Smooth rationalization

Cristián Ugarte*

November 15, 2022

Most Recent Version here

Abstract

Economic models usually endow agents with (well-behaved) differentiable utilities. However, the behavioral implications of such an assumption are unclear. We study conditions under which consumer choices can be rationalized by a differentiable utility, which we denote as smooth rationalization. The main consequence of assuming a differentiable utility is that choices that are revealed indifferent to each other must have the same marginal rate of substitution. To include this requirement, we propose a modification of the original data set using the revealed indifference relation and show that rationalizing the modified data set is equivalent to smooth rationalization. The same reasoning applies when differentiability into several more restricted families of utility functions is required. We develop tests for smooth rationalization in the cases of strictly concave, homothetic, and quasilinear utilities. We also show that the existence of second- and higher-order derivatives of the utility function cannot be tested. We implement our test of a differentiable utility into several experimental data sets. Our results suggest that, in most cases, assuming differentiability is not costly from an empirical standpoint. However, our data set also presents subjects whose behavior closely follows common utility functions that are not differentiable.

*Department of Economics, University of California Berkeley. cugarte@berkeley.edu. I thank Shachar Kariv, Chris Shannon, Federico Echeñique, and participants of the UC Berkeley Theory lunch and the 2022 SAET Conference (Revealed Preferences session) for many thoughtful comments and suggestions. All mistakes are my own.

The axioms of the [consumer] theory must be formulated in terms of observable choices made by a consumer among commodity vectors

Gérard Debreu (1972, p.605)

1 Introduction

Differentiability of the utility function is one of the more elusive assumptions in economics. It is a widespread and helpful assumption in that it allows us to characterize optimal choices through first-order conditions, which significantly simplifies the mathematical analysis and adds tractability to economic models, comparative statics, and their applications. However, there is no clear interpretation of the behavioral implications of assuming a differentiable utility function. Unlike other assumptions, such as monotonicity, continuity, or concavity, the axioms for a differentiable utility function are motivated by geometry, not behavior. In this paper, we take a different approach and study the assumption of a differentiable utility function from an empirical (rather than an axiomatic) perspective. We provide necessary and sufficient conditions on demand data under which the observed choices can be rationalized by a well-behaved (i.e., strictly increasing, continuous, and concave) differentiable utility function.

There are two different but related benefits of providing a characterization of smooth rationalization. First, it provides a theoretical understanding of the restrictions imposed by a differentiable utility on the behavior of a consumer who makes choices in a competitive market. Given that the impositions of these assumptions over preference relations are not well understood, this exercise is particularly relevant. Second, the conditions that we provide can be used to test for a differentiable utility in actual data. Because there is a vast applied economics literature that relies on first-order conditions to characterize results, testing for differentiability of the utility function is a critical exercise in understanding the validity of the conclusions of this literature.

Our characterization of smooth rationalization is framed in the revealed preference literature that starts with Samuelson (1938). Afriat's theorem, the seminal result in the revealed preference literature (Afriat, 1967; Varian, 1982), states that choices can be rationalized by a well-behaved utility if and only if they satisfy the Generalized Axiom of Revealed Preferences (GARP). Focusing on differentiability and strict convexity, Chiappori and Rochet (1987) show that an invertible demand (i.e., two different choices need to arise from different prices) is sufficient to achieve a differentiable utility function.[1] However, we show that an invertible demand is too strong a requirement in the case of a strictly concave utility, and is not sufficient if the utility is concave instead of strictly concave. Instead, we show that *revealed indifferences* are the main components that are able to add differentiability to such a utility.

In the classical consumer setting, revealed indifferences arise if two different choices are revealed preferred to each other. The main observation for our analysis is that, whenever two choices are revealed indifferent to each other, concavity of the utility function implies that the indifference set between such choices must be flat. Hence, both choices have the same marginal rate of substitution (MRS). Furthermore, we show that when several observed choices are revealed indifferent to one another, each choice is optimal from not only its own price, but also the meet of all the prices at which the indifferent choices were made.[2] This observation allows us to characterize necessary and sufficient conditions to rationalize the data with a differentiable utility by a simple modification of the data set: each price is replaced with the meet of the prices among choices that are revealed indifferent. Furthermore, if the new prices imply further indifferences, the modification has to be done again until a fixed point is reached. Rationalization by a well-behaved and differentiable utility, which we call *smooth rationalization*, can be tested by checking GARP in this modified data set. Furthermore, we show that whenever we can smoothly rationalize the choices, we can do so with an infinitely differentiable utility. Hence, to assume the existence of second and further derivatives of the utility function comes at no cost from an empirical perspective.

From an axiomatic standpoint, Debreu (1972) presents a characterization of differentiability of the utility function. In the words of Neilson (1991), differentiability "requires that indifference sets be smooth and 'vary smoothly' "; that is, indifference sets do not have kinks, and the change in utility level must also be smooth. Our characterization of a smooth utility function rationalizing demand data implies that only the first of Debreu's conditions (smooth indifference sets) is relevant from an empirical perspective, and that smooth variation of the utility level cannot be tested.

As in Afriat's theorem, our result provides a characterization of smooth rationalization through a modified version of the Afriat inequalities, i.e., through the existence of numbers satisfying a given set of inequalities. Usually, the Afriat inequalities can be explained using the concavity of

[1] Surprisingly, the conditions proposed by Chiappori and Rochet (1987) are usually referred to as necessary and sufficient to obtain a differentiable utility in the case of the Strong Axiom of Revealed Preferences (SARP). However, Chiappori and Rochet (1987) do not claim the necessity of their conditions.

[2] For a given set of K-dimension vectors, the meet of such vectors is the vector resulting from taking the coordinate-wise minimum among the set.

the utility function. However, they can also be motivated through the Lagrangian of the utility maximization problem (assuming differentiability) if the nonnegativity constraints are omitted. The inequalities in our characterization start from the complete Lagrangian of the utility maximization problem, i.e., they include the nonnegativity constraints for every good. Furthermore, we need to include inequalities characterizing the vector of marginal utilities (which may not exist in the nondifferentiable case) and the complementary slackness condition for the nonnegativity restrictions (as they might not be binding, unlike the budget constraint).

We use our characterization of rationalization to test smooth rationalization on experimental data sets from several sources. In total, we test the differentiability assumption on 4,958 subjects who make choices either between risky assets or in a social environment where they choose between own consumption and consumption by an anonymous third party. We find that more than 92% of subjects who satisfy GARP also satisfy smooth rationalization. Moreover, we use the Houtman and Maks (1985) Index (HM Index) to recover preferences for subjects who fail GARP and find that adding differentiability to the rationalization requirements, in general, does not increase the value of the index. On the other hand, we find subjects whose choice patterns (approximately) follow well-known families of utility functions that are not differentiable and accumulate most (or all) of their choices in the nondifferentiable portion of the utility function.[3] An additional finding of our analysis is that corner solutions are a common feature of both choices under risk and social choices.

We show that our main idea, that revealed indifferences are the key component in testing for differentiability of the utility function, applies to several types of rationalization. We first focus on a utility that is strictly convex. From Matzkin and Richter (1991), we know that the necessary and sufficient condition for the existence of a well-behaved and strictly concave utility is for choices to satisfy Houthakker's (1950) Strong Axiom of Revealed Preferences (SARP). In this case, we use revealed indifferences to modify our data set in the same form as in the case of GARP (i.e., by taking the meet of prices among choices that are revealed indifferent to each other), and show that a necessary and sufficient condition for smooth rationalization by a strictly concave data set is SARP of the modified data set. We link our result to the invertible demand requirement proposed by Chiappori and Rochet (1987), which they name Strong SARP (SSARP). The motivation for

[3]Because we assume that the utility is concave, we know it is differentiable almost everywhere (Theorem 25.5 in Rockafellar, 2015). However, it is still possible for the nondifferentiable part of the utility to accumulate most of the choices. An extreme example is the Leontief utility, which accumulates all the choices on its nondifferentiable part.

SSARP is that a characterization of the optimal consumption is the equality between the MRS and the price ratio. Hence, a noninvertible data set implies that the same choice arises from different price vectors, and this equality cannot hold. The difference between SSARP and our characterization of smooth rationalization in the case of SARP is that equality between the MRS and the price ratio does not need to hold if the optimal bundle is a corner solution, i.e., it does not have to hold for goods with zero demand. Empirically, we find that corner solutions are common, and hence most (83%) of the subjects who satisfy SARP and fail SSARP present choices that can be rationalized by a differentiable utility. This difference is also relevant when recovering preferences using the HM Index.

We finish the paper by analyzing two more common restrictions to the utility function. We first analyze rationalization by a homothetic utility and then by a quasilinear utility. The rationalization test for a homothetic utility is the Homothetic Axiom of Revealed Preferences (HARP Varian, 1983a), and that for a quasilinear utility is cyclical monotonicity (Brown & Calsamiglia, 2007). In both cases, the specific structure of the utility function allows us to infer further indifferences than the ones obtained using the classical revealed preferences approach. We show that smooth rationalization can be tested in these settings by including the new indifferences in the modification of the data set, and then by using the specific test for each type of utility in the modified data set. In this sense, the intuition of the test in the GARP case still applies: revealed indifference, along with concavity, implies that the indifference curve must be flat, and hence all the choices that are revealed indifferent to each other must have the same MRS and be optimal from the price resulting from the meet of their respective prices.

Related Literature

Our work contributes to the revealed preference literature. The problem of rationalizing a data set of choices starts with Samuelson (1938). The seminal result of this literature is Afriat's theorem (Afriat, 1967; Diewert, 1973; Varian, 1984), which has led to a vast literature on rationalizing observed choices. Fostel et al. (2004), Polisson and Renou (2016), and Beggs (2021) study new proofs of Afriat's theorem that highlight different aspects and interpretations of the result.

