim$) but allows for the possibility that bundles in $B^t \setminus A^t$ are optimal; in other words, it simply requires that

for all $x \in A^t$*, we have* $x \succsim y$ *for all* $y \in B^t$.

Afriat's Theorem (and its generalizations to nonlinear domains in Forges and Minelli (2009) and Nishimura, Ok and Quah (2017)) characterize datasets that satisfy this concept of rationalization. Loosely speaking, the first notion is the one commonly used in the theoretical revealed preference literature, while empirical work using revealed preference have mostly relied on the second (weaker) notion, which is unsurprising since the second concept does not posit that the observer has observed all the optimal choices, but only one, or some, of them.

A **third concept** of rationalization has been studied by Fishburn (1976). This concept involves a *different* interpretation of the dataset, where the observations are thought to be *coarse*. By this we mean that the observer knows that, when presented with the budget B^t, the agent has chosen from among the bundles in

A^t, but does not know precisely which bundle in A^t was chosen. To be precise, Fishburn requires the existence of a preference $\succsim$ such that for every $t \in T$,

$$\textit{there exists } x^t \in A^t \textit{ with } x^t \succsim y \textit{ for all } y \in B^t \textit{ and } x^t \succ y \textit{ for all } y \in B^t \setminus A^t.$$

In other words, Fishburn's concept of rationalization relaxes the first concept (of rationalization) by allowing some elements of A^t to be nonoptimal, but it retains the requirement that nothing outside of A^t is optimal. This suggests that a **fourth concept** of rationalization may be useful in empirical applications: one that allows for the possibility that some elements in A^t are nonoptimal (following Fishburn) and also that some elements outside of A^t are optimal (following Afriat). Formally, this concept simply requires the existence of a preference $\succsim$ such that for all $t \in T$,

$$\textit{there exists } x^t \in A^t \textit{ with } x^t \succsim y \textit{ for all } y \in B^t.$$

This rationalization concept, which we shall refer to as *coarse rationalization*, is the focus of our paper.

The revealed preference literature since the 1970s have by and large neglected Fishburn's rationalization concept. We think that Fishburn's concept, as well as the relaxation of that concept which we just proposed, deserves notice because they are relevant to empirical applications of revealed preference. These concepts are applicable whenever the observer knows (or hypothesizes) that there is an optimal choice found in A^t, but is agnostic about precisely which alternatives within A^t are optimal. There are at least three scenarios in which it is useful to think of coarse rationalization.

(1) The most obvious cases are those where the bundles chosen are known to be imprecise. For example, a researcher may have information on how much is spent on broad categories of goods, without knowing the allocation within each category. An economist can estimate a worker's total income based on his hourly wage, but may only have a rough idea of his choices regarding leisure and consumption goods. Alternatively, a researcher may have records on a consumer's credit card purchases, which puts a *lower* bound on how much is spent each month on different goods. Since there could be goods bought with cash, relying merely on such records, the researcher may not be able to recover the consumer's precise breakdown of monthly

expenditure on each good.

(2) There could be situations where some alternative x^t is recorded as the choice from B^t but, in testing for rationality or estimating the preference, the researcher may wish to accommodate the possibility that choices were observed with error; this could be accomplished by defining a neighborhood A^t around x^t (in some sense appropriate to the specific context) and then checking if $\mathcal{O}$ is coarsely rationalizable.

(3) In experimental settings, it is common to find subjects whose choice behavior are not exactly consistent with rationality. Since the choices x^t are typically observed perfectly, the rationality violations are not due to observational errors. Nonetheless, one could still use the size of the neighborhood A^t around x^t (suitably measured) as a way of comparing the rationality of different experimental subjects; those who require larger A^ts to rationalize their behavior can be deemed less rational.

In Section 2 of this paper we formulate a condition called the *never-covered property* (NCP) which is necessary and sufficient for a dataset $\mathcal{O} = \{(A^t, B^t)\}_{t\in T}$ to admit a coarse rationalization (i.e., rationalization according to the fourth concept) by a continuous and strictly increasing utility function.[1] This result could be thought of as a generalization of Afriat's Theorem, which characterizes rationalizability by a continuous and strictly increasing utility function in the case where A^t is a singleton.[2] It is well-known that in Afriat's Theorem, rationalization is characterized by the generalized axiom of revealed preference (GARP); this concept coincides with the never-covered property when A^ts are singleton sets. It is well-known that GARP can be easily checked; we show that there is a computationally efficient way of checking NCP, which facilitates the use of the concept in empirical applications.

Our concept of coarse rationalization pertains to situations in which a consumer's choices are not perfectly observed. A related and natural question is how to test the rationalizability of a consumer's choices when the budget set is not precisely known, perhaps because prices are imperfectly observed. In this case

[1] By this we mean that $\mathcal{O}$ has a coarse rationalization by a preference that can be represented by such a utility function.

[2] In this case, the second and fourth concepts of rationalization coincide.

a dataset has the form $\mathcal{O}^* = \{(x^t, \{B^{t,s}\}_{s\in G_t})\}_{t\in T}$ where for each observation $t \in T$, the bundle x^t is the observed choice made by the consumer, while the true budget set from which x^t is chosen is only known to be a set in the collection $\{B^{t,s}\}_{s\in G_t}$. We could then ask whether $\mathcal{O}^*$ has a *dual coarse rationalization* in the sense that there is a selection $s_t \in G_t$, for every observation $t \in T$, and a preference such that x^t is optimal in B^{t,s_t} for all $t \in T$.

In Section 3, we show that when the budget sets $B^{t,s}$ are classical budget sets (so $B^{t,s} = \{x \in \mathbb{R}^n_+ : p^{t,s} \cdot x \le 1\}$, for some vector of strictly positive prices $p^{t,s}$), then the dual coarse rationalizability of $\mathcal{O}^*$ is equivalent to the coarse rationalizability of some dataset $\mathcal{O}^{**} = \{(\underline{A}^t, \underline{B}^t)\}_{t\in T}$ (which can be straightforwardly constructed from $\mathcal{O}^*$). Therefore, the characterization result and the efficient algorithm that we developed in Section 2 can be applied to ascertain if $\mathcal{O}^*$ admits a dual coarse rationalization by a continuous and strictly increasing utility function.

Section 4 applies our results to the computation of rationality indices. We assume that A^t is a singleton and that B^t is a classical budget set, so the dataset has the form $\mathcal{D} = \{(x^t, L(p^t))\}_{t\in T}$, where $x^t \in \mathbb{R}^n_+$ is the observed choice from the budget set $L(p^t) = \{x \in \mathbb{R}^n_+ : p^t \cdot x \le 1\}$, where p^t is a vector of strictly positive prices. Afriat's Theorem tells us that GARP is a necessary and sufficient condition for $\mathcal{D}$ to be rationalized by a continuous and strictly increasing utility function. However, in most empirical applications, it is common for subjects to fail GARP. Various indices have been proposed to measure the severity of a subject's departure from rationality and we focus on two such indices.

One index, which we call the *perturbation index* measures the seriousness of rationality violations by measuring the extent to which the observed bundles x^t have to be perturbed for the dataset to be rationalizable (while keeping the budget set at observation t fixed at $L(p^t)$). Such an approach is intuitive and very similar to the one adopted by Varian (1985) to measure deviations from cost-minimizing factor demand in a production model (see Appendix for a fuller discussion). Another index, the *price misperception index* was proposed by de Clippel and Rozen (2023) and measures the extent to which the price vectors p^t have to be altered (which can be interpreted as misperception by the consumer) in order to restore rationality. Our main results enable both indices to be calculated with ease. We also provide an illustration of the performance of our algorithm by computing the perturbation indices of subjects in a portfolio choice experiment carried out by Choi et al. (2007).

2 Coarse Rationalizability

In this section, we formulate the notion of coarse rationalization and develop a necessary and sufficient condition — called the never-covered property — under which a dataset admits such a rationalization.

2.1 Basic Concepts

Let the consumption space be $\mathbb{R}^n_+$. A consumer's *preference* $\succsim$ is a binary relation on $\mathbb{R}^n_+$ that is complete and transitive. We use $\succ$ to denote the asymmetric part of $\succsim$, and refer to it as a *strict preference*.[3] A preference $\succsim$ is *continuous* if, whenever $x \succ y$, there are open neighborhoods N_x and N_y of x and y respectively, such that for all $x' \in N_x$ and $y' \in N_y$, $x' \succ y'$. It is well-known that any continuous preference on $\mathbb{R}^n_+$ admits a continuous utility representation $U : \mathbb{R}^n_+ \to \mathbb{R}$, i.e., $U(x) \geq U(y)$ if and only if $x \succsim y$ (see Debreu (1954)). A preference $\succsim$ is *locally nonsatiated* if, for every $x \in \mathbb{R}^n_+$ and every open neighborhood N_x of x, there is $x' \in N_x$ such that $x' \succ x$. Local nonsatiation holds if $\succsim$ is *increasing*; by this we mean that $x \succsim y$ if $x \geq y$ and $x \succ y$ if $x \gg y$.[4] The preference $\succsim$ is *strictly increasing* if $x \succ y$ whenever $x > y$. Clearly, if $\succsim$ admits a utility representation $U : \mathbb{R}^n_+ \to \mathbb{R}$, then U will be an increasing (strictly increasing) function if $\succsim$ is increasing (strictly increasing).[5] We refer to U as *regular* if it is continuous and strictly increasing.

[3] A *binary relation* R on X is a nonempty subset of $X \times X$. We write xRy to mean that $(x, y) \in R$. The binary relation R is *reflexive* if for all $x \in X$, xRx, *transitive* if for all $x, y, z \in X$, xRy and yRz imply xRz, and *complete* if for all $x, y \in X$, either xRy or yRx holds. The *asymmetric part of* R is the binary relation P on X such that xPy if xRy but not yRx.

[4] For any two n-dimensional vectors x and y, $x \geq y$ means that for each $i = 1, ..., n$, $x_i \geq y_i$, $x > y$ means that $x \geq y$ and $x \neq y$, and $x \gg y$ means that for each $i = 1, ..., n$, $x_i > y_i$.

[5] To be precise, we refer to a function $g : \mathbb{R}^n_+ \to \mathbb{R}$ as increasing if for all $x, y \in \mathbb{R}^n_+$, $x \geq y$ implies $g(x) \geq g(y)$ and $x \gg y$ implies $g(x) > g(y)$. We refer to g as being *strictly increasing* if, for all $x, y \in \mathbb{R}^n_+$, $x > y$ implies $g(x) > g(y)$.

For $x \in \mathbb{R}^n_+$ and $B \subseteq \mathbb{R}^n_+$, we write $x \succsim B$ if $x \succsim y$ for all $y \in B$ and $x \succ B$ if $x \succ y$ for all $y \in B$. We say that x is $\succsim$*-optimal* in B if $x \in B$ and $x \succsim y$ for all $y \in B$; the set of $\succsim$-optimal elements in B is denoted by $\max(B; \succsim)$. Obviously, if $\succsim$ admits a utility function U, then for all $x \in \max(B; \succsim)$ and $y \in B$, we have $U(x) \geq U(y)$.

Suppose a researcher has collected a finite set of observations $\mathcal{O} = \{(A^t, B^t)\}_{t \in T}$ (indexed by t), where A^t and B^t are nonempty subsets of $\mathbb{R}^n_+$ and $A^t \subseteq B^t$. We interpret B^t as the budget set from which the consumer chooses at observation t and A^t as the subset of B^t from which the choice was made. Following Forges and Minelli (2009), we assume that, for each t,

$$B^t = \{x \in \mathbb{R}^n_+ \,|\, g^t(x) \leq 0\}$$

where $g^t : \mathbb{R}^n_+ \to \mathbb{R}$ is a continuous and increasing function. Clearly, this formulation covers the case where B^t is a classical linear budget set; in this case, there is a price vector $p^t \gg 0$ prevailing at observation t, such that B^t consists of those bundles that cost less than the agent's wealth (normalized at 1), i.e., $B^t = L(p^t)$ where, for any $p \gg 0$,[6]

$$L(p) = \{x \in \mathbb{R}^n_+ \,|\, p \cdot x \leq 1\}. \tag{1}$$

In this case, $g^t(x) = p^t \cdot x - 1$, which is a continuous and strictly increasing function of x.

The following rationalization concept on $\mathcal{O}$ is the focus of our paper.

Definition 1. *$\mathcal{O} = \{(A^t, B^t)\}_{t \in T}$ has a* coarse rationalization *by the preference $\succsim$ on $\mathbb{R}^n_+$ if for all $t \in T$,*

$$\max(B^t; \succsim) \cap A^t \neq \emptyset.$$

When we refer to $\mathcal{O}$ as being coarsely rationalized by a utility function U, we mean that it is being rationalized by the preference U induces. When A^t is a singleton for all t, a coarse rationalization by a preference $\succsim$ would simply require

[6] It is always without loss of generality to normalize the agent's wealth to 1: for any price vector p^t and wealth $I^t > 0$ with which the associated budget set is $B^t = \{x \in \mathbb{R}^n_+ \,|\, p^t \cdot x \leq I^t\}$, we can consider an alternative price vector $\hat{p}^t = \frac{1}{I^t} p^t$ such that $B^t = L(\hat{p}^t)$.

the unique element in A^t to be $\succsim$-optimal; in this case, we shall simply refer to a coarse rationalization as a *rationalization*.

2.2 The Never-Covered Property

When does a dataset $\mathcal{O} = \{(A^t, B^t)\}_{t\in T}$ admit a coarse rationalization? As stated, our question has a trivial answer, since any dataset can be coarsely rationalized by a preference where all bundles are deemed to be indifferent. However, as we shall see, once we require the preference to be locally nonsatiated, then it is no longer the case that every dataset is coarsely rationalizable. Our main result is a generalization of the well-known theorem of Afriat (1967), which characterizes those datasets which are rationalizable by a locally nonsatiated preference, in the special case where A^t is a singleton.

To motivate our characterization of coarse rationalizability, we first suppose that $\mathcal{O} = \{(A^t, B^t)\}_{t\in T}$ has a coarse rationalization by a locally nonsatiated preference $\succsim$. By definition, there is $x \in A^t$ with $x \succsim B^t$. Since $\succsim$ is locally nonsatiated, $x \notin \mathring{B}^t$.[7] Thus, A^t cannot be covered by (in other words, contained in) $\mathring{B}^t$.

This argument could be generalized to more than one observation. For any nonempty $T' \subseteq T$, let

$$A(T') := \bigcup_{t\in T'} A^t \quad \text{and} \quad B(T') := \bigcup_{t\in T'} B^t.$$

Notice that if $\hat{x}$ satisfies $\hat{x} \succsim B^t$ for all $t \in T'$ then $\hat{x} \notin \mathring{B}(T')$. Such an $\hat{x} \in A(T')$ indeed exists: for each $t \in T'$, pick $x^t \in A^t$ such that $x^t \succsim B^t$ and let $\hat{x} \in \max(\{x^t\}_{t\in T'}; \succsim) \subseteq A(T')$. Thus, $A(T')$ cannot be covered by $\mathring{B}(T')$. Let $\hat{t}$ be an observation at which $\hat{x} \in A^{\hat{t}}$.

We are now ready to introduce the procedure that we call *the iterated exclusion of dominated observations*. Given a nonempty $T' \subseteq T$, let $\Phi^0(T') := \emptyset$, and let $\Phi^1(T')$ consist of t such that A^t is covered by $\mathring{B}(T')$, i.e.,

$$\Phi^1(T') := \left\{t \in T' : A^t \subseteq \mathring{B}(T')\right\}.$$

[7] For any set K, we write $\mathring{K}$ to denote its interior.

Since $\hat{t} \notin \Phi^1(T')$, we have $\Phi^1(T') \neq T'$. Let

$$\Phi^2(T') := \left\{t \in T' : A^t \subseteq \mathring{B}(T') \bigcup B(\Phi^1(T'))\right\}.$$

Obviously, $\Phi^1(T') \subseteq \Phi^2(T')$. Since $\hat{x} \succ \mathring{B}(T')$, we obtain $\hat{x} \succ A^t$ for each $t \in \Phi^1(T')$; since $A^t \cap \max(B^t; \succsim) \neq \emptyset$, we know that $\hat{x} \succ B^t$ for each $t \in \Phi^1(T')$. Thus, $\hat{x} \succ B(\Phi^1(T'))$. We conclude that $A(T')$ (which contains $\hat{x}$) cannot be covered by $\mathring{B}(T') \bigcup B(\Phi^1(T'))$ and so $\hat{t} \notin \Phi^2(T')$. We may repeat this argument for $m = 2, 3, \ldots$, where

$$\Phi^{m+1}(T') := \left\{t \in T' : A^t \subseteq \mathring{B}(T') \bigcup B(\Phi^m(T'))\right\}.$$

Since $\Phi^m(T')$ is an increasing sequence in m (in the set inclusion sense) and T' is finite, the procedure stops at some m^* when $\Phi^{m^*}(T') = \Phi^{m^*+1}(T')$. Let $\Phi(T') := \Phi^{m^*}(T')$; we refer to $\Phi(T')$ as the set of *revealed dominated observations* (or simply *dominated observations*) in T'. Since $\mathring{B}(T') \bigcup B(\Phi(T'))$ cannot contain $\hat{x}$, we obtain $\hat{t} \notin \Phi(T')$. Thus, $\Phi(T')$ is a strict subset of T'.

Definition 2. *$\mathcal{O} = \{(A^t, B^t)\}_{t \in T}$ satisfies* the never-covered property *(NCP) if for every nonempty $T' \subseteq T$, $\Phi(T') \neq T'$.*

We have shown that NCP is a necessary condition for $\mathcal{O}$ to be coarsely rationalizable. Theorem 1 below shows that it is also sufficient. Indeed, whenever $\mathcal{O}$ satisfies NCP then it can be coarsely rationalized by a continuous and increasing utility function.

Theorem 1. *The following statements on $\mathcal{O} = \{(A^t, B^t)\}_{t \in T}$ are equivalent:*

(1) *$\mathcal{O}$ can be coarsely rationalized by a locally nonsatiated preference.*

(2) *$\mathcal{O}$ satisfies NCP.*

(3) *$\mathcal{O}$ can be coarsely rationalized by a continuous and increasing utility function.*

Our proof of Theorem 1 makes use of known results in the case where A^t is a singleton. We refer to a dataset with this property as a *standard dataset* and write it as $\mathcal{D} = \{(x^t, B^t)\}_{t \in T}$. It is well-known that $\mathcal{D}$ is rationalizable by a locally nonsatiated preference $\succsim$ (in the sense that $x^t \succsim B^t$ for all $t \in T$) if and only if it obeys the *generalized axiom of revealed preference* (GARP). To understand this property, let $Y = \{x^t\}_{t \in T}$. For x^t and $x^{t'}$ in Y, we say that x^t is *revealed preferred to* $x^{t'}$ and denote it by $x^t \, R \, x^{t'}$ if $x^{t'} \in B^t$, and we say that x^t is *revealed strictly preferred to* $x^{t'}$ and denote it by $x^t \, P \, x^{t'}$ if $x^{t'} \in \mathring{B}^t$. GARP requires that there does not exist x^{t_1}, $x^{t_2} \ldots, x^{t_n}$ in Y such that

$$x^{t_1} \, R \, x^{t_2} \, R \cdots R \, x^{t_n} \text{ and } x^{t_n} \, P \, x^{t_1}. \tag{2}$$

It is straightforward to check that if $\mathcal{D}$ is rationalizable by a locally nonsatiated preference then it must obey GARP and, in fact, the converse is also true (see Forges and Minelli (2009) and Nishimura, Ok and Quah (2017)).[8] To prove Theorem 1, it suffices to show that when the coarse dataset $\mathcal{O} = \{(A^t, B^t)\}_{t \in T}$ satisfies NCP, then we can find $x^t \in A^t$ such that $\mathcal{D} = \{(x^t, B^t)\}_{t \in T}$ satisfies GARP. The details of how this is done are found in the Appendix.

In the case where B^t are linear budget sets (see (1)) with $p^t \gg 0$, we can sharpen Theorem 1. As usual, we can find $x^t \in A^t$ such that $\mathcal{D} = \{(x^t, B^t)\}_{t \in T}$ satisfies GARP. The eponymous theorem of Afriat (1967) is then applicable and guarantees that $\mathcal{D}$ can be rationalized by a utility function that is regular (in the sense of being continuous and strictly increasing) and concave.