Afriat's theorem has been extended in several directions. A seminal example is Matzkin and Richter (1991), who show that SARP is necessary and sufficient for rationalization by a strictly concave utility. The main consequence of a strictly concave utility is that the demand it generates is a single-valued function instead of a correspondence. Taking a further step, Lee and Wong

(2005) show that under SARP, the rationalizing utility can be chosen to generate a demand that is not only single-valued but also infinitely differentiable. Chiappori and Rochet (1987) study the problem of smooth rationalization under SARP and provide SSARP as a sufficient condition. The paper by Chiappori and Rochet (1987) is closest to ours regarding the research question and the methodology. There are two main differences between their work and ours. First, we develop sufficient and necessary conditions, allowing us to characterize and test smooth rationalization and obtain a behavioral interpretation of the differentiability assumption. Second, our analysis is more general as it is not restricted to a strictly concave utility, although it can still be applied to this case.

Afriat's theorem has been extended in several directions apart from the basic assumptions of strict concavity and differentiability. One notable direction is to make it more general: Forges and Minelli (2009) characterize the rationalization of choices under nonlinear budget sets, and Nishimura et al. (2017) provide a general characterization for a more abstract consumption space and notion of desirability. In all cases, the characterization of a rationalizable data set is remarkably similar to GARP. Another extension of Afriat's theorem is the analysis of rationalization by commonly used functional forms of the utility function. For example, Varian (1983a) studies rationalization by a homothetic utility, Brown and Calsamiglia (2007) by a quasilinear utility, and Varian (1983a), Diewert and Parkan (1985), and Quah (2014) by a weakly separable utility. In the specific domain of choices under risk, Green and Srivastava (1986), Varian (1983b), and Kubler et al. (2014) provide a characterization of choices rationalizable by an expected utility function when beliefs are objective, and Echenique and Saito (2015) extend this analysis to the case of subjective expected utility, i.e., to the case when probabilities are unknown to the econometrician.

From an axiomatic perspective, the seminal characterization of a differentiable utility function is provided by Debreu (1972). He shows that continuous and increasing preferences can be represented by a differentiable utility if, and only if, their indifference sets are smooth manifolds. Alternatively, Rubinstein (2012) provides a characterization of differentiable utilities as preferences having a property that can be interpreted as the existence of a marginal utility vector.[4] Although such axiomatizations lack an appealing economic interpretation, the convenience of a differentiable utility

[4] In addition, in the specific case of expected utility, Nielsen (1999) and Nakamura (2015) find conditions for obtaining a differentiable Bernoulli function. Intuitively, they exploit the fact that a linear function can locally approximate a differentiable function. Hence, we can think of the agent as being approximately risk-neutral for small lotteries.

means the assumption is widely used in applied research.

Finally, our empirical analysis relies on the experimental design pioneered by Choi et al. (2007b). Several tools are proposed to recover preferences in the case that a subject fails GARP, among them Afriat (1973), Houtman and Maks (1985), Varian (1990), de Clippel and Rozen (2021), and Ugarte (2022a). We choose the methodology proposed by Houtman and Maks (1985) because it is the only one among the most popular ones (the others two being Afriat, 1973; Varian, 1990) that allows us to differentiate between differentiable and nondifferentiable utilities (see Ugarte, 2022b, for a detailed discussion).

The remainder of the paper proceeds as follows: Section 2 introduces the problem and presents the main definitions for our analysis. Section 3 studies the problem of rationalization by smooth rationalization under SARP, i.e., rationalization by a strictly concave and differentiable utility. We start by analyzing the SARP case because it is analytically simple, but still includes most of the conceptual components of our main result, and because it provides a direct comparison with the seminal paper by Chiappori and Rochet (1987). Section 4 presents the general characterization of smooth rationalization. Section 5 implements our smooth rationalization test into experimental data. Section 6 analyzes smooth rationalization for both homothetic and quasilinear utilities. Finally, Section 7 concludes. All the proofs are provided in the appendices.

2 Preliminaries

Notation We work with the following notation. $\mathbb{N}$ is the set of natural numbers (excluding zero), $\mathbb{R}$ is the set of real numbers, $\mathbb{R}_+$ is the set of positive numbers including zero, and $\mathbb{R}_{++}$ excludes it. For any $M \in \mathbb{N}$, $[M] = \{1, 2, \ldots, M\}$ is the set of the first M naturals. A vector $x \in \mathbb{R}^M$ is $x = (x_1, x_2, \ldots, x_M)$. The K-dimensional zero vector is $\mathbf{0}$. For any two vectors $x, y \in \mathbb{R}^M$ we write $x \geq y$ if $x_i \geq y_i$ for all $i \in [M]$, $x > y$ if $x \geq y$ and $x \neq y$, and $x \gg y$ if $x_i > y_i$ for all $i \in [M]$ ($<$, $\leq$, and $\ll$ are defined similarly). A function $f : \mathbb{R}^M \to \mathbb{R}$ is strictly increasing if $x > y$ implies $f(x) > f(y)$. For $x, y \in \mathbb{R}^M$, $||x - y||$ denotes the Euclidean distance, and $x \cdot y$ the dot product.

Our analysis starts from an agent's demand data set, comprised of N different observations. The agent consumes bundles of K nonnegative goods, i.e., the consumption space is $\mathbb{R}^K_+$. In each observation $i \in [N]$, the agent faces a strictly positive price vector $p^i \in \mathbb{R}^K_{++}$ and spends an amount of money normalized to one. The expenditure normalization is without loss of generality and greatly simplifies the notation for our exposition. Given p^i, the agent chooses a bundle from the budget set $\{x \in \mathbb{R}^K_+ : p^i \cdot x \leq 1\}$; we denote such a choice as x^i. Together, prices and choices form the data set $\mathcal{D} = (x^i, p^i)_{i \in [N]}$, which is the primitive of our problem. We refer to both rounds $i \in [N]$ and pairs (p^i, x^i) as observations and to bundles x^i as chosen bundles or choices. As standard in the revealed preference literature (and unavoidable for rationalization by any meaningful utility), we assume that the agent spends all her income on each choice, i.e., $p^i \cdot x^i = 1$.

Revealed preferences are the main tool for learning about the agent's preferences. For two bundles x, y we can infer that x is preferred to y if when x was chosen y was also available; this is, if there is $i \in [N]$ such that $x = x^i$ and $p^i \cdot y \leq 1$. Moreover, the idea of monotonicity, that "more is better", leads us to infer that if x was chosen and y was not only available but also strictly cheaper, then x has to be strictly preferred to y. Along with the idea of transitivity, the previous notion leads to the definition of revealed preferences.

Definition 1. For a data set $\mathcal{D}$ and two choices x^i and x^j, x^i is

- directly revealed preferred to x^j, denoted $x^i \succsim^\star x^j$, if $p^i \cdot x^j \leq 1$;
- directly revealed strictly preferred to x^j, denoted $x^i \succ^\star x$, if $p^i \cdot x^j < 1$;
- revealed preferred to x^j, denoted $x^i \succsim x^j$, if there is a sequence of observations $(m_\ell)_{\ell \in [L]}$ such that

$$x^i \succsim^\star x^{m_1} \succsim^\star x^{m_2} \succsim^\star \ldots \succsim^\star x^{m_L} \succsim^\star x^j; \text{ and}$$

- revealed indifferent to each x^j, denoted $x^i \sim x^j$, if $x^i \succsim x^j$ and $x^j \succsim x^i$.

The definition of revealed preferences above presents some slight differences from the standard definition. First, and without loss of generality, it only focuses on comparisons between choices. It also does not use transitivity to infer strict preferences but uses it to infer indifference. The reason for this is that we only need revealed strict preferences that are direct for our exposition.[5] Since the relation of revealed indifference is an equivalence relation (when restricted to choices in the data) whenever $x^i \sim x^j$, we also say that x^i and x^j are revealed indifferent to each other.

Afriat (1967) uses revealed preferences to show that we can interpret the choices in $\mathcal{D}$ as being driven by a locally nonsatiated utility function (this is, there is a locally nonsatiated U satisfying $U(x^i) \geq U(x)$ whenever $p^i \cdot x \leq 1$) if, and only if, it satisfies a condition he called cyclical consistency. Furthermore, the rationalizing utility can always be chosen to be strictly increasing, continuous, and concave; hence with linear prices, these properties cannot be falsified.[6] The most popular

[5] Typically, the definition of revealed preferences also includes the revealed strict preference relation, where x^i is revealed strictly preferred to x^j if $x^i \succsim x^j$ and one of the relations in the sequence linking x^i and x^j is strict.

[6] Forges and Minelli (2009) discuss how to test strict concavity when budget sets are not linear.

version of cyclical monotonicity is the Generalized Axiom of Revealed Preferences, proposed by Varian (1982).

Definition 2. $\mathcal{D}$ satisfies the *Generalized Axiom of Revealed Preferences*, *GARP*, if for any pair of choices x^i and x^j,

$$x^i \succsim x^j \implies x^j \not\succ^\star x^i \,.$$

The intuition of GARP as a necessary condition to rationalize the data is clear. If a utility function drives the choices, the revealed preferences will agree with the ordering established by such a utility. Hence we cannot infer that x^i is (weakly) preferred to x^j and that x^j is strictly preferred to x^i. The proof that GARP is also sufficient is constructive. It relies on showing that GARP implies the existence of a set of numbers that satisfy relations known as the Afriat inequalities, which are used to construct a utility function that is strictly increasing, continuous, and concave.[7]

Usually, applied economic research imposes more structure on the utility function than the one obtained in Afriat's Theorem. The two more common additional assumptions are strict concavity and differentiability. A strictly concave utility has a unique optimal choice from each price vector, which implies that the induced demand is a single-valued function.[8] Differentiability implies that optimal choices can be characterized by first-order conditions, which make optimal choices tractable to perform comparative statics. Matzkin and Richter (1991) study the empirical content of assuming a strictly concave utility; they show that Houthakker's (1950) Strong Axiom of Revealed Preferences is necessary and sufficient to rationalize a data set by a strictly increasing, continuous, and strictly concave utility.