Theorem 2. *The following statements on $\mathcal{O} = \{(A^t, L(p^t))\}_{t \in T}$ are equivalent:*

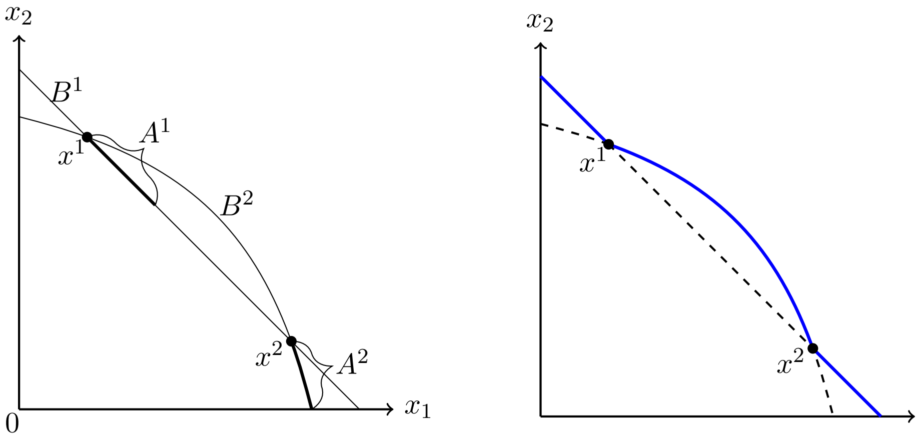
(1) *$\mathcal{O}$ can be coarsely rationalized by a locally nonsatiated preference.*

(2) *$\mathcal{O}$ satisfies NCP.*

(3) *$\mathcal{O}$ can be coarsely rationalized by a strictly increasing, continuous and concave utility function.*

Example 1. The coarse dataset depicted in Figure 1a consists of two observations, so we may write $T = \{1, 2\}$. For $T' = \{1\}$, we have $\Phi(T') = \emptyset$ since A^1 is on the boundary of B^1 and so $\Phi(T') \neq T'$ as required by NCP. Clearly, the same holds for $T' = \{2\}$. For $T' = \{1, 2\}$, notice that $x^1 \in A^1$ is not contained in $\mathring{B}^1 \cup \mathring{B}^2$ and neither is $x^2 \in A^2$. Thus $\Phi(T') = \Phi^1(T') = \emptyset \neq T'$. We conclude that this dataset satisfies NCP. It follows that there ought to be a selection from A^1 and A^2 so that

[8] Proofs of this result when B^t are linear are found in Afriat (1967) and Varian (1982). The term *generalized axiom of revealed preference* follows Varian (1982).



(a) $\mathcal{O} = \{(A^1, B^1), (A^2, B^2)\}$ (b) Indifference curve through x^1 and x^2

Figure 1: Coarse Rationalization

the resulting dataset obeys GARP; indeed $\mathcal{D} = \{(x^1, B^1), (x^2, B^2)\}$ obeys GARP. Figure 1b depicts the indifference curve passing through x^1 and x^2 of a preference that rationalizes the data; notice that this indifference curve is not convex. Indeed it is quite clear that *any* locally nonsatiated preference that coarsely rationalizes $\mathcal{O}$ must have both x^1 and x^2 as optimal bundles in B^1 (and in B^2) and that such a preference cannot be convex. This does not contradict Theorem 2 since B^2 is not a linear budget set. □

Variations on Coarse Rationalization. Nishimura, Ok and Quah (2017) extend Afriat's Theorem by characterizing $\mathcal{D} = \{(x^t, B^t)\}_{t \in T}$ that can be rationalized by a utility function that is continuous and increasing with respect to a given preorder (with the product order as a special case). For example, when an agent is choosing among bundles of contingent consumption, with states having commonly known probabilities, it would be natural to require a rationalizing utility function to be increasing with respect to first order stochastic dominance; there is a suitably modified version of GARP that could test if $\mathcal{D}$ admits such a rationalization. Theorem 1 can be similarly extended to characterize, via suitably modified versions of NCP, coarse rationalizability with respect to the types of utility families studied in Nishimura, Ok and Quah (2017). For details, see Hu et al. (2022).

Partial Congruence Axiom. Fishburn (1976) studies when $\mathcal{O} = \{(A^t, B^t)\}_{t \in T}$ can be rationalized by a preference $\succsim$ in the following sense: for all $t \in T$, there is $x^t \in A^t$ with $x^t \succsim B^t$ and $x^t \succ B^t \setminus A^t$. It shows that this holds if and only if $\mathcal{O}$ satisfies the *partial congruence axiom*, which requires

$$A(T') \not\subseteq \bigcup_{t \in T'} (B^t \setminus A^t) \text{ for all nonempty } T' \subseteq T. \tag{3}$$

Notice that Fishburn's rationalization concept is not comparable with coarse rationalization as characterized by Theorem 1. It is weaker in the sense that $\succsim$ is not required to be locally nonsatiated or strictly increasing, but it is stronger in the sense that it requires elements in $B^t \setminus A^t$ to be *non*-optimal. Correspondingly, the partial congruence axiom is neither stronger nor weaker than NCP. For example, a single observation (A^1, B^1) with A^1 in the interior of B^1 violates NCP but satisfies the partial congruence axiom. On the other hand, the dataset depicted in Figure 1a violates (3) (and hence the partial congruence axiom) for $T' = T$. This is consistent with the fact that x^1 and x^2 are in both B^1 and B^2 and thus there cannot be a preference $\succsim$ for which x^1 is optimal in B^1 but not x^2 and x^2 is optimal in B^2 but not x^1.[9]

Coarse Rationalization as an Alternative to Product Aggregation. In studies of consumer demand, a researcher would often not have information on the demand for every relevant good. One way of addressing this issue is to perform an aggregation procedure across goods, even though this approach is strictly valid only under stringent conditions on the utility function and/or the pattern of prices changes. To be specific, suppose that at observation t, the information available consists of the prices of all goods $p^t \in \mathbb{R}^n_{++}$, the demand for the first $m-1$ goods, and the total expenditure on the remaining goods (which we denote by $c^t_{m,n}$). In other words, the specific demands for goods m, $m+1, \ldots, n$ are not observed. To get round this problem, the researcher could construct a price index for those

[9] Fishburn's result allows T to be an infinite set and A^t and B^t to be nonempty sets in an arbitrary space of alternatives. We follow Afriat's Theorem by requiring T to be finite and A^t and B^t to be in a Euclidean space because we want to impose topological conditions on the rationalization. In our setting, rationalization is by a *continuous* utility function; this guarantees important features such as the existence of optima on compact sets.

goods, $\bar{p}_m^t$, which would be a function of their prices $(p_m^t, p_{m+1}^t, \ldots, p_n^t)$, with the corresponding demand for the composite good being $\bar{x}_m^t = c_{m,n}^t / \bar{p}_m^t$. In this way, the researcher creates a dataset of the standard form, with observation t consisting of the price vector $(p_1^t, p_2^t, \ldots, \bar{p}_m^t)$ and the demand $(x_1^t, x_2^t, \ldots, \bar{x}_m^t)$ for the m goods.

Coarse rationalization offers an alternative approach to tackle this problem. At observation t, since x_i^t for $i = 1, \ldots, m-1$ and $c_{m,n}^t$ are observed, the bundle chosen by the consumer must lie in the set

$$A^t = \left\{ x \in \mathbb{R}_+^n : x_i = x_i^t \text{ for } i = 1, \ldots, m-1 \text{ and } \textstyle\sum_{i=m}^n p_i x_i = c_{m,n}^t \right\}.$$

The corresponding coarse dataset is $\mathcal{O} = \{(A^t, B^t)\}_{t \in T}$, where

$$B^t = \left\{ x \in \mathbb{R}_+^n : p^t \cdot x \leq \sum_{i=1}^{m-1} p_i^t x_i^t + c_{m,n}^t \right\}.$$

Theorem 1 can be used to ascertain the rationalizability of this coarse dataset.

As an illustration, suppose that $\mathcal{O}$ consists of two observations where

$$\begin{aligned} &p^1 = (2, 2.5, 3.5), \quad x_1^1 = 1.5, \quad c_{2,3}^1 = 9; \\ &p^2 = (4, 3, 3), \qquad x_1^2 = 3, \quad\;\; c_{2,3}^2 = 4.5. \end{aligned}$$

This dataset is coarsely rationalizable. Indeed, $\tilde{x} = (1.5, 9/2.5, 0)$ is in A^1 but $p^2 \cdot \tilde{x} = 16.8 > 16.5$, so it is not in B^2. This guarantees that $\mathcal{O}$ satisfies NCP. On the other hand, suppose we aggregate goods 2 and 3 into a composite commodity, with the price of the composite being 3 (the average price of its constituent goods) at both observations 1 and 2. Then the demand for the composite good at these observations are $\bar{x}_2^1 = 9/3 = 3$ and $\bar{x}_2^2 = 4.5/3 = 1.5$. The corresponding two-good dataset has

$$\begin{aligned} &p^1 = (2, 3), \quad x^1 = (1.5, 3), \quad I^1 = 12; \\ &p^2 = (4, 3), \quad x^2 = (3, 1.5), \quad I^2 = 16.5. \end{aligned}$$

It is straightforward to check that this dataset violates GARP.

2.3 Algorithm

Given a subset T', it is usually straightforward to check whether $\Phi(T') = T'$. Thus, Theorem 1 provides a way of checking if a dataset is coarsely rationalizable: we need to check whether $\Phi(T') \neq T'$ for all $T' \subseteq T$. However, this is not promising as an empirical strategy, since for a dataset with m observations, we would have to go through all $2^m - 1$ nonempty subsets of T to guarantee NCP. In this subsection, we provide a simple algorithm to check whether NCP holds. This algorithm requires us to check whether $\Phi(T') \neq T'$ for at most m subsets of T. Thus, NCP can be checked in an efficient manner.

Following the convention in the computer science literature, we use k' to denote the updated value of a variable k.

Algorithm I. Set $T^0 := T$. Set $k := 1$.

START. Derive $T^k := \Phi(T^{k-1})$. Consider the following mutually exclusive cases:

(a). $T^k = \emptyset$: Stop and output *NCP holds.*

(b). $\emptyset \neq T^k \subsetneq T^{k-1}$: Go to START with $k' = k + 1$.

(c). $\emptyset \neq T^k = T^{k-1}$: Stop and output *NCP fails.*

Note that Algorithm I is effectively checking whether $\Phi(T^k) = T^k$ for an endogenous sequence of subsets of T. We emphasize that, for a dataset with m observations, Algorithm I necessarily terminates within m steps, and we only need to check at most m subsets of T.

Proposition 1 below provides the justification for Algorithm I.

Proposition 1. $\mathcal{O} = \{(A^t, B^t)\}_{t\in T}$ *satisfies NCP if and only if Algorithm I outputs* "NCP holds."

3 Dual Coarse Rationalizability

The notion of coarse rationalizability is founded on the imperfect observability of choice: instead of observing exactly the bundle chosen from a budget set B^t, the observer could only tell that it is drawn from $A^t \subseteq B^t$. But it is also possible for observations of prices to be imperfect and, in this section, we explain how this leads to a dual notion of coarse rationalizability.

We consider a dataset of the form

$$\mathcal{O}^* = \Big\{(x^t, \{L(p^{t,s})\}_{s\in G_t})\Big\}_{t\in T}$$

where $T \neq \emptyset$ is finite and, for every $t \in T$, $x^t \in \cap_{s \in G_t} L(p^{t,s})$. We interpret x^t as the choice observed to be made by a consumer in observation t and $\{L(p^{t,s})\}_{s \in G_t}$ as the collection of possible budget sets from which the consumer's choice was made. In other words, the researcher observes perfectly the bundle chosen by the consumer but may have imprecise information on the corresponding price vector. We are interested in identifying those datasets where there is a preference such that, at each t, the bundle x^t is optimal according to that preference for some budget set among those in $\{L(p^{t,s})\}_{s \in G_t}$. This is stated formally as follows.

Definition 3. *$\mathcal{O}^* = \{(x^t, \{L(p^{t,s})\}_{s \in G_t})\}_{t \in T}$ has a* dual coarse rationalization *by the preference $\succsim$ on $\mathbb{R}^n_+$ if there exists a selection $s_t \in G_t$ for every $t \in T$ such that for all $t \in T$,*

$$x^t \in \max\left(L(p^{t,s_t}); \succsim\right).$$

We refer to $\mathcal{O}^*$ as being dual coarsely rationalized by a utility function U if it is dual coarsely rationalized by the preference induced by U.

We are interested in the non-trivial case in which the preference $\succsim$ that dual-coarsely rationalizes the dataset $\mathcal{O}^*$ is locally nonsatiated. Therefore, without loss of generality, we can consider a dataset such that for every $t \in T$ and $s \in G_t$, $x^t \cdot p^{t,s} = 1$: whenever we have $x^t \cdot p^{t,\bar{s}} < 1$, we can simply remove $L(p^{t,\bar{s}})$ from the set of candidate budget sets at observation t, since local nonsatiation implies that x^t cannot be the optimal bundle in $L(p^{t,\bar{s}})$ in such a case.

To motivate our next result, suppose that $\mathcal{O}^* = \{(x^t, \{L(p^{t,s})\}_{s \in G_t})\}_{t \in T}$ admits a dual coarse rationalization by a locally nonsatiated preference. This means that there is a selection $L(p^{t,s_t})$ from $\{L(p^{t,s})\}_{s \in G_t}$ such that the standard dataset $\mathcal{D} = \{(x^t, L(p^{t,s_t}))\}_{t \in T}$ can be rationalized by a locally nonsatiated preference. But Afriat's Theorem guarantees that $\mathcal{D}$ can also be rationalized by a regular utility function U; furthermore, with no loss of generality, we can choose U to be bounded above by $\bar{u} \in \mathbb{R}$.[10] For each $p \in \mathbb{R}^n_{++}$, we can associate it with an indirect utility

[10] The boundedness of U can be ensured by taking the monotone and continuous transformation $\arctan(\cdot)$ for U.

level, denoted by $V(p)$, such that

$$V(p) = \max_{x \in L(p)} U(x).$$

We further extend V to the whole domain of $\mathbb{R}^n_+$ such that for all $\hat{p} \in \mathbb{R}^n_+ \backslash \mathbb{R}^n_{++}$, $V(\hat{p}) = \bar{u} + 1$, and define $W : \mathbb{R}^n_+ \to \mathbb{R}$ such that $W = -V$.

It is straightforward to check that W is locally nonsatiated.[11] Furthermore, for every $t \in T$ and $p \in \mathbb{R}^n_{++}$ with $x^t \cdot p \leq 1$, we have

$$V(p^{t,s_t}) = U(x^t) \leq \max_{x \in L(p)} U(x) = V(p),$$

where the first equality is by the fact that U rationalizes $\mathcal{D}$, and the inequality holds since $x^t \in L(p)$ whenever $x^t \cdot p \leq 1$. In other words, p^{t,s_t} is the vector that maximizes W in $\{p \in \mathbb{R}^n_+ : x^t \cdot p \leq 1\}$.

Therefore, a *necessary* condition for $\mathcal{O}^*$ to admit a dual coarse rationalization by a locally nonsatiated preference is the existence of a locally nonsatiated utility function W that coarsely rationalizes the dataset

$$\mathcal{O}^{**} := \Big\{ \{p^{t,s}\}_{s \in G_t}, \{p \in \mathbb{R}^n_+ : x^t \cdot p \leq 1\} \Big\}_{t \in T};$$

less obviously, we show that this property on $\mathcal{O}^{**}$ is also *sufficient* for the dual coarse rationalization of $\mathcal{O}^*$. This characterization gives us a test of the dual coarse rationalizability of $\mathcal{O}^*$ since we know (from Theorem 2) that the coarse rationalizability of $\mathcal{O}^{**}$ can be checked through the never-covered property. The following theorem states these claims formally.

Theorem 3. *The following statements on* $\mathcal{O}^* = \{(x^t, \{L(p^{t,s})\}_{s \in G_t})\}_{t \in T}$ *are equivalent:*

(1) $\mathcal{O}^*$ *can be dual coarsely rationalized by a locally nonsatiated preference.*

(2) $\mathcal{O}^{**}$ *satisfies NCP.*

(3) $\mathcal{O}^*$ *can be dual coarsely rationalized by a strictly increasing, continuous and concave utility function.*

Example 2. The dataset $\mathcal{O}^*$ depicted in Figure 2a contains two observations:

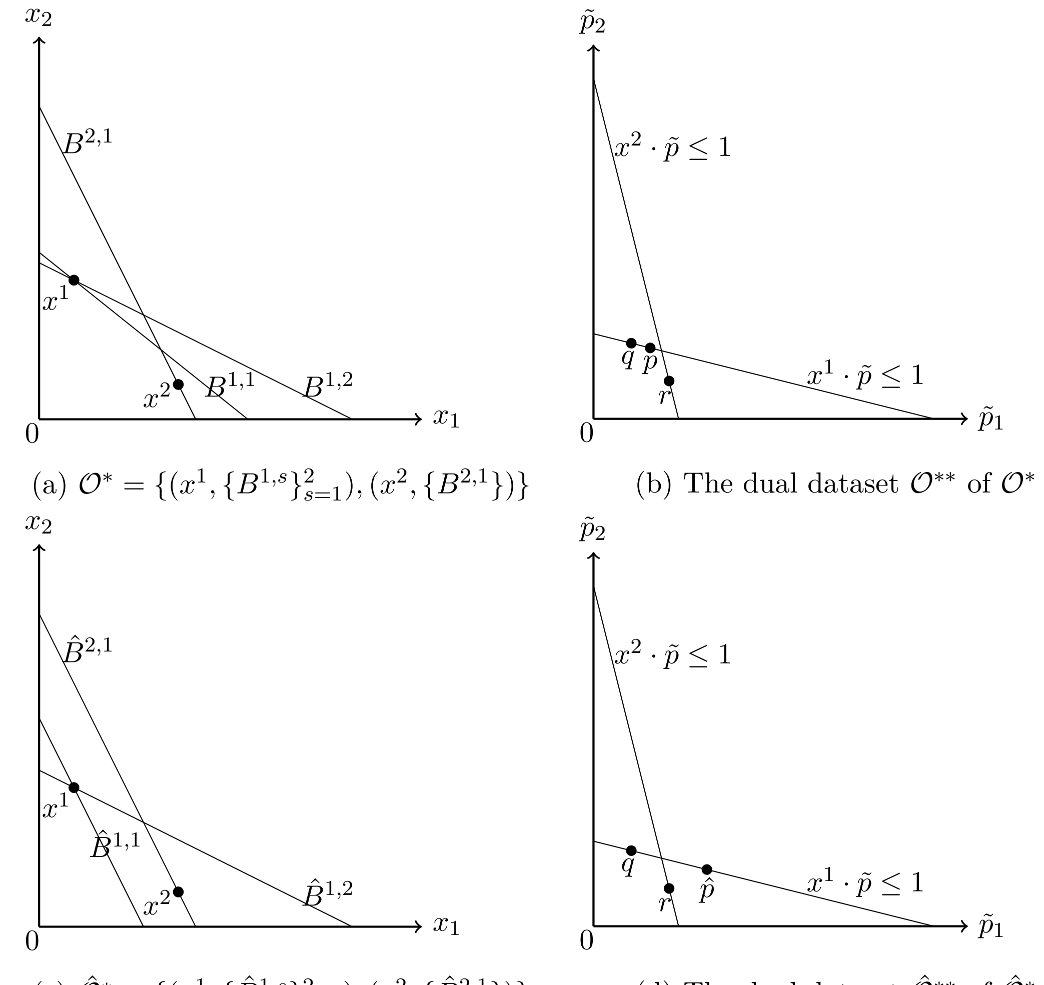
[11] It suffices to show that for all $p, \hat{p} \in \mathbb{R}^n_+$, if $p \gg \hat{p}$, then $W(p) > W(\hat{p})$. Indeed, $p \gg \hat{p}$ implies either (i) $\hat{p} \in \mathbb{R}^n_+ \backslash \mathbb{R}^n_{++}$ and $p \in \mathbb{R}^n_{++}$, indicating $W(p) \geq -\bar{u} > -\bar{u} - 1 = W(\hat{p})$, or (ii) $\hat{p}, p \in \mathbb{R}^n_{++}$, which further implies $W(p) > W(\hat{p})$ as a result of the monotonicity of U.

$(x^1, \{B^{1,1}, B^{1,2}\})$ and $(x^2, \{B^{2,1}\})$, where

$$x^1 = (1,4), \quad B^{1,1} = L(p) = L(\tfrac{1}{6}, \tfrac{5}{24}), \quad B^{1,2} = L(q) = L(\tfrac{1}{9}, \tfrac{2}{9}); \text{ and}$$

$$x^2 = (4,1), \quad B^{2,1} = L(r) = L(\tfrac{2}{9}, \tfrac{1}{9}).$$

Note that no matter whether $B^{1,1}$ or $B^{1,2}$ is the true budget set, we can always reveal that x^1 is strictly better than x^2 and vice versa, indicating that $\mathcal{O}^*$ is not dual-coarsely rationalizable by a locally non-satiated preference. The dual dataset $\mathcal{O}^{**}$ of $\mathcal{O}^*$ is depicted in Figure 2b. Since r is in the interior of $\{\tilde{p} \in \mathbb{R}^n_+ : x^1 \cdot \tilde{p} \leq 1\}$ and both p and q are in that of $\{\tilde{p} \in \mathbb{R}^n_+ : x^2 \cdot \tilde{p} \leq 1\}$, we conclude that NCP fails for $\mathcal{O}^{**}$.