Definition 3. $\mathcal{D}$ satisfies the *Strong Axiom of Revealed Preferences*, *SARP*, if for any two different choices $x^i \neq x^j$,

$$x^i \succsim x^j \implies x^j \not\succsim^\star x^i \,.$$

The relation between SARP and strict concavity of the utility function can be understood using the following characterization.

[7]Specifically, GARP implies the existence of numbers $u^i \in \mathbb{R}$ and $\lambda^i > 0$ such that $u^i \geq u^j + \lambda^i(1 - p^i \cdot x^j)$ for every i, j. The classical utility rationalizing the data in the proof of Afriat's Theorem is $U(x) = \min_{i \in [N]} u^i - \lambda^i(1 - p^i \cdot x)$. Nishimura et al. (2017) present a nonconstructive proof of Afriat's Theorem for a general choice environment and criterion of desirability; their proof does not include concavity as a property of the utility function.

[8]Suppose x, y are two optimal choices from price p and $x \neq y$, and the utility function U is strictly concave. Then the bundle $z = x/2 + y/2$ satisfies $p \cdot z = 1$ and, by strict concavity $U(z) > U(x)/2 + U(y)/2 = U(x) = U(y)$. Then z yields a higher utility than x and y and is affordable at price p, contradicting the optimality of x and y.

Remark 1. $\mathcal{D}$ satisfies SARP if, and only if, it satisfies GARP and $x^i \not\sim x^j$ whenever $x^i \neq x^j$.

All the remarks' proofs are in Appendix A. Rationalization with a strictly concave utility U implies $U(x^i) > U(x)$ whenever $x \neq x^i$ and $p^i \cdot x \leq 1$, which is known as strong rationalization. Hence, it does not allow for two different chosen bundles to be indifferent to each other. If such a relation occurs, then there would be two choices $x^i \neq x^j$ that are indifferent to each other, and such that $p^i \cdot x^j = 1$, implying that both x^i and x^j are optimal choices when the price vector is p^i. This relation contradicts a single-valued demand, i.e., contradicts strong rationalization.

This paper studies the empirical content of the utility function's differentiability assumption. We know that any concave utility function is generically differentiable, i.e., differentiable in a dense subset of its domain (Rockafellar, 2015, Theorem 25.5). However, that does not translate into a zero probability of observing choices in the nondifferentiable portion of such utility. An extreme example of such case is the Leontief utility $U(x) = \min_{k\in[K]} x_k$. For any price p, the optimal choice of the Leontief utility satisfies $x_1 = x_2 = \ldots = x_K$. Hence all choices lay in the nondifferentiable portion of the function. Although the Leontief utility is generically differentiable, the first-order conditions never characterize the utility-maximizing bundle.

Since basic notions of differentiability are defined in open sets, but the utility function is usually defined in $\mathbb{R}^K_+$, which is closed, we use a definition of differentiability that allows for a well-defined derivative at the boundary.

Definition 4. Take $X \subset \mathbb{R}^K$ and $x \in X$ a limit point of X. The function $f : X \to \mathbb{R}$ is *differentiable at* x if there is a vector $\nabla f(x) \in \mathbb{R}^K$ such that for every sequence of points $x^n \in X$ satisfying $x^n \to x$ and $x^n \neq x$ we have

$$\lim_{n\to\infty} \frac{f(x^n) - f(x) - \nabla f(x) \cdot (x^n - x)}{||x^n - x||} = 0$$

The function f is *differentiable* if it is differentiable at every limit point $x \in X$.

The only difference between Definition 4 and the standard definition of a (Fréchet) derivative is that the derivative is also defined at the boundary of the domain. This difference is relevant as it allows us to express corner solutions in terms of their first-order conditions. In our results, we obtain differentiability on the boundary by finding a differentiable function in a superset of $\mathbb{R}^K_+$, then defining the utility as the restriction of such function into the consumption set.

Following Afriat's Theorem, we focus on utilities that are also strictly increasing and concave (we omit continuity since it is implied by differentiability). We refer to this property as smooth rationalization.

Definition 5. $\mathcal{D}$ is *smoothly rationalizable* if there is a differentiable, strictly increasing, and concave utility function $U : \mathbb{R}^K_+ \to \mathbb{R}$ such that, for every $i \in [N]$, $U(x^i) \geq U(x)$ whenever $p^i \cdot x \leq 1$. Such function *smoothly rationalizes* $\mathcal{D}$.

3 SARP and Differentiable Utility

In this section, we analyze conditions to smoothly rationalize the observed choices by a strictly concave utility, i.e., for smooth rationalization restricted to SARP. We start focusing on SARP for two reasons. First, for historical reasons: Chiappori and Rochet (1987), the seminal work on conditions to rationalize choices by a differentiable utility, focuses on SARP. Hence the comparison of their results and ours is more explicit in this case. Second, strict concavity simplifies the exposition and examples without losing the most relevant features. After presenting the SARP analysis, the GARP extension should be more straightforward for the reader to follow.

3.1 Strong SARP: A sufficient condition for smooth rationalization

Chiappori and Rochet (1987) study conditions under which the data can be smoothly rationalized. Their motivation is the data set in panel (a) of Figure 1, in which two different budget sets have the same optimal choice. The reason such data is not smoothly rationalizable is explicit: the indifference curve going through the bundle x^t has to have a kink at x^t. If not, it would go to the interior of at least one of the budget sets, which is impossible if choices are optimal. Since the indifference curve has a kink, the utility function cannot be differentiable. In terms of the first-order conditions, a differentiable utility rationalizing the data implies that, in each choice, the ratio of prices has to be equal to the marginal rate of substitution (MRS). This condition fails if the same bundle is chosen from two different price vectors since the expenditure normalization requires that two prices have (some) different price ratios.[9]

To avoid situations like the one in panel (a) of Figure 1, Chiappori and Rochet (1987) require for the observed demand to be invertible; this is, for different prices to have different choices. They call this condition Strong SARP.

Definition 6 (Chiappori and Rochet, 1987)**.** $\mathcal{D}$ satisfies the *Strong version of SARP* (SSARP) if it satisfies SARP and for all observations i, j

[9]Without expenditure normalization, the requirement would be for prices and expenditures not to be proportional to each other.

(CR) $p^i \neq p^j$ implies $x^i \neq x^j$.

The main result from Chiappori and Rochet (1987) is that SSARP is sufficient for smooth rationalization by a strictly concave utility. Furthermore, they show that the rationalizing utility can be chosen to be infinitely differentiable.

Theorem (Chiappori and Rochet, 1987)**.** *If $\mathcal{D}$ satisfies SSARP, it is smoothly rationalizable by a strictly concave and infinitely differentiable utility function defined on a compact subset of $\mathbb{R}^K_+$.*

As noted by Matzkin and Richter (1991, Theorem 1^∞), the restriction of the function to a compact subset of $\mathbb{R}^K_+$ is unnecessary; this is, if $\mathcal{D}$ satisfies SSARP, it can be smoothly rationalized by a strictly concave and infinitely differentiable utility.

The following example illustrates why SSARP is not a necessary condition: it presents a simple data set that fails SSARP and a strictly concave and differentiable function that rationalizes it.

Example 1. Suppose there are two commodities, and $\mathcal{D}$ comprises two observations ($K = N = 2$). The observations are $(p^1, x^1) = ((2,1), (1,0))$, and $(p^2, x^2) = ((4,1), (0,1))$. As $p^1 \neq p^2$ and $x^1 = x^2$, $\mathcal{D}$ fails SSARP. However, it is smoothly rationalizable by $U(x) = \sqrt{(x_1+1)(x_2+1)}$, which is strictly increasing, strictly concave, and infinitely differentiable.

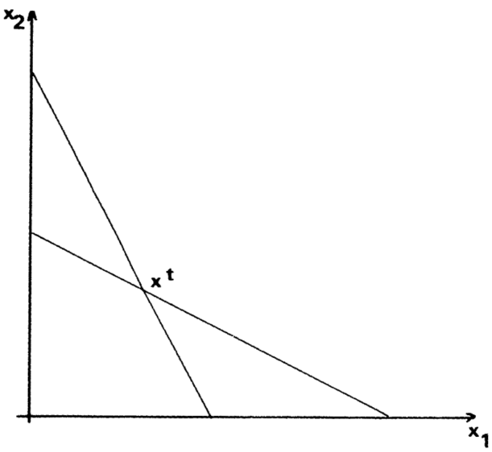
Example 1 is shown in panel (b) of Figure 1. The main reason why SSARP is not necessary is that the equality between MRS and price ratio does not have to hold for corner solutions, i.e., for goods whose consumption is equal to zero. Hence, an invertible demand is required whenever consumption is positive but not when it is not. Surprisingly, an invertible demand for interior solutions (i.e., strictly positive consumption) is insufficient to ensure a differentiable demand. The combination of strict concavity and differentiability also imposes testable requirements on the corner solutions.