(a) $\mathcal{O}^* = \{(x^1, \{B^{1,s}\}_{s=1}^2), (x^2, \{B^{2,1}\})\}$

(b) The dual dataset $\mathcal{O}^{**}$ of $\mathcal{O}^*$

(c) $\hat{\mathcal{O}}^* = \{(x^1, \{\hat{B}^{1,s}\}_{s=1}^2), (x^2, \{\hat{B}^{2,1}\})\}$

(d) The dual dataset $\hat{\mathcal{O}}^{**}$ of $\hat{\mathcal{O}}^*$

Figure 2: Dual Coarse Rationalization

Next, consider the dataset $\hat{\mathcal{O}}^*$ depicted in Figure 2c with two observations:

$(x^1, \{\hat{B}^{1,s}\}_{s=1}^2)$ and $(x^2, \{\hat{B}^{2,1}\})$, where

$$\hat{B}^{1,1} = L(\hat{p}) = L(\tfrac{1}{3}, \tfrac{1}{6}),\ \hat{B}^{1,2} = L(q) = L(\tfrac{1}{9}, \tfrac{2}{9});\text{ and}$$
$$\hat{B}^{2,1} = L(r) = L(\tfrac{2}{9}, \tfrac{1}{9}).$$

That is, $\hat{\mathcal{O}}^*$ is the same as $\mathcal{O}^*$ except that we replace p in $B^{1,1}$ with $\hat{p}$. By selecting $\hat{B}^{1,1}$, one can show that x^2 is revealed preferred to x^1 while the inverse is not true. Thus, the dataset $\hat{\mathcal{O}}^*$ is dual coarsely rationalizable by a locally nonsatiated preference. In Figure 2d, the dual dataset $\hat{\mathcal{O}}^{**}$ of $\hat{\mathcal{O}}^*$ satisfies NCP as $\hat{p}$ is clearly not covered. □

4 Rationality Indices

For a standard dataset of the form $\mathcal{D} = \{(x^t, L(p^t))\}_{t\in T}$, where $x^t \in \mathbb{R}^n_+$ is the observed choice from the budget set $L(p^t) = \{x \in \mathbb{R}^n_+ : p^t \cdot x \leq 1\}$, Afriat's Theorem tells us that GARP is a necessary and sufficient condition for $\mathcal{D}$ to be rationalized by a regular utility function. However, in most empirical applications, it is common for subjects to fail GARP. Various indices have been proposed to measure the severity of a subject's departure from rationality, with the most commonly used measure being the *critical cost efficiency index* (CCEI) due to Afriat (1973).[12] The CCEI is defined as

$$e^* := \sup\{e : \mathcal{D} \text{ is rationalized at efficiency level } e \text{ by a regular utility function}\},$$

where a utility function U *rationalizes* $\mathcal{D}$ *at cost efficiency level* $e \in (0,1]$ if $U(x^t) \geq U(x)$ for $x \in \mathbb{R}^n_+$ that satisfies $p^t \cdot x \leq e$. Obviously, if $\mathcal{D}$ is rationalized by a regular utility function, then its CCEI is 1. One reason for the popularity of this index is its simplicity: it is easy to calculate because rationalizability at any efficiency level e can be ascertained by a modified version of GARP.

In this section, we discuss two alternative rationality indices that, following from our results in the earlier sections, are also easy to compute and which also constitute natural ways of measuring departures from rationality.

[12] Papers that use this concept or a related version due to Varian (1990) include Harbaugh, Krause and Berry (2001), Andreoni and Miller (2002), Choi et al. (2007), Choi et al. (2014), Fisman, Kariv and Markovits (2007), Carvalho, Meier and Wang (2016), and Halevy, Persitz and Zrill (2018).

4.1 Perturbation Index

A natural way of measuring the severity of departures from rationality is to measure the extent to which the chosen bundles x^t have to be perturbed before the perturbed dataset becomes rationalizable. Such a measure was used by Varian (1985) in his analysis of production data, where the chosen bundles are bundles of factor demand. In that context, it is natural to assume that output levels at each bundle of factors are observable, which considerably simplifies the calculation of such an index, whereas the same exercise in the context of consumption is more complicated, essentially because it is *not* reasonable to assume that utility levels are observable (see Appendix for more details). Nonetheless, our extension of Afriat's Theorem (with Algorithm I for checking NCP) makes it feasible to implement such a rationality measure.

To be precise, suppose the researcher records the consumer choosing x^t from $L(p^t)$; to accommodate the possibility that x^t was observed with error, the researcher could allow the true consumption bundle to be in the set

$$A^{t,\kappa} := \left\{ x \in L(p^t) : \ p^t \cdot x = 1 \text{ and } |p^t_i x_i - p^t_i x^t_i| \leq \kappa \text{ for all } i \right\}, \tag{4}$$

where $\kappa \in [0, 1]$. In other words, the true expenditure on good i is allowed to deviate from $p^t_i x^t_i$ but not by more than the fraction κ of income. In experimental settings, where there is no question that x^t is indeed the observed choice, we could interpret κ as a measure of the extent to which we allow the subject to make mistakes. Whatever the interpretation, we can test, via NCP, whether there is a regular utility function $U : \mathbb{R}^n_+ \to \mathbb{R}$ that coarsely rationalizes

$$\mathcal{O}^\kappa = \left\{ \left(A^{t,\kappa}, L(p^t) \right) \right\}_{t \in T}.$$

If so, then there are bundles $\tilde{x}^t$, such that $|p^t_i \tilde{x}^t_i - p^t_i x^t_i| \leq \kappa$ for all i and t, and $\widetilde{\mathcal{D}} = \{(\tilde{x}^t, L(p^t)\}_{t \in T}$ is rationalized by U. This is illustrated in Figure 3, where the actual dataset $\{(x^1, L^1), (x^2, L^2)\}$ is not rationalizable, but $\{(A^1, L^1), (A^2, L^2)\}$ (as depicted) is coarsely rationalizable. Indeed, so long as we choose $y \in A^1 \setminus L^2$, then $\mathcal{D} = \{(y, L^1), (x^2, L^2)\}$ obeys GARP and is rationalizable by a regular utility function.

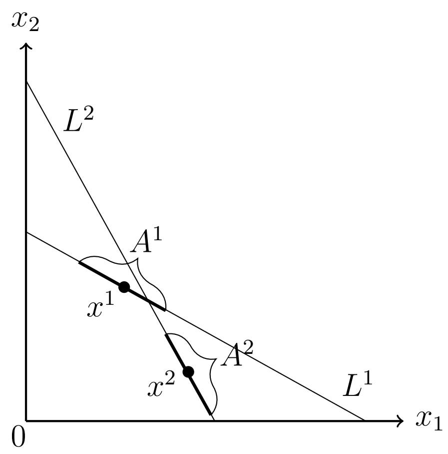


Figure 3: $\{(x^1, L^1), (x^2, L^2)\}$ is not rationalizable, but $\{(A^1, L^1), (A^2, L^2)\}$ (as depicted) is coarsely rationalizable.

Since we can determine the coarse rationalizability of $\mathcal{O}^\kappa$, we can calculate the *critical perturbation index* (or *perturbation index*, for short)

$$\kappa^* := \inf\{\kappa : \mathcal{O}^\kappa \text{ is coarsely rationalizable by a regular utility function}\}$$

by binary interpolation.[13] The index gives the smallest perturbation needed to guarantee that the coarsened dataset $\mathcal{O}^\kappa$ admits a coarse rationalization by a regular utility function. Obviously, if $\mathcal{D}$ is rationalizable to begin with, then its perturbation index equals zero.[14]

As we indicated in Section 2.2, there are modified versions of NCP that could test if a coarse dataset such as $\mathcal{O}^\kappa$ admits a coarse rationalization by a utility function belonging to certain families. Thus perturbation indices with respect to such utility families could also be calculated.

4.1.1 Computing the Perturbation Index

By Theorems 1 and 2, $\mathcal{O}^\kappa$ can be coarsely rationalized by a regular utility function if and only if NCP holds. The latter could be ascertained using Algorithm I, which requires checking whether $T' = \Phi(T')$ for a nested sequence of observations $T' \subseteq T$.

[13] If $\kappa = 1$, then $A^t = L(p^t) \setminus \mathring{L}(p^t)$ for all $t \in T$ and $\mathcal{O}^1$ is coarsely rationalizable. Thus the perturbation index is always well-defined.

[14] There are other implementable variations on the perturbation index. These are discussed in the Appendix. The Appendix also discusses an alternative way of computing the perturbation index via the Afriat inequalities.

For a given T', the check of whether $T' = \Phi(T')$ requires calculating an increasing sequence of subsets $\Phi^1(T')$, $\Phi^2(T')$, etc, that terminates at $\Phi(T')$.

Implementing this procedure is straightforward because checking if an observation belongs to $\Phi^m(T')$ (for a given m) reduces to checking if there is a solution to a system of linear inequalities. Recall that

$$\Phi^1(T') = \left\{ t \in T' : A^t \subseteq \bigcup_{t \in T'} \mathring{L}(p^t) \right\}.$$

We calculate $\Phi^1(T')$ by checking whether each A^s (for $s \in T'$) is contained in $\bigcup_{t \in T'} \mathring{L}(p^t)$ and this in turn can be verified by checking if there is a solution $z \in \mathbb{R}^n$ to the following system of linear inequalities:

$$\begin{aligned} z &\geq 0 \\ p^s \cdot z &= p^s \cdot x^s \\ |p_i^s z_i - p_i^s x_i^s| &\leq \kappa \quad \text{for each good } i \\ p^t \cdot z &\geq 1 \quad \text{for all } t \in T'. \end{aligned} \tag{5}$$

Clearly, $A^s \subseteq \bigcup_{t \in T'} \mathring{L}(p^t)$ if and only if there is *no* solution to this system of linear inequalities.

Having calculated $\Phi^1(T')$ we can then calculate $\Phi^2(T')$ (which contains $\Phi^1(T')$), where

$$\Phi^2(T') := \left\{ t \in T' : A^t \subseteq \left(\bigcup_{t \in T'} \mathring{L}(p^t) \right) \bigcup \left(\bigcup_{t \in \Phi^1(T')} L(p^t) \right) \right\}.$$

This can be done using a similar inequality system, with (5) replaced by

$$p^t \cdot z \geq 1 \text{ for all } t \in T' \qquad \text{and} \qquad p^t \cdot z > 1 \text{ for all } t \in \Phi^1(T').$$

We continue in this manner until $\Phi^m(T') = \Phi^{m+1}(T')$ at which point we can check whether $\Phi(T') := \Phi^m(T')$ is a strict subset of T'.

Checking whether $\mathcal{O}^\kappa$ satisfies NCP can be accomplished in *polynomial time*. This is because, for each $m = 1, 2, \ldots$, calculating $\Phi^m(T')$ requires us to solve at most $|T'| + 1 - m$ systems of linear inequalities. Thus establishing whether $T' = \Phi(T')$ involves solving at most $|T'|(|T'|+1)/2$ linear problems and establishing if $\mathcal{O}^\kappa$ satisfies NCP involves solving no more than

$$\frac{|T|(|T|+1)}{2} + \frac{(|T|-1)|T|}{2} + \frac{(|T|-2)(|T|-1)}{2} + \ldots = \frac{|T|(|T|+1)(|T|+2)}{6}$$

linear problems.[15]

4.1.2 Empirical Illustration

We study the data collected from the portfolio choice experiment in Choi et al. (2007). The experiment was performed on 93 undergraduate subjects at UC Berkeley. Every subject was asked to make choices in 50 decision problems. In each problem, the subject divided her budget between two Arrow-Debreu securities, with each security paying 1 token (equivalent to US$0.50) if the corresponding state was realized, and 0 otherwise. We focus on the symmetric treatment where each state of the world occurred with a commonly known probability of 1/2. This treatment was applied to 47 subjects (subject ID 201-219 and 301-328). The prices of the Arrow-Debreu securities were chosen at random (over some compact interval) and varied across problems and subjects, with income normalized at 1 throughout.