3.2 SARP and Smooth Rationalization

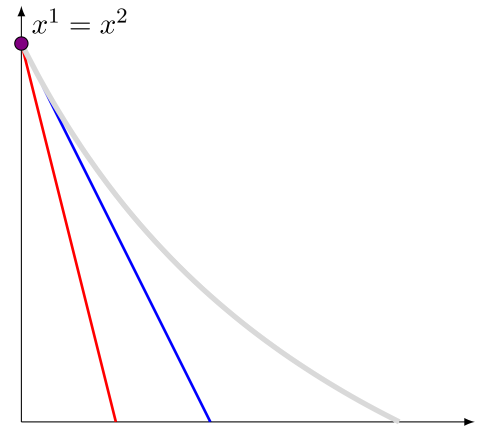
This section presents a test for smooth rationalization by a strictly concave utility. To simplify our exposition, we introduce the following definitions: for each $i \in [N]$:

$$q^i = \bigwedge_{\{j \in [N]: x^j = x^i\}} p^j = \left(\min_{\{j: x^j = x^i\}} p^j_1, \min_{\{j: x^j = x^i\}} p^j_2, \ldots, \min_{\{j: x^j = x^i\}} p^j_K \right); \text{ and}$$

$$\mathcal{D}^S_\wedge = \left(q^i, x^i\right)_{i \in [N]} . \tag{1}$$



(a) Figure 1 in Chiappori and Rochet (1987). Same choice is made from two different budget sets. The data set satisfies SARP but is not smoothly rationalizable.



(b) Representation of Example 1. $\mathcal{D}$ is not invertible, hence fails SSARP, but is smoothly rationalizable.

Figure 1: Two noninvertible data sets. The one in panel (a) cannot be rationalized by a differentiable utility; the one in panel (b) can be rationalized by a differentiable utility.

The vector q^i is given by the meet of price vectors over the set of all observations whose choice is equal to x^i (including observation i). $\mathcal{D}^S_\wedge$ is a data set in which each price p^i is replaced with q^i. In this section, we refer to $\mathcal{D}^S_\wedge$ as the *modified data set*; in contrast, $\mathcal{D}$ is the original data set. The modified data set plays a crucial role in our characterization of smooth rationalization.

The main result of this section presents necessary and sufficient conditions for smooth rationalization by a strictly concave utility. It also presents a modified version of the Afriat inequalities that introduce the possibility of corner solutions. Finally, it shows that, given a differentiable utility function, the existence of second and further derivatives has no empirical content.

Theorem 1. *The following are equivalent:*

Sa-1) $\mathcal{D}$ is smoothly rationalizable by a strictly concave utility.

Sa-2) $\mathcal{D}$ satisfies SARP, and the following two conditions hold

(D1) If $x^i = x^j$ and $x^i_k > 0$, then $p^i_k = p^j_k$.

(D2) $\mathcal{D}^S_\wedge$ satisfies SARP.

Sa-3) For all $i \in [N]$ there exists numbers $u^i \in \mathbb{R}$, $\lambda^i > 0$ and K-dimensional vectors $\mu^i \geq \mathbf{0}$ such that

$$u^i > u^j + \lambda^i(1 - p^i \cdot x^j) + \mu^i \cdot x^j \qquad \textit{whenever } x^i \neq x^j \tag{S1}$$

$$u^i = u^j \qquad \textit{whenever } x^i = x^j \tag{S2}$$

$$\lambda^i p^i - \mu^i = \lambda^j p^j - \mu^j \qquad \textit{whenever } x^i = x^j \tag{S3}$$

$$\lambda^i p^i - \mu^i \gg \mathbf{0} \qquad \textit{for all } i \in [N] \tag{S4}$$

$$\mu^i \cdot x^i = 0 \qquad \textit{for all } i \in [N] \tag{S5}$$

Sa-4) $\mathcal{D}$ is smoothly rationalizable by a strictly concave and infinitely differentiable utility.

The proof of this result is in Appendix C. An intuitive explanation of why Sa-1) implies Sa-2) and why Sa-2) implies Sa-3) is presented below. Showing that Sa-3) implies Sa-4) is constructive: we use the numbers in Sa-3), along with the techniques introduced by Chiappori and Rochet (1987) and by Matzkin and Richter (1991), to construct an strictly increasing, strictly concave, and differentiable utility function. As this function is also infinitely differentiable, we get the equivalence between Sa-1) and Sa-4).

Statements Sa-2) and Sa-3) provide two different characterizations of smooth rationalization by a strictly concave utility. Sa-2) provides conditions on the observed choices in terms similar to SARP and SSARP. Like many results in the revealed preference literature, Sa-3) characterizes rationalization as a modified version of Afriat inequalities. The main focus in our explanation of Theorem 1 is on Sa-2) since this characterization is the one that yields a behavioral interpretation.

First, we note that the conditions in Sa-2) are strictly weaker than SSARP. A data set that satisfies Sa-2) and fails SSARP is shown in Example 1. To see that SSARP implies Sa-2) note that SSARP implies that for every $i \neq j$ we have either $p^i \neq p^j$ and $x^i \neq x^j$, or $p^i = p^j$ and $x^i = x^j$. Hence (D1) is satisfied, and $q^i = p^i$ for all i, which implies that $\mathcal{D}^S_\wedge = \mathcal{D}$ and $\mathcal{D}^S_\wedge$ satisfies SARP. Also, it is easy to see that the requirement that $\mathcal{D}$ satisfies SARP is superfluous, as it is implied by $\mathcal{D}^S_\wedge$ satisfying the same property; we include it in the result only for expositional purposes.

Besides SARP, Sa-2) imposes two conditions on the data. Condition (D1) has a clear interpretation; it requires equality between MRS and price ratio, as in SSARP, but only for interior solutions. By contrapositive, condition 6 in the definition of SSARP is equivalent to the following condition: if $x^i = x^j$, then $p^i = p^j$. Condition (D1) relaxes this requirement by imposing it only for the goods consumed in a strictly positive amount.

From the difference between panel (a) and (b) in Figure 1, the reader could think that SARP and (D1) together are sufficient for smooth rationalization under SARP. However, this is not the case. Instead, strict concavity and differentiability impose restrictions on the corner solutions, i.e., on goods that are not consumed. Such restrictions translate into (D2). Specifically, these two properties of the utility function imply that indifference curves must be differentiable and strictly convex, which implies that (a scaled version of) q^i is a supergradient of the utility at x^i.[10] One reason why (D2) might be harder to interpret that (D2) is because, given SARP and (D1), it is vacuously satisfied when bundles are composed of two goods (see Remark 4 in Section 5). The following section presents a detailed explanation of the conditions in Sa-2) and an example to illustrate them.

3.3 Motivation for Theorem 1

We analyze the implications of smooth rationalization on consumer choices starting from the classical utility maximization problem. Suppose a consumer's choices are driven by a strictly increasing, strictly concave, and differentiable utility function U, subject to the budget constraint $p \cdot x \leq 1$, where x is a bundle of nonnegative commodities. The maximization problem is

$$\max_{x \in \mathbb{R}^K_+} U(x) \quad \text{s.t. } 1 - p \cdot x \geq 0 \tag{UM}$$

We can solve this problem using the method of Lagrange multipliers and the Karush–Kuhn–Tucker conditions. The Lagrangian is

$$\mathcal{L}(x, \lambda, \mu) = U(x) + \lambda(1 - p \cdot x) + \mu \cdot x\,.$$

Although the function U is defined only on $\mathbb{R}^K_+$, to assure sufficiency of the first order conditions, we need to add Lagrange multipliers for the nonnegativity of x, as in corner solutions, the equality between MRS and price ratio is not necessary.[11] At the optimal, λ is the marginal utility of

[10] Given a strictly concave function $f : \mathbb{R}^K_+ \to \mathbb{R}$, a vector $v \in \mathbb{R}^K$ is a *supergradient* of f at x if $f(x) > f(y) + v \cdot (x - y)$ for every $y \neq x$. If f its differentiable then for every $x \gg 0$ its gradient $\nabla f(x)$ is the only supergradient at x; this is not necessarily true if x is in the boundary of $\mathbb{R}^K_+$.

[11] An alternative approach is to think of maximizing a differentiable function f defined on set containing $\mathbb{R}^K_+$, such that f equals U on $\mathbb{R}^K_+$, subject to the budget constraint and the constraint $x \geq \mathbf{0}$. Since U is differentiable, concave, and monotone, it is continuously differentiable. Hence the existence of such f is given by Theorem 7 in Aversa et al. (1985).

income, and each component of the vector μ is the shadow cost of the corresponding nonnegativity restriction. The first-order conditions are

$$\begin{aligned} &\nabla_x \mathcal{L} = \nabla U(x) - \lambda\, p + \mu = \mathbf{0} && && \\ &\partial\mathcal{L}/\partial\lambda = 1 - p \cdot x \geq 0\,, && \lambda \geq 0\,, && \lambda\, \partial\mathcal{L}/\partial\lambda = 0 \\ &\nabla_\mu \mathcal{L} = x \geq \mathbf{0}\,, && \mu \geq \mathbf{0}\,, && \mu_k\, x_k = 0 \text{ for all } k \in [K] \end{aligned}$$

where $\nabla_x\mathcal{L}$ and $\nabla_\mu\mathcal{L}$ refer to the vectors of partial derivatives of $\mathcal{L}$ with respect to x and μ, respectively, and $\nabla U(x)$ is the gradient of U at x. By dual feasibility we have $\lambda \geq 0$ and $\mu \geq \mathbf{0}$. Moreover, as U is strictly increasing we can rule out $\lambda = 0$, so $\lambda > 0$ and $p \cdot x = 1$. Finally, $\mu \geq \mathbf{0}$ and $x \geq \mathbf{0}$ imply that $\mu_k\, x_k = 0$ for all k is equivalent to $\mu \cdot x = 0$.

Motivation for Sa-2) The counterexample in Figure 1 suggests that SSARP has to be relaxed to allow for corner solutions. Specifically, we need equality between price ratio and MRS only for goods with nonzero demand. Instead of requiring $p^i = p^j$ whenever $x^i = x^j$, the following weaker condition allows for differences between MRS and price ratio only in corner solutions:

$$\text{whenever } x^i = x^j, \text{ for every } k \text{ such that } x^i_k > 0 \text{ we require } p^i_k = p^j_k\,.$$

This is (D1).

If $\mathcal{D}$ satisfies SARP, (D2) is a condition imposed uniquely on the corner solutions. Although corner solutions do not require equality between MRS and price ratio, from the condition $\mu \geq \mathbf{0}$, we can obtain an upper bound on the vector of marginal utilities ∇U. As the utility is strictly concave, this bound generates additional data requirements, which are illustrated through Example 2 (shown in Figure 2).

Example 2. Take $\mathcal{D}$ comprised of three observations: $(p^1, x^1) = ((1/5, 4/3, 2/5), (5, 0, 0))$, $(p^2, x^2) = ((1/5, 1/2, 1), (1, 0, 0))$, and $(p^3, x^3) = ((1/7, 1, 1/2), (7/4, 1/2, 1/2))$. SARP holds as $x^3 \succsim^\star x^1 = x^2$ is the only revealed preference; the induced utility satisfies $U(x^3) > U(x^1) = U(x^2)$. Moreover, $x^1 = x^2 = (1, 0, 0)$ and $p^1_1 = p^2_1 = 1$, so (D1) holds as well. However, $q^3 \cdot x^1 = 5/7 \leq 1$ and $q^1 \cdot x^3 = 4/5 \leq 1$, hence (D2) fails.

Panel (a) of Figure 2 presents the data set in Example 2, which satisfies SARP and (D1), but fails (D2). As $\mathcal{D}$ satisfies SARP, we can interpret the choices as being driven by a strictly convex utility U, and from the revealed preferences, we have $U(x^3) > U(x^1) = U(x^2)$. In panel (b), we focus on the indifference set at bundle $x^1 (= x^2)$; specifically, we project it over the $x_2 = 0$ and

$x_3 = 0$ planes, the dimensions over which the consumer chooses zero consumption. Its projection into the $x_2 = 0$ plane is given because the indifference set is strictly convex and cannot intersect the interior of x^1's budget set. Similarly, as $x^1 = x^2$, the indifference set cannot intersect the interior of x^2's budget set either, which explains its projection into the $x_3 = 0$ plane. The reason why x^1's budget set is the one binding along the $x_2 = 0$ plane is that over that plane, the consumer can exchange x_1 for x_3, and x_3 is cheaper in the first observation (red budget set) than in the second (blue budget set), i.e., $p_3^1 < p_3^2$. Similarly, the blue budget set is the one bounding the projection along the $x_3 = 0$ plane, since in this plane the trade-off is between x_1 and x_2, and x_2 is cheaper in the second observation ($p_2^2 < p_2^1$). As $p_1^1 = p_1^2$, the slopes of the projections at the bundle are given by p_3^1/p_1^1 (along the $x_2 = 0$ plane) and p_2^2/p_1^1 (along the $x_3 = 0$ plane).[12] The generalization of this idea, that the lowest price is the one bounding the projection of the indifference class, drives the focus into the vector q^i for the definition of $\mathcal{D}_\wedge^S$ and (D2).

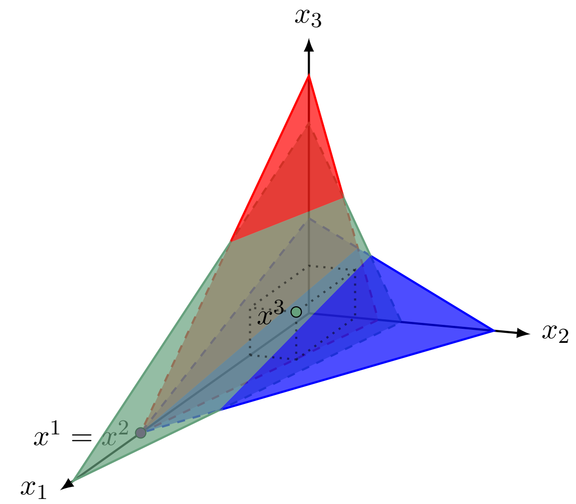
As U is strictly concave, we can find a supergradient at x^1, i.e., a plane that is tangent to x^1's indifference set at x^1, and such that any other bundle x in this plane satisfies $U(x) < U(x^1)$. If the utility is also differentiable, the indifference set also is. Hence, we can use the slopes of its projections to construct the tangent plane. Specifically, as the slopes of the indifference set at x^1 are bounded by $p_3^1/p_1^1 = q_3^1/q_1^1$ along $x_2 = 0$ and $p_2^2/p_1^1 = q_2^1/q_1^1$ along $x_3 = 0$, then one possible hyperplane is $q^1 \cdot x = 1$. Therefore we have $U(x^1) > U(x)$ for every x satisfying $q^1 \cdot x \leq 1$. As $q^1 \cdot x^3 < 1$, differentiability implies $U(x^1) > U(x^3)$, which contradicts the revealed preferences (since they imply $U(x^3) > U(x^1)$). Therefore $\mathcal{D}$ cannot be smoothly rationalized.[13]

Condition (D2) can also be explained in terms of the first-order conditions of the utility maximization problem. As $\mu \geq \mathbf{0}$, we can bound the vector of marginal utilities by $\nabla U(x) = \lambda p - \mu \leq \lambda p$. Denoting by $U_k(x)$ the partial derivative of U with respect to x_k, we have $U_k(x) \leq \lambda p_k$. Now suppose our data set has two observations i, j such that $x^i = x^j$; then $U_k(x^i) = U_k(x^j)$ implies $U_k(x^i) \leq \min\{\lambda^i p^i; \lambda^j p^j\}$. Furthermore, it is easy to see that $\lambda^i = \lambda^j$,[14] hence $U_k(x^i) \leq \lambda^i \min\{p^i; p^j\}$.

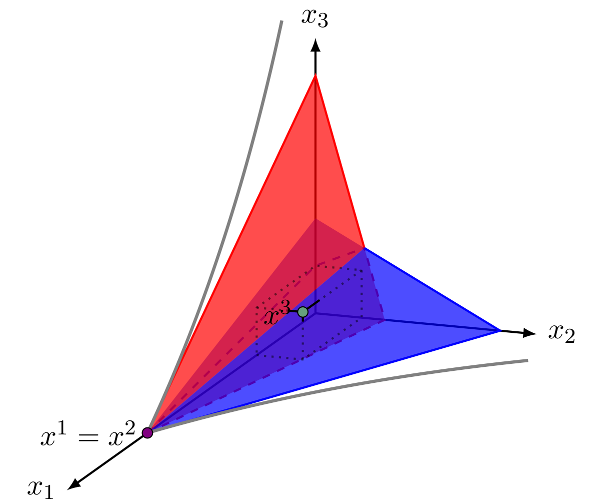
[12] The indifference set in Figure 2 is only one possibility. In general, the projections into the $x_2 = 0$ and $x_3 = 0$ planes do not need to be tangent to the budget set at the bundle x^1; they only need to be strictly convex and not go into the interior of the budget set.

[13] If the utility were not differentiable, then the hyperplane $q^1 \cdot x \leq 1$ is not necessarily below x^1's indifference set. As both x^1's and x^2's budget sets are below the indifference set, both p^1 and p^2 (and any convex combination of both) are still supergradients of U at x^1.

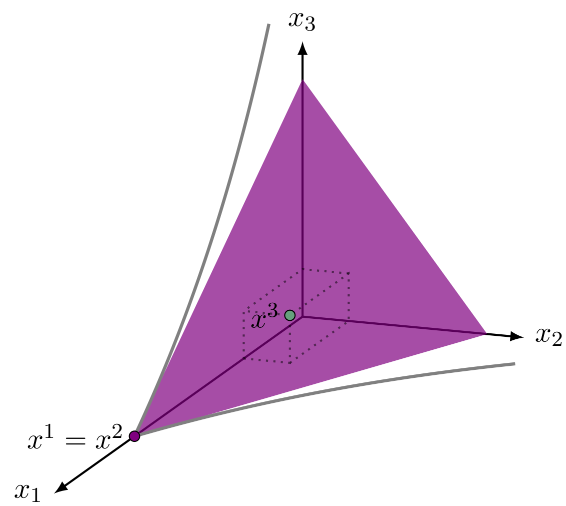
[14] This can be shown as follows: take $k \in [K]$ such that $x_k^i > 0$. Then $\mu_k^i = \mu_k^j = 0$, which implies (as $U_k(x^i) = U_k(x^j)$) that $\lambda^i p_k^i = \lambda^j p_k^j$. From (D1) we have $p_k^i = p_k^j$, therefore $\lambda^i = \lambda^j$. A formal argument (without requiring (D1)) is presented in the proof that Sa-1) implies Sa-2) in the Appendix.



(a) $\mathcal{D}$ satisfies SARP and (D1). Revealed preferences imply $U(x^3) > U(x^1) = U(x^2)$.



(b) Projections of x^1's indifference set into the $x_2 = 0$ and $x_3 = 0$ planes (in gray).



(c) Plane below x^1's indifference set (in purple). Strict concavity and differentiability imply $U(x^1) = U(x^2) > U(x^3)$.

Figure 2: Representation of Example 2. The data satisfies SARP and (D1), but is not smoothly rationalizable as it fails (D2).