For each state $s \in \{1, 2\}$, let x_s denote the demand for the security that pays off in that state and let p_s denote its price. For each subject and in each decision problem $t \in T = \{1, \ldots, 50\}$, the state prices $p^t = (p_1^t, p_2^t)$ were randomly chosen and the subject faced a budget set

$$L(p^t) = \{x \in \mathbb{R}_+^2 : p_1^t x_1 + p_2^t x_2 \leq 1\}.$$

Thus the set of observations collected from a subject can be written as $\mathcal{D} = \{(x^t, L(p^t))\}_{t=1}^{50}$, where x^t is the subject's choice in $L(p^t)$.

In calculating the perturbation index, we apply Algorithm I on $\mathcal{O}^\kappa$ for different values of κ; the index can then be obtained by binary interpolation. As an illustration, Table 1 shows the steps involved when Algorithm I is applied to Subject 201's data. The algorithm involves calculating the set of revealed dominated observations $T^1 := \Phi(T)$ and then checking whether $T^1 = T$; if not, it calculates $T^2 := \Phi(T^1)$ and checks whether $T^2 = T^1$; and so forth until either $T^k = T^{k-1}$ (in which case $\mathcal{O}^\kappa$ fails to satisfy NCP) or $T^k = \emptyset$ (in which case $\mathcal{O}^\kappa$ satisfies NCP). T^1, T^2, $T^3 \ldots$ form a nested sequence of sets; the number of elements in each set is

[15] For algorithms to check the solvability of a system of linear inequalities see Karmarkar (1984).

indicated in Table 1. We see that when $\kappa = 0.2$, $\mathcal{O}^\kappa$ fails NCP because $T^{11} = T^{10}$ and is nonempty while $\mathcal{O}^\kappa$ satisfies NCP when $\kappa = 0.3$, because T^{16} is empty. Thus κ^* lies between 0.2 and 0.3. Indeed, by binary interpolation, we find that $\kappa^* = 0.2151$.

	$\kappa = 0.2$	$\kappa = 0.3$
$\lvert T^1 \rvert$	48	47
$\lvert T^2 \rvert$	47	46
$\lvert T^3 \rvert$	45	44
$\lvert T^4 \rvert$	41	39
$\lvert T^5 \rvert$	39	36
$\lvert T^6 \rvert$	37	31
$\lvert T^7 \rvert$	33	25
$\lvert T^8 \rvert$	27	20
$\lvert T^9 \rvert$	21	18
$\lvert T^{10} \rvert$	18	16
$\lvert T^{11} \rvert$	18	13
$\lvert T^{12} \rvert$		8
$\lvert T^{13} \rvert$		4
$\lvert T^{14} \rvert$		2
$\lvert T^{15} \rvert$		1
$\lvert T^{16} \rvert$		0

Table 1: Testing NCP on Subject 201.

Similar calculations are carried out for the other 46 subjects. We find that 12 subjects pass the test exactly and so have an index of 0. The median value of the index is 0.0778, with 0.1504 and 0 being the 75th and 25th percentiles respectively. The cumulative distribution of the perturbation index (κ^*) is depicted in Figure 4. For each $r \in [0, 1]$, we plot the percentage of subjects whose index are less than or equal to r.

How does the index compare with Afriat's CCEI? Since the CCEI is increasing with rationality while the opposite holds for the perturbation index, the two indices are naturally negatively correlated. We find that the rank correlation coefficient is

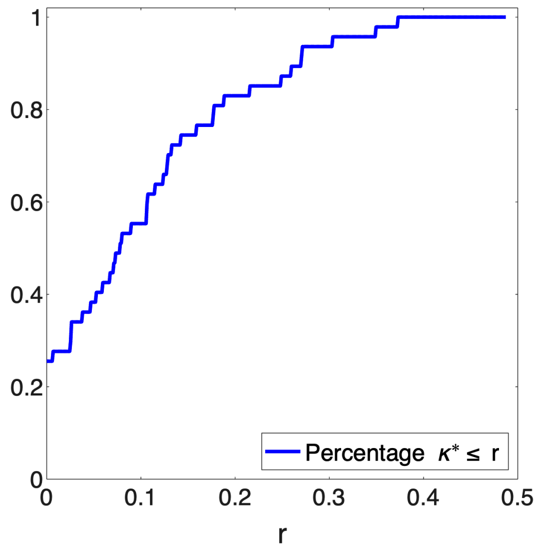


Figure 4: Distribution of κ^*.

-0.919 while the linear correlation coefficient is -0.795 (using the CCEI calculated for these subjects in Polisson, Quah and Renou (2020)). So the two indices are correlated, but not perfectly. In empirical work where conclusions are drawn based on the CCEI, one could use the perturbation index as a way of checking the robustness of those conclusions. For example, there is evidence that rationality, as measured by the CCEI, can be helpful in explaining broader economic outcomes, including family wealth (see Choi et al. (2014)); it is interesting to check whether conclusions such as these are sensitive to the rationality index used.

4.2 FOC-Departure Index

In this section, we demonstrate how our results can be applied to compute the first order condition (FOC)-departure index introduced and axiomatized by de Clippel and Rozen (2023). In contrast to the perturbation index, where the bundles are perturbed in order to restore rationality, this index could be thought of as measuring a dataset's closeness to rationality by measuring the extent to which prices have to be perturbed to recover rationality.

Consider an observation $(x^t, L(p^t))$ where the consumer chooses bundle x^t under price vector p^t. If x^t is an interior bundle in $\mathbb{R}^n_+$ and the consumer is maximizing a regular and quasi-concave utility function U that is differentiable at

x^t with $g = \partial u(x_t)$ being the corresponding gradient, then the first order condition requires $p = g$. By contrast, if the consumer's choice x^t is not the optimal one in $L(p^t)$, then $p \neq g$. Motivated by this observation, de Clippel and Rozen (2023) propose to use the function

$$\delta(p, g) = \max_{i,j \in \{1,\ldots,n\}: i \neq j} \left\{ \frac{p_i/p_j}{g_i/g_j}, \frac{g_i/g_j}{p_i/p_j}. \right\} \tag{6}$$

as a measure of the consumer's departure from rationality. Note that δ is always greater than or equal to 1: $\delta = 1$ means that there is no deviation from rationality, while a larger value of δ corresponds to a larger deviation from rationality.

A notion of approximate rationalizability can then be developed based on δ. A dataset $\mathcal{D} = \{(x^t, L(p^t))\}_{t \in T}$ is said to be *ϵ-rationalizable* if there is a regular and quasi-concave utility function U such that for every $t \in T$, there is a vector g in the quasi-gradient set $\partial u(x^t)$[16] such that

$$\delta(p^t, g) \leq \frac{1}{1 - \epsilon}. \tag{7}$$

Afriat's Theorem tells us that $\mathcal{D}$ can be rationalized by a regular and concave utility function if and only if it satisfies GARP; in this case, $\mathcal{D}$ is 0-rationalizable. When $\mathcal{D}$ is not 0-rationalizable, we can enlarge ϵ until we can find some regular and quasi-concave U that ϵ-rationalizes $\mathcal{D}$. The *FOC-departure index*, $\text{FDI}(\mathcal{D})$, is defined as the infimum of such ϵ, i.e.,

$$\text{FDI}(\mathcal{D}) := \inf\{\epsilon \in [0, 1] : \mathcal{D} \text{ is } \epsilon\text{-rationalizable by a regular and quasi-concave } U\}.$$

Following de Clippel and Rozen (2023), we can relate the FOC-departure index with the *price misperception index*. For a given dataset $\mathcal{D}$, the price misperception index of $\mathcal{D}$, denoted by $\text{PMI}(\mathcal{D})$, is the infimum over all $\epsilon \in [0, 1]$ that satisfies the

[16] The quasi-gradient set is defined as

$$\partial u(x^t) = \{g \in \mathbb{R}^n_{++} : \forall y \in \mathbb{R}^n_{+}, u(y) \geq u(x^t) \Rightarrow g \cdot y \geq g \cdot x^t\}.$$

following condition:

$$\begin{aligned}
&\exists \text{ regular and quasi-concave } U \text{ and } \{\hat{p}^t\}_{t\in T} \subseteq \mathbb{R}^n_{++} \text{ such that} \\
&\forall t \in T,\ x^t \cdot \hat{p}^t = 1, \\
&\forall t \in T, \forall i,j \in \{1,...,n\} \text{ with } i \neq j,\ 1-\epsilon \leq \frac{p^t_i/p^t_j}{\hat{p}^t_i/\hat{p}^t_j} \leq \frac{1}{1-\epsilon}, \\
&\forall y \in L(\hat{p}^t),\ U(x^t) \geq U(y).
\end{aligned} \tag{8}$$

We can interpret $\hat{p}^t$ as the misperceived price vector of the bundle x^t, so that the consumer is maximizing utility with respect to those misperceived prices rather than the true prices p^t. The price misperception index is the minimum misperception we must allow the consumer in order for the data to be rationalized by some regular and quasi-concave utility function.

Proposition 2 of de Clippel and Rozen (2023) shows that $\text{PMI}(\mathcal{D}) = \text{FDI}(\mathcal{D})$ for all datasets $\mathcal{D}$. This identity is useful because calculating $\text{PMI}(\mathcal{D})$ — and hence $\text{FDI}(\mathcal{D})$ — is straightforward given our results. Indeed, for a given ϵ, checking whether condition (8) holds is equivalent to checking the dual coarse rationalizability of the dataset

$$\mathcal{O}^*_\epsilon = \left\{ (x^t, \{L(\hat{p}^t)\}_{\hat{p}^t \in Z_{\delta,\epsilon}(p^t)}) \right\}_{t\in T}, \tag{9}$$

where for each t,

$$Z_{\delta,\epsilon}(p^t) := \left\{ \hat{p}^t \in \mathbb{R}^n_{++} : \delta(p^t, \hat{p}^t) \leq \frac{1}{1-\epsilon} \text{ and } x^t \cdot \hat{p}^t = 1 \right\}.$$

By Theorem 3, this is equivalent to the coarse rationalizability of

$$\mathcal{O}^{**}_\epsilon = \left\{ (Z_{\delta,\epsilon}(p^t), L(x^t)) \right\}_{t\in T},$$

where $L(x^t) := \{p \in \mathbb{R}^n_+ : x^t \cdot p \leq 1\}$, which is in turn equivalent to checking that $\mathcal{O}^{**}_\epsilon$ obeys NCP.

Using Algorithm I, we know that NCP can be checked in polynomial time. Indeed, the procedure outlined in Section 4.1.1 applies here, mutatis mutandi. In particular, suppose that we want to check whether $\Phi(T') = T'$ for a nonempty subset $T' \subseteq T$ and assume that we have obtained $\Phi^k(T')$ with $\Phi^k(T') \subsetneq T'$. Then the set $\Phi^{k+1}(T')$ is given by

$$\Phi^{k+1}(T') = \left\{ t \in T' : Z_{\delta,\epsilon}(p^t) \subseteq \left(\bigcup_{t\in T'} \mathring{L}(x^t) \right) \bigcup \left(\bigcup_{t\in \Phi^k(T')} L(x^t) \right) \right\}. \tag{10}$$

To check whether some given $s \in T'$ belongs to $\Phi^{k+1}(T')$, we can check if there is a solution $p \in \mathbb{R}^n$ to the following system of inequalities which, crucially, are linear in p:

$$\begin{aligned} p &\geq 0 \\ x^s \cdot p &= 1 \\ (1-\epsilon)p_j \cdot \frac{p_i^s}{p_j^s} &\leq p_i \leq \frac{p_j}{1-\epsilon} \cdot \frac{p_i^s}{p_j^s} \quad \text{for all goods } i \neq j \\ x^t \cdot p &\geq 1 \quad \text{for all } t \in T' \\ x^t \cdot p &> 1 \quad \text{for all } t \in \Phi^k(T'). \end{aligned} \tag{11}$$

Clearly, $t \in \Phi^{k+1}(T')$ if and only if there is *no* solution to this system of linear inequalities.[17]

Appendix

Proof of Theorem 1. Since every increasing utility function generates a increasing, thus locally nonsatiated, preference, (3) implies (1). We have already shown in the main text that (1) implies (2). It remains to show that (2) implies (3). We do this by first setting out the procedure with which x^t can be chosen from A^t so that $\mathcal{D} = \{(x^t, B^t)\}_{t \in T}$ obeys GARP.

Denote by $\mathcal{E}(T')$ the set of bundles that are revealed to be dominated through the procedure of iterated exclusion of dominated observations, i.e.,

$$\mathcal{E}(T') := \mathring{B}(T') \bigcup B\left(\Phi(T')\right).$$

Since $\mathcal{O}$ satisfies NCP, for any nonempty $T' \subseteq T$, $\Phi(T')$ is a strict subset of T', which implies that $A\left(T'\right) \setminus \mathcal{E}(T') \neq \emptyset$.

Let $T_1 := T$ and $S_1 := A(T_1) \setminus \mathcal{E}(T_1)$. We proceed by induction. Suppose that we have constructed T_k and S_k for some $k \geq 1$. If $T_k \neq \emptyset$, construct T_{k+1} and S_{k+1}:

$$T_{k+1} := \Phi(T_k) = \{t \in T_k : A^t \subseteq \mathcal{E}(T_k)\} \text{ and } S_{k+1} := A(T_{k+1}) \setminus \mathcal{E}(T_{k+1}).$$

Since $\mathcal{O}$ satisfies NCP, if $T_k \neq \emptyset$, then $T_{k+1} = \Phi(T_k) \subsetneq T_k$ and $S_k = A(T_k) \setminus \mathcal{E}(T_k) \neq \emptyset$. The construction stops when $T_{k^*} \neq \emptyset$ and $T_{k^*+1} = \emptyset$ for some k^*.

[17] See the Appendix for a discussion of alternative methods for calculating PMI/FDI.

We are now ready to select x^t in A^t for each $t \in T$ such that $\mathcal{D} = \{(x^t, B^t)\}_{t\in T}$ obeys GARP. For each $1 \leq k \leq k^*$, let $V_k := T_k \setminus T_{k+1}$ denote the collection of observations that are eliminated when constructing T_{k+1} from T_k. Clearly, $\{V_k\}_{k=1}^{k^*}$ is a partition of T. By definition, for each k and each $t \in V_k = T_k \setminus T_{k+1}$, we have $A^t \setminus \mathcal{E}(T_k) \neq \emptyset$ and hence $A^t \cap S_k = A^t \cap (A(T_k) \setminus \mathcal{E}(T_k)) \neq \emptyset$.

For each k and each $t \in V_k = T_k \setminus T_{k+1}$, we pick an arbitrary $x^t \in A^t \cap S_k$. We claim that $\mathcal{D} = \{(x^t, B^t)\}_{t\in T}$ obeys GARP. Let $k(t)$ be the corresponding index k such that $t \in V_k$. It suffices to show that (1) $x^t R x^{t'}$ implies that $k(t) \leq k(t')$; and (2) $x^t P x^{t'}$ implies that $k(t) < k(t')$. Suppose that $x^t R x^{t'}$ but $k(t) > k(t')$. Then $t \in \Phi(T_{k(t')})$ due to the construction of $\{V_k\}_{k=1}^{k^*}$. It follows that $A^t \subseteq \mathcal{E}(T_{k(t')})$ and $B^t \subseteq \mathcal{E}(T_{k(t')})$. Since $x^t R x^{t'}$, we have $x^{t'} \in B^t \subseteq \mathcal{E}(T_{k(t')})$, which contradicts with $x^{t'} \in S_{k(t')} = A(T_{k(t')}) \setminus \mathcal{E}(T_{k(t')})$. Hence, $x^t R x^{t'}$ implies $k(t) \leq k(t')$. Suppose that $x^t P x^{t'}$ but $k(t) \geq k(t')$. If $k(t) > k(t')$, then we have the same contradiction as argued above. If $k(t) = k(t') = k$ for some k, then both x^t and $x^{t'}$ belong to S_k. Since $S_k = A(T_k) \setminus \mathcal{E}(T_k)$ and $\mathring{B}(T_k) \subseteq \mathcal{E}(T_k)$, we have

$$x^t, x^{t'} \in S_k \subseteq A(T_k) \setminus \mathring{B}(T_k).$$

But this is impossible since $x^t P x^{t'}$ implies $x^{t'} \in \mathring{B}^t$. Hence, $x^t P x^{t'}$ implies $k(t) < k(t')$.

By Proposition 3 of Forges and Minelli (2009), there are numbers v^t and $\lambda^t > 0$ such that, for all t, $s \in T$,

$$v^s \leq v^t + \lambda^t g^t(x^s),$$

and the utility function

$$U(x) = \min_{t\in T}\{v^t + \lambda^t g^t(x)\} \tag{12}$$

rationalizes $\mathcal{D}$ and hence coarsely rationalizes $\mathcal{O}$. Since g^t are continuous, so is U, and since g^t are increasing, U is also increasing. □

Proof of Theorem 2. Given Theorem 1, the only part of this result that still needs proving is the claim that the utility function that coarsely rationalizes $\mathcal{O}$ can be chosen to be *strictly* increasing, continuous *and concave*. But this is clear from the form of the utility function (12). Indeed, $L(p^t) = \{x \in \mathbb{R}^n_+ : g^t(x) \leq 0\}$, where

$g^t(x) = p^t \cdot x - 1$. Since g is continuous, strictly increasing, and linear, U is strictly increasing, continuous and concave. $\square$

Proof of Proposition 1. **"Only if":** If $\mathcal{O}$ satisfies NCP, then $\Phi(T') \neq T'$ for any nonempty $T' \subseteq T$. Thus, Case (c) never occurs when we run Algorithm I on this dataset. Furthermore, T^k is strictly decreasing in k in the set inclusion sense, and $T^{k^*} = \emptyset$ for some k^*. Therefore, Algorithm I outputs *NCP holds*.

"If": We first claim that the operator $\Phi(\cdot)$ is monotonically increasing in the set inclusion sense, i.e., if $T' \subseteq T''$ then $\Phi(T') \subseteq \Phi(T'')$. Indeed, the iteration procedure in Section 2.2 that defines $\Phi(\cdot)$ satisfies, inductively, $\Phi^m(T') \subseteq \Phi^m(T'')$ for each $m = 1, 2, \dots$, which results in $\Phi(T') \subseteq \Phi(T'')$.

Now suppose that Algorithm I outputs *NCP holds*. We then have a sequence of subsets of T, $\{T^0, T^1, \dots, T^{k^*}\}$, where $T^k = \Phi(T^{k-1}) \subsetneq T^{k-1}$ for all $k = 1, 2, \dots, k^*$ and $T^{k^*} = \emptyset$. For any nonempty $T' \subseteq T$, there exists some k such that $T' \subseteq T^k$ and $T' \not\subseteq T^{k+1}$. Since $\Phi(\cdot)$ is monotonically increasing, $\Phi(T') \subseteq \Phi(T^k) = T^{k+1}$. Since $T' \not\subseteq T^{k+1}$, we have $\Phi(T') \neq T'$. Thus, the dataset satisfies NCP. $\square$

Proof of Theorem 3. It suffices to show that if $\mathcal{O}^{**}$ is coarsely rationalizable (by a locally nonsatiated preference), then $\mathcal{O}^*$ is dual-coarsely rationalizable (by a locally nonsatiated preference). To see this, note that if $\mathcal{O}^{**}$ is coarsely rationalizable, then by Theorem 1, there exists a selected dataset

$$\mathcal{D}^{**} = \left\{(p^{t,s_t}, \{p \in \mathbb{R}^n_+ : x^t \cdot p \leq 1\})\right\}_{t \in T},$$

with $s_t \in G_t$ for every $t \in T$, such that it is rationalizable. Define the revealed preference relations (R^{**}, P^{**}) among $\{p^{t,s_t}\}_{t \in T}$ such that for all $t, \hat{t} \in T$,

(1) $p^{t,s_t} R^{**} p^{\hat{t},s_{\hat{t}}}$ if and only if $x^t \cdot p^{t,s_t} \geq x^t \cdot p^{\hat{t},s_{\hat{t}}}$, and

(2) $p^{t,s_t} P^{**} p^{\hat{t},s_{\hat{t}}}$ if and only if $x^t \cdot p^{t,s_t} > x^t \cdot p^{\hat{t},s_{\hat{t}}}$.[18]

By Afriat's Theorem, the rationalizability of $\mathcal{D}^{**}$ is equivalent to the acyclicity of (R^{**}, P^{**}). However, note that $p^{t,s_t} R^{**} p^{\hat{t},s_{\hat{t}}}$ is equivalent to $1 \geq x^t \cdot p^{\hat{t},s_{\hat{t}}}$ which is further equivalent to $x^{\hat{t}} \cdot p^{\hat{t},s_{\hat{t}}} \geq x^t \cdot p^{\hat{t},s_{\hat{t}}}$, and similarly, $p^{t,s_t} P^{**} p^{\hat{t},s_{\hat{t}}}$ is equivalent to $x^{\hat{t}} \cdot p^{\hat{t},s_{\hat{t}}} > x^t \cdot p^{\hat{t},s_{\hat{t}}}$. It follows from the acyclicity of (R^{**}, P^{**}) that (R, P) is also

[18] See Deb et al. (2023) for a similar revealed preference relation among prices.

acyclic, where R and P are revealed preference relations defined over $\{x^t\}_{t\in T}$ such that for all $t, \hat{t} \in T$, $x^t R x^{\hat{t}}$ if and only if $x^t \cdot p^{t,s_t} \geq x^{\hat{t}} \cdot p^{t,s_t}$, and $x^t P x^{\hat{t}}$ if and only if $x^t \cdot p^{t,s_t} > x^{\hat{t}} \cdot p^{t,s_t}$. By Afriat's Theorem, the dataset $\mathcal{D} = \{x^t, L(p^{t,s_t})\}_{t\in T}$ is rationalizable. Therefore, $\mathcal{O}^*$ is dual-coarsely rationalizable. □

Perturbation Index and Varian (1985). The idea of relaxing a revealed preference test to allow for measurement error was also explored by Varian (1985). That paper uses this idea to test the hypothesis that a firm is cost minimizing. Using our language, the introduction of measurement error leads to a coarsening of the dataset, where the true choice made by the firm is allowed to be in a ball around the observed choice x^t. Because the hypothesis being tested is different (cost-minimization rather than utility-maximization), the test used in that paper is *not* related to the never-covered property. To be precise, Varian (1985) assumes that the observer has information on factor prices, factor demand (imperfectly observed) and the output level. In the context of consumption data, the analog to the output level would be the utility level, but notice that we do not require this information (even in an ordinal form) to be part of our observations. It is the absence of this information that makes testing the utility-maximization hypothesis (with or without coarsening) a different exercise from testing cost-minimization.