Generalizing this argument we have

$$\nabla U(x^i) \leq \lambda^i q^i \,.$$

This is, $\lambda^i q^i$ is an upper bound for $\nabla U(x^i)$. By (D1), the prices of two observations with the same bundle differ only on the goods with zero demand. Using this fact and after some algebra (see equation (13) in the Appendix) we can conclude that strict concavity implies that, whenever $x \neq x^i$, we have $U(x^i) - U(x) > \lambda^i q^i \cdot (x^i - x)$. Thus $\lambda^i q^i$ is a supergradient of U at x^i. As $q^i \cdot x^i = 1$ and $\lambda^i > 0$, then $q^i \cdot x \leq 1$ implies $U(x^i) > U(x)$. Hence, a violation of SARP in $\mathcal{D}^S_\wedge$ makes it impossible to smoothly rationalize $\mathcal{D}$; smooth rationalization of $\mathcal{D}$ implies that $\mathcal{D}^S_\wedge$ satisfies SARP.

Motivation for Sa-3) The version of the Afriat inequalities in Sa-3) follow directly from the maximization problem and the first order conditions. Motivated by the maximization problem, but neglecting the nonnegativity constraints in the Lagrangian, Chiappori and Rochet (1987) and Matzkin and Richter (1991) use the numbers $u^i \in \mathbb{R}$ and $\lambda^i > 0$ satisfying the inequalities $u^i > u^j + \lambda^i(1 - p^i \cdot x^j)$ whenever $x^i \neq x^j$, and $u^i = u^j$ whenever $x^i = x^j$.[15] We follow the same approach but include the representation of the nonnegativity restrictions of the Lagrangian.

If the data is rationalizable, then each choice is optimal, and since U is strictly concave, each choice is the unique optimal solution given the price vector. Furthermore, from the first-order conditions the vector of marginal utilities $\nabla U(x)$ is equal to $\lambda\, p - \mu$. These conditions, and replacing $U(x^i) = u^i$ for all $i \in [N]$, are sufficient to motivate the existence of the Afriat inequalities:

(S1): Since U is strictly concave, each choice is the unique optimal solution from its budget set, i.e., $\mathcal{L}(x^i, \lambda^i, \mu^i) > \mathcal{L}(x^j, \lambda^i, \mu^i)$ whenever $x^i \neq x^j$. The condition follows from $p^i \cdot x^i = 1$ and $\mu^i \cdot x^i = 0$.

(S2): As $x^i = x^j$, both choices have the same utility.

(S3): As $x^i = x^j$, both choices have the same marginal utility.

(S4): As U is strictly increasing, marginal utilities are strictly positive.

(S5): It follows from the complementary slackness condition $\mu \cdot x = 0$.

From the Envelope Theorem, we know that the Lagrange multiplier λ associated with the budget constraint is the marginal utility of income. A consequence of conditions (S3) and (S5) is

[15] Chiappori and Rochet (1987) does not present the second relation, as the assumption of an invertible demand implies that we can assume that all observations have different choices without loss of generality.

that two observations with the same choice have the same marginal utility of income, even if the prices from which they were chosen are different.

Remark 2. If (S3) and (S5) hold, then $\lambda^i = \lambda^j$ whenever $x^i = x^j$.

The formal proof of the existence of the Afriat inequalities (S1) to (S5) starts from a data set that satisfies Sa-2). We construct the numbers using a variation of the algorithm in Varian's (1982) proof of Afriat's theorem.

3.4 Sufficiency of the modified data set for a differentiable utility

Since the vectors q^i, scaled by the marginal utility of income, bound the vector of marginal utilities ∇U, they might convey enough information to test for smooth rationalization. However, (D2) is not enough to obtain a differentiable utility: (D1) is also needed. The reason is that the definition of SARP rules out indifferences and therefore does not compare a bundle with itself (since each bundle is, by completeness, indifferent to itself).

Chiappori and Rochet (1987) leading example, presented in panel (a) of Figure 1, shows a data set that is not rationalizable but satisfies (D2). Although (D2) holds, the modified data set $\mathcal{D}^S_\wedge$, shown in Figure 3, presents the following problem: according to the vector q, which play the role of prices in $\mathcal{D}^S_\wedge$, the consumer does not spend all her income. Hence, according to q, a choice is directly revealed strictly preferred to itself, which is a clear contradiction. SARP does not incorporate a requirement for each choice to spend all its income because such property is necessary for any meaningful rationalization and therefore assumed in the original data set. However, the characteristics of our modifications to obtain $\mathcal{D}^S_\wedge$ may lead to an observation in which the consumer is "burning money", i.e., does not spend all her income, which is in clear contradiction to the monotonicity assumption. If this happens, i.e., if the consumer burns money, Remark 1 no longer holds, as a data set can satisfy SARP but will always fail GARP.

The following result shows that if the consumer spends all her money on the modified data set, i.e., if $q^i \cdot x^i = 1$ for all i, then such a data set is sufficient to test for a differentiable utility.

Proposition 1. *$\mathcal{D}$ satisfies Sa-2) if, and only if, $\mathcal{D}^S_\wedge$ satisfies SARP and $q^i \cdot x^i = 1$ for all $i \in [N]$.*

The proof of Proposition 1 is in Appendix D. This result shows that smooth rationalization by a strictly concave utility can be characterized using solely the characteristics of the modified data set.

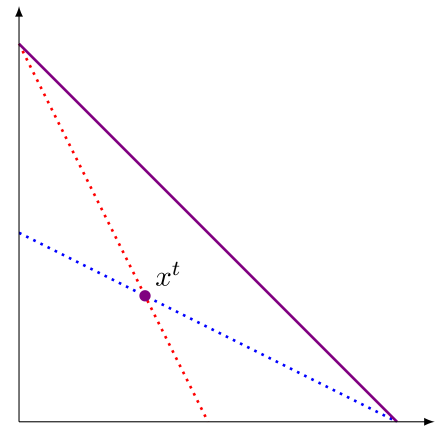


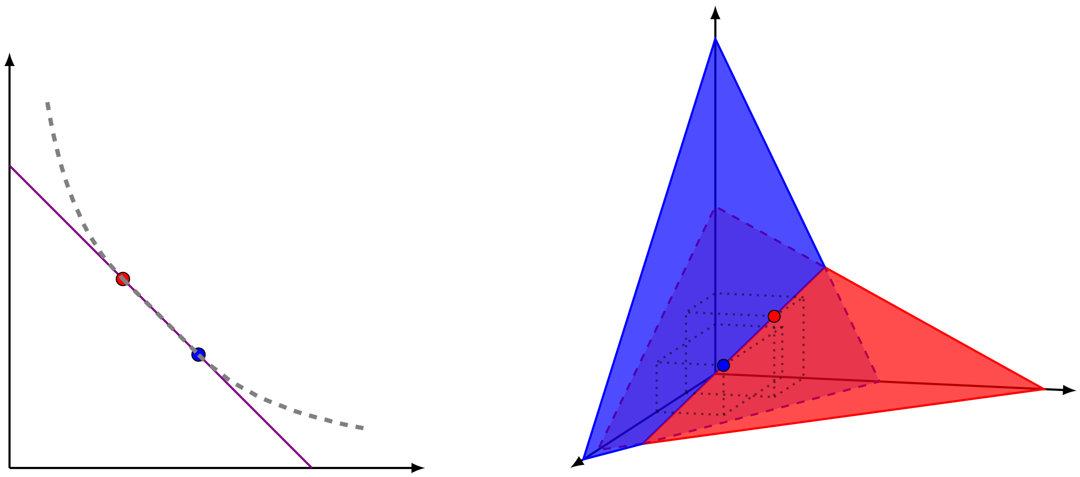
Figure 3: Modified data set of panel (a) in Figure 1.

4 The general case: GARP and Differentiable Utility

In this section, we characterize the most general case in our paper: rationalization by a strictly increasing, concave, and differentiable utility. Unlike the case of SARP, an invertible demand is not necessary nor sufficient for smooth rationalization.[16] Figure 4 presents two examples of data sets that satisfy GARP, fail SARP, and are invertible. Panel (a), which is motivated in Figure 2 in Chiappori and Rochet (1987), presents a data set in which two observations have the same price but different choices, i.e., it is not invertible. This data set is smoothly rationalizable, and the indifference curve between both choices has to be flat. On the other hand, panel (b) presents an invertible data set in which both choices are different and come from different prices. However, this data set cannot be smoothly rationalized since the two choices are revealed indifferent to each other; hence both have to be optimal solutions from the two prices in the data. Since both choices are interior solutions, each has to satisfy equality between MRS and price ratio for both prices, which is impossible as the two prices are different.

Panel (b) of Figure 4 shows that focusing on an invertible demand, i.e., on cases in which the same choice comes from two different prices, is not sufficient to characterize smooth rationalization. Furthermore, it suggests that violations of this property might arise from two choices that are revealed indifferent to each other. Our result in Theorem 2 shows that choices that are revealed indifferent to each other (including equal choices coming from different prices) are the only

[16] In footnote 1, Chiappori and Rochet (1987) suggest that a data set that satisfies GARP and presents an invertible demand can be rationalized by a differentiable utility: *"(...) if the data satisfies both GARP and condition (ii) of Definition 3* [$p^i \neq p^j \implies x^i \neq x^j$]*, there exists an infinitely differentiable, strictly increasing, concave utility function which rationalizes them (...)"* Panel (b) of Figure 4 shows that such a suggestion is incorrect.



(a) $\mathcal{D}$ satisfies smooth rationalization. (b) $\mathcal{D}$ fails smooth rationalization.

Figure 4: Examples of invertible data sets that satisfy GARP and fail SARP.

additional requirement over GARP to recover a differentiable utility. In order to include such differences, we first present a generalization of the modified data set used in the previous section, study its properties, and then use them to develop a test for smooth rationalization.

4.1 A modified data set using indifferences

As in the case of SARP, we propose a data set that replaces the original prices with new ones. There are two differences between the previous modification and the one we propose here. First, this one includes not only observations with the same choice but all observations that are revealed indifferent to each other. This idea leads to the following definition.

Definition 7. Take a data set $\mathcal{D} = (p^i, x^i)_{i\in[N]}$ and let $I(i) = \{j \in [N] : x^j \sim x^i\}$ be the set of observations with choices that are revealed indifferent to x^i, and $\tilde{q}^i = \bigwedge_{j\in I(i)} p^j$. The *one-step modification* of $\mathcal{D}$, $\Gamma(\mathcal{D})$, is given by $\Gamma(\mathcal{D}) = (\tilde{q}^i, x^i)_{i\in[N]}$.

The second difference between the modified data set we propose here and the one in the previous section is that the one-step modification Γ might create new revealed indifferences as prices decrease after the modification. If this is the case, the new revealed indifferences imply that a new modification is required. Hence, our modified data set needs to be defined as a fixed point.

Definition 8. For a data set $\mathcal{D}$, the *modified data set*, $D_\wedge$, is the fixed point of the one-step modification Γ, using $\mathcal{D}$ as the starting point.

We refer to the prices q^i in the modified data set as the modified prices and to the prices p^i as the original ones.

A natural question arising from Definition 8 is whether the aforementioned fixed point exists, i.e., if every data set has a corresponding modified data set. The answer to such a question is assured to be positive by the specific definition of revealed preferences we use. In Definition 1, we construct the revealed preferences by looking at the cost of a bundle at a price and assuming an expenditure of 1. However, it is possible for a data set $\Gamma(\mathcal{D}) = (\tilde{q}^i, x^i)_{i\in[N]}$ to have $\tilde{q}^i \cdot x^i < 1$. According to our definition of revealed preferences, we still construct this relation using an expenditure of one as a reference point; this is, x^i is revealed to x^j if $\tilde{q}^i \cdot x^j \leq 1$ even if $\tilde{q}^i \cdot x^i < 1$. This definition assures the existence of the fixed point since the one-step modification Γ expands the set of revealed indifferences. Therefore, in a worst-case scenario, the fixed point is reached when all choices are revealed indifferent to each other.

The following remark shows that the modified data set defined in this section is a generalization of the one used in the case of SARP.

Remark 3. If $\mathcal{D}$ is smoothly rationalizable by a strictly concave utility, then $\mathcal{D}_\wedge = \mathcal{D}_\wedge^S$.

The motivation for the modified data set comes from the idea that if x^i and x^j are revealed indifferent to each other ($x^i \sim x^j$) and x^j is affordable at prices p^i, then x^j is also an optimal choice when the price is p^i. Furthermore, suppose $\mathcal{D}$ is smoothly rationalizable. In that case, p^i gives us information about the vector of marginal utilities (i.e., the gradient of the utility function) at x^j, and we can combine the information from both p^i and p^j. As with SARP, the specific way the information from both prices has to be combined is by taking the minimum price in each dimension. Furthermore, we can establish a cyclicality argument: if $x^i \sim x^j$, then p^i is informative about the vector of marginal utilities at x^j, even if $p^i \cdot x^j > 1$ (GARP rules out $p^i \cdot x^j < 1$), and such information can be used by taking the coordinate-wise minimum of both prices.

The following result specifies the information we can obtain from modifying our data set.

Proposition 2. *If $\mathcal{D}$ is smoothly rationalizable by U, then $\mathcal{D}_\wedge$ also is.*

The proofs of Proposition 2 and of Corollary 1 below are in Appendix E. The proof of Proposition 2 proceeds by first showing that if U smoothly rationalizes $\mathcal{D}$ then it also smoothly rationalizes $\Gamma(\mathcal{D})$, and then uses that $\mathcal{D}_\wedge$ is a finite iteration of Γ on $\mathcal{D}$. The proof of the first part, that $\Gamma(\mathcal{D})$ is smoothly rationalizable by U, relies on a characterization of the set of the solutions of convex minimization problems provided by Jeyakumar et al. (2004).[17]

The main implication of the previous result is that a necessary condition to smoothly rationalize a data set is to be able to interpret choices as optimal not only from the original prices but also from the modified ones. In other words, $\mathcal{D}_\wedge$ has to satisfy GARP. This result formalizes the intuition of why the examples in Figure 3 and in panel (b) of Figure 5 are not smoothly rationalizable. In Figure 3 we have that, according to $\mathcal{D}_\wedge$, we have both $x^1(= x^2) \succ^\star x^3$ and $x^3 \succ^\star x^1(= x^2)$, a violation of GARP. In panel (b) of Figure 5 we have not only violations of GARP when comparing different observations, but also for every observation i we have $q^i \cdot x^i < 1$, which implies $x^i \succsim^\star x^i$ and $x^i \succ^\star x^i$, a violation of GARP (since x^i is revealed strictly preferred to itself).

The second consequence of Proposition 2 is that it provides tighter bounds on the marginal utilities than the ones obtained by looking at the original data set. Although the bounds on the marginal utilities have an unobservable variable (the marginal utility of expenditure), such bounds can still be used to estimate MRSs.

Corollary 1. *Suppose $\mathcal{D}$ is smoothly rationalized by U, and let λ^i be the marginal utility of money (i.e., the Lagrange multiplier of the budget set) when choosing x^i from price p^i. Then $\nabla U(x^i) \leq \lambda^i q^i$, and $U_k(x^i) = \lambda^i q^i_k$ whenever $x^i_k > 0$.*

The following section presents a characterization of smooth rationalization. The main component of such characterization is that GARP of the modified data set is both necessary and sufficient to rationalize $\mathcal{D}$ using a differentiable utility.

4.2 Smooth Rationalization

Theorem 2 establishes the equivalence between smooth rationalization and GARP of $\mathcal{D}_\wedge$. As in the case of SARP, smooth rationalization is also characterized by a modified version of the Afriat inequalities; for this characterization, we use the revealed indifferences that we infer using the modified data set $\mathcal{D}_\wedge$. Also similar to the case of SARP, we show that second- and higher-order derivatives have no empirical content.

Theorem 2. *The following are equivalent:*

Ga-1) $\mathcal{D}$ is smoothly rationalizable.

[17] Since the utility U is concave, $-U$ is convex. Hence the maximization of U has the same properties as the minimization of a convex function.

Ga-2) $\mathcal{D}_\wedge$ *satisfies GARP.*

Ga-3) Let $\sim_\wedge$ *denote the revealed indifferences according to the prices in* $\mathcal{D}_\wedge$*. There are numbers* $u^i \in \mathbb{R}$*,* $\lambda^i > 0$ *and* K*-dimensional vectors* $\mu^i \geq \mathbf{0}$ *such that*

$$u^i > u^j + \lambda^i(1 - p^i \cdot x^j) + \mu^i \cdot x^j \qquad \text{whenever } x^i \not\sim_\wedge x^j \tag{G1}$$

$$u^i = u^j \qquad \text{whenever } x^i \sim_\wedge x^j \tag{G2}$$

$$\lambda^i p^i - \mu^i = \lambda^j p^j - \mu^j \qquad \text{whenever } x^i \sim_\wedge x^j \tag{G3}$$

$$\lambda^i p^i - \mu^i \gg \mathbf{0} \qquad \text{for all } i \in [N] \tag{G4}$$

$$\mu^i \cdot x^i = 0 \qquad \text{for all } i \in [N] \tag{G5}$$

Ga-4) $\mathcal{D}$ *is smoothly rationalizable by an infinitely differentiable utility.*

The proof of this result is in Appendix F. The proof that Ga-1) implies Ga-2) follows from Proposition 2; the proof that Ga-2) implies Ga-3) proceeds by showing that similar numbers exist for the modified data set $\mathcal{D}_\wedge$, and then uses the replacement $q^i = p^i - \mu^i/\lambda^i$; the proof that Ga-3) implies Ga-4) follows closely the techniques in the proof of Chiappori and Rochet (1987), but using instead the numbers introduced in Ga-3); finally, that Ga-4) implies Ga-1) is immediate.

The first conclusion we can draw from the equivalence between statements Ga-1) and Ga-2) is that GARP of $\mathcal{D}_\wedge$ is necessary and sufficient for smooth rationalization. This equivalence gives us a simple test to check for differentiability of the utility function. After constructing the modified data set, the only further step is to check whether such a data set satisfies GARP. Unlike the characterization of smooth rationalization by a strictly concave utility in Theorem 1 and Proposition 1, statement Ga-2) has no requirements over the modified data set beyond GARP. This difference is because, in its definition, GARP includes comparing a bundle with itself, while SARP does not. Hence, the extra requirement $q^i \cdot x^i = 1$ for all $i \in [N]$ is implied by $\mathcal{D}_\wedge$ satisfying GARP.[18]

The original Afriat inequalities for rationalization are numbers $v^i \in \mathbb{R}$ and $\gamma^i > 0$ such that $v^i \geq v^j + \gamma^j(1 - p^i \cdot x^j)$. Several insights about smooth rationalization can be concluded from comparing such numbers to those in Ga-3). First, we can interpret both u^i and v^i as the utility levels and λ^i and γ^i as the marginal utility of expenditure. Then, the inequalities in the original Afriat's theorem are equivalent to (G1) and (G2), in the sense that they can be seen as inequalities

[18]We can see that GARP of $\mathcal{D}_\wedge$ implies $q^i \cdot x^i = 1$ by contrapositive. If $q^i \cdot x^i < 1$, then $x^i \succsim x^i \succ^\star x^i$, a violation of GARP.

coming from comparing the value of Lagrangian at the optimal versus its value at other bundle but using the same Lagrange multipliers.