Alternative Methods for Calculating Rationality Indices. We know from Afriat's Theorem that a dataset $\mathcal{D} = \{(x^t, L(p^t))\}_{t\in T}$ obeys GARP (and is thus rationalizable) if and only if there are u^t and λ^t (for $t \in T$) that solve the *Afriat inequalities*

$$\begin{aligned} \lambda^t &> 0 \\ u^s &\leq u^t + \lambda^t p^t \cdot (x^s - x^t) \quad \text{for } s \neq t. \end{aligned}$$

To calculate the perturbation index, we need to determine if $\mathcal{O}^\kappa$ is rationalizable. Instead of checking this with NCP, we can appeal to the Afriat inequalities and check whether there are u^t, λ^t, and bundles z^t (for all $t \in T$) that solve

$$z^t \geq 0$$

$$\begin{aligned} |p_i^s z_i^s - p_i^s x_i^s| &\leq \kappa \quad \text{for each good } i \\ \lambda^t &> 0 \\ u^s &\leq u^t + \lambda^t p^t \cdot (z^s - z^t) \quad \text{for } s \neq t. \end{aligned}$$

This approach is not satisfactory because the system of inequalities is bilinear in the unknowns and solving a bilinear system of inequalities is, in general, an NP-hard problem (see Toker and Ozbay (1995)). On the other hand, we have explained in the paper that checking whether $\mathcal{O}^\kappa$ is rationalizable can in fact be accomplished in polynomial time.

Similarly, a crucial step in calculating the price misperception index is to determine the dual coarse rationalizability of $\mathcal{O}^*_\epsilon$ (see (9)) for a given $\epsilon > 0$. As observed by de Clippel and Rozen (2023), this can be obtained by solving the Afriat inequalities in a procedure similar to the one described above, except that one will have to solve for a set of unknown prices (rather than a set of unknown bundles). Once again, such an approach is not satisfactory because the system of inequalities will be bilinear in the unknowns when, in fact, one could ascertain the dual coarse rationalizability of $\mathcal{O}^*_\epsilon$ in polynomial time.

In the case where there are just two goods, de Clippel and Rozen (2023) propose a different and simple method for calculating the price misperception index. That method relies crucially on the well-known result (see Banerjee and Murphy (2006)) that when there are just two goods, every revealed preference cycle with a strict relation (see (2)) must contain a shorter cycle of length two. That approach will not work in settings where there are three or more goods (see, for example, Ahn et al. (2014)).

Variations on the Perturbation Index. The perturbation and price misperception indices share a characteristic with Afriat's CCEI: in all three cases the permissible 'deviation' to restore rationality is *uniform* across all observations. In the case of the perturbation index, the bound of κ is uniformly applied to all observed bundles (see (4)); in the price misperception index, the price misperception is measured by δ and is uniformly bounded between $1/(1-\epsilon)$ and $1-\epsilon$ at each observation (see (8)); lastly, in the CCEI, a common cost efficiency e is imposed on all observations. Varian (1990) proposes a modification of the CCEI where the cost efficiency level at different observations are allowed to vary and a dataset's departure from rationality is measured by some average of the cost efficiency across all observations. A similar idea can be applied to both the perturbation and price misperception indices.

We explain this more carefully in the context of the perturbation index. Let

$$A^{t,\kappa^t} := \left\{x \in L(p^t) : \; p^t \cdot x = 1 \text{ and } |p_i^t x_i - p_i^t x_i^t| \leq \kappa^t \text{ for all } i\right\},$$

where $\kappa^t \in [0, 1]$. (This definition coincides with that given in (4)) except that the bound κ^t is allowed to vary with the observation t.) For any $\boldsymbol{\kappa} = (\kappa^1, \kappa^2, \ldots, \kappa^T)$, we can test, via NCP, whether there is a regular utility function $U : \mathbb{R}_+^n \to \mathbb{R}$ that coarsely rationalizes

$$\mathcal{O}^{\boldsymbol{\kappa}} = \left\{\left(A^{t,\kappa^t}, L(p^t)\right)\right\}_{t \in T}.$$

Let $S : [0, 1]^T \to \mathbb{R}$ be an aggregation rule defined on different values of $\boldsymbol{\kappa}$. For example, if we set $S(\boldsymbol{\kappa}) = \sum_{t=1}^T \kappa^t / T$, then we are simply taking the arithmetic average over κ^t. Based on S, we have a *generalized perturbation index*

$$S^* := \inf\{S(\boldsymbol{\kappa}) : \mathcal{O}^{\boldsymbol{\kappa}} \text{ is coarsely rationalizable by a regular utility function}\}$$

This index corresponds to the (standard) perturbation index if $S(\boldsymbol{\kappa}) = \max_{t \leq T} \kappa^t$.

Calculating the index S^* is, in general considerably more difficult than calculating the (standard) perturbation index κ^*. Checking if $\mathcal{O}^{\boldsymbol{\kappa}} = \{A^{t,\kappa^t}, L(p^t)\}_{t \in T}$ satisfies NCP is straightforward using our algorithm; the fact that the bounds κ^t are allowed to vary with t does not create any difficulty. However, finding S^* requires searching over $\{\kappa^t\}_{t \in T}$, which lives in a T-dimensional space.[19] This computational difficulty is exactly analogous to the difficulty involved in calculating Varian's index. Different approaches have been developed that are effective in calculating Varian's index (see Halevy, Persitz and Zrill (2018)) (by searching through the T-dimensional space more efficiently) and those approaches could also be used to calculate S^*.

[19] Note that this is unlike the case of calculating κ^* where we can require κ^t to be equal across observations with no loss of generality, effectively reducing the search to a one-dimensional problem.

References

Afriat, Sydney N. 1967. "The construction of utility functions from expenditure data." *International Economic Review*, 8(1): 67–77.

Afriat, Sydney N. 1973. "On a System of Inequalities in Demand Analysis: An Extension of the Classical Method." *International Economic Review*, 460–472.

Ahn, David, Syngjoo Choi, Douglas Gale, and Shachar Kariv. 2014. "Estimating ambiguity aversion in a portfolio choice experiment." *Quantitative Economics*, 5(2): 195–223.

Andreoni, James, and John Miller. 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica*, 70(2): 737–753.

Banerjee, Samiran, and James H Murphy. 2006. "A simplified test for preference rationality of two-commodity choice." *Experimental Economics*, 9: 67–75.

Carvalho, Leandro S, Stephan Meier, and Stephanie W Wang. 2016. "Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday." *American Economic Review*, 106(2): 260–284.

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv. 2007. "Consistency and Heterogeneity of Individual Behavior under Uncertainty." *American Economic Review*, 97(5): 1921–1938.

Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman. 2014. "Who is (More) Rational?" *American Economic Review*, 104(6): 1518–1550.

Deb, Rahul, Yuichi Kitamura, John K.-H. Quah, and Jörg Stoye. 2023. "Revealed Price Preference: Theory and Empirical Analysis." *The Review of Economic Studies*, 90(2): 707–743.

Debreu, Gerard. 1954. "Representation of a preference ordering by a numerical function." *Decision Processes*, 3: 159–165.

de Clippel, Geoffroy, and Kareen Rozen. 2023. "Relaxed Optimization: How Close Is a Consumer to Satisfying First-Order Conditions?" *Review of Economics and Statistics*, 105(4): 883–898.

Fishburn, Peter. 1976. "Representable Choice Functions." *Econometrica*, 44(5): 1033–1043.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual Preferences for Giving." *American Economic Review*, 97(5): 1858–1876.

Forges, Francoise, and Enrico Minelli. 2009. "Afriat's Theorem for General Budget Sets." *Journal of Economic Theory*, 144(1): 135–145.

Halevy, Yoram, Dotan Persitz, and Lanny Zrill. 2018. "Parametric Recoverability of Preferences." *Journal of Political Economy*, 126(4): 1558–1593.

Harbaugh, William T, Kate Krause, and Timothy R Berry. 2001. "GARP for Kids: On the Development of Rational Choice Behavior." *American Economic Review*, 91(5): 1539–1545.

Hu, Gaoji, Jiangtao Li, John K.-H. Quah, and Rui Tang. 2022. "A theory of revealed indirect preference." *Available at SSRN 3776049.*

Karmarkar, Narendra. 1984. "A new polynomial-time algorithm for linear programming." In *Proceedings of the sixteenth annual ACM symposium on Theory of computing.* 302–311.

Nishimura, Hiroki, Efe A. Ok, and John K.-H. Quah. 2017. "A Comprehensive Approach to Revealed Preference Theory." *American Economic Review*, 107(4): 1239–1263.

Polisson, Matthew, John K.-H. Quah, and Ludovic Renou. 2020. "Revealed Preferences Over Risk and Uncertainty." *American Economic Review*, 110(6): 1782–

1820.

Richter, Marcel K. 1966. “Revealed Preference Theory.” *Econometrica*, 34(3): 635–645.

Toker, Onur, and Hitay Ozbay. 1995. “On the NP-Hardness of Solving Bilinear Matrix Inequalities and Simultaneous Stabilization with Static Output Feedback.” Vol. 4, 2525–2526, IEEE.

Varian, Hal R. 1982. “The Nonparametric Approach to Demand Analysis.” *Econometrica: Journal of the Econometric Society*, 945–973.

Varian, Hal R. 1985. “Non-parametric Analysis of Optimizing Behavior with Measurement Error.” *Journal of Econometrics*, 30(1-2): 445–458.

Varian, Hal R. 1990. “Goodness-of-fit in optimizing models.” *Journal of Econometrics*, 46(1-2): 125–140.