The first difference between the original Afriat inequalities and the ones in Ga-3) is that here we compare bundles using two relations, a strict inequality in (G1) and an equality in (G2). In contrast, the original rationalization result uses only one weak inequality. In other words, smooth rationalization can fail if we assume indifference between choices that are not revealed indifferent to each other, even if such an assumption does not contradict the revealed preferences. We explain this result through the following example: take two observations i, j such that $x^i \succsim^\star x^j$ and $x^j \not\succsim x^i$; this is, we can only infer that x^i is revealed preferred to x^j. Furthermore, suppose x^j is an interior solution. Since $x^j \not\succsim x^i$, we conclude that $p^i \neq p^j$ and $x^i \neq x^j$. If we were to conclude that x^i and x^j are indifferent to each other, then we would need $p^i \cdot x^j = 1$ (if not, x^j is strictly cheaper and equally desirable as x^i under prices p^i). Then x^j is optimal at both prices p^i and p^j. However, as x^j is an interior solution, then the first order conditions imply equality between x^j's MRS and both p^i and p^j, which is impossible as p^i and p^j are different. A simple representation is presented in Figure 5. Not being able to assume indifferences is a strict restriction over the ordering between choices in GARP. If we only require rationalization, we can impose any ordering between the choices that do not violate the revealed preferences (Theorem 1 in Quah, 2014); imposing differentiability of the utility function imposes more restrictions (no further indifferences) on such ordering.

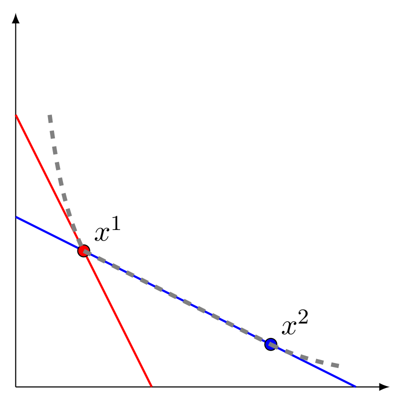
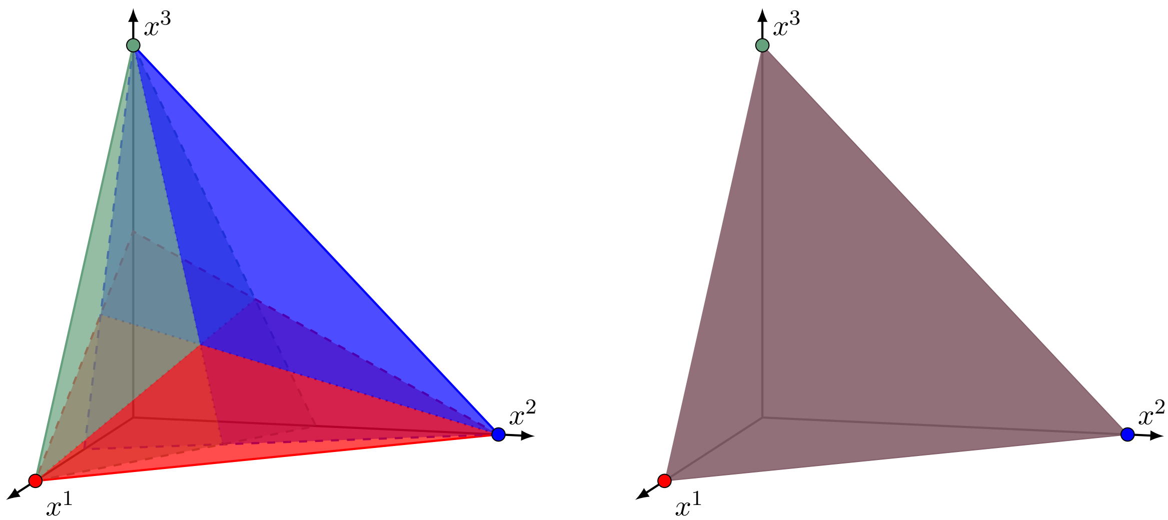


Figure 5: Any utility rationalizing $\mathcal{D}$ under which both choices are indifferent is not differentiable: smooth rationalization requires $u(x^2) > u(x^1)$.

As in Theorem 1, we can interpret the value $\lambda^i p^i - \mu^i$ in Ga-3) as the vector of marginal utilities at x^i. An implication of (G3) is that choices that are revealed indifferent to each other have not only the same utility but also the same marginal utility. This relation has two implications. First, since two choices that are revealed indifferent to each other have the same MRS, their indifference set has to be flat (i.e., an affine function) between them by the concavity of the utility. Examples of this property can be found in panel (a) of Figure 4 for two choices, and in Figure 6 for three choices. A second implication is that all choices have not only the same MRS but also the same marginal utility of expenditure. We show that choices indifferent to each other have the same marginal utility and the marginal utility of expenditure by relying on the characterizations of the set of optimal solutions from Jeyakumar et al. (2004). Finally, (G5) is the empirical equivalent to the complementary slackness condition of the nonnegativity restrictions in the utility maximization problem.



(a) Original data set. (b) Indifference set going through three choices.

Figure 6: Example of a data set with three different choices revealed indifferent to each other. The only possible indifference curve going through the three choices is the plane defined by them.

A final comment about Theorem 2 is that the only empirical component of a differentiable demand is the differentiability of the indifference sets. From an axiomatic perspective, smooth indifference sets are not sufficient for a smooth utility, as we also need the utility level to "vary smoothly". Dekel (1986) shows an example of a preference relation whose indifference sets are hyperplanes, therefore smooth, but cannot be represented by a differentiable utility. Furthermore, Neilson (1991) shows that differentiable indifference sets assure differentiability of the Hicksian demand but not of the Marshallian demand. On the other hand, a differentiable utility assures that both types of demand are differentiable. Such distinction cannot be made in empirical terms, as if the indifference sets are smooth, we can always construct a smooth utility.

Theorem 2 presents the broadest characterization in our paper. Besides differentiability, we only assume strict monotonicity and concavity of the utility function; these properties are usually assumed in applied work and are not testable in Afriat's theorem.[19]

5 Empirical Implementation

In this section we study the empirical relevance of smooth rationalization. In particular, we study smooth rationalization under SARP (Theorem 1) and the difference between smooth rationalization and SSARP. We limit our empirical analysis to the case of SARP because all subjects that satisfy GARP in our data set also satisfy SARP.[20] To simplify the exposition, we will refer to smooth rationalization by a strictly concave utility as Smooth SARP.

We analyze choices of 4,958 subjects coming from several sources. They all follow the design of Choi et al. (2007b): subjects have to choose between two different goods, observe a graphical representation of the budget set on the computer screen, and use the mouse to select a point in the budget line. We analyze two different choice environments. The first one is choices under risk, for which there are two states of the world, and the goods are Arrow securities for each state. The second environment is social choices, where the goods are the own consumption and consumption of an (anonymous) second subject; this situation is usually known as the *dictator game*, as the second subject cannot choose their consumption. Subjects make several choices in each type of experiment from randomly-drawn budget sets. After the experiment, one choice is randomly drawn from a uniform distribution, and payments are made according to that choice. In the case of risk choices, one state of the world is also randomly drawn.

For each choice environment (Risk and Social), we split our data set into two samples: one in which subjects are either undergraduate or graduate students, which we denote Students, and another in which the subjects are representative of the general population, denoted General. Our four samples are:

[19]It is unclear whether these properties, specially concavity, can be assumed without loss of generality. In other words, it is not evident whether there are data sets that can be rationalized by either a concave or a differentiable utility but not by one that is both differentiable and concave.

[20]As our data comes from experiments in which subjects choose in a two-good environment ($K = 2$), the only option for a subject to satisfy GARP and ail SARP is by facing the same price vector twice and making different choices. In our data set, no subject faces the same price vector twice.

- The Risk-Students sample: this sample comprises 1,020 subjects making 50 different choices each. Of the total, 974 subjects make choices in a symmetric environment where both states of the world are equally likely. The remaining 46 subjects face an asymmetric environment where one state has probability 2/3 and the other has probability 1/3. The data of subjects facing a symmetric problem is taken from the experiments in Choi et al. (2007a), Zame et al. (2020), Cappelen et al. (2021), and Dembo et al. (2021), and the data of subjects facing an asymmetric problem is taken from Choi et al. (2007a). All the subjects are undergraduate students in different universities.
- The Risk-General sample: this sample comprises 1,182 subjects making 25 choices each. The source of this data set is the experiment in Choi et al. (2014), where all subjects face a symmetric environment (equally likely states of the world). The sample is representative of the Dutch-speaking population in the Netherlands.
- The Social-Students sample: this sample comprises 1,058 subjects making 50 choices each. The sources for this data are the experiments in Fisman et al. (2007), Fisman, Jakiela, Kariv, and Markovits (2015), Fisman, Jakiela, and Kariv (2015), and Li et al. (2017). The samples in Fisman et al. (2007) and Fisman, Jakiela, and Kariv (2015) are comprised of undergraduate students; the one in Fisman, Jakiela, Kariv, and Markovits (2015) of Law students; and the one in Li et al. (2017) of medical students.
- The Social-General sample: this sample comprises 1,698 subjects making 50 choices each. The sources are the experiments in Fisman et al. (2017), comprised of 1,002 subjects, and the experiment in Fisman et al. (2022), of 696 subjects. Both experiments are embedded in the American Life Panel (ALP), an internet survey administered by the RAND Corporation to adult Americans.

Since our data set comprises experiments in which subjects choose only between two goods, the comparison between Smooth and Strong SARP indicates the relevance of corner solutions. The reason for this is that choices satisfying Smooth SARP and failing Strong SARP have to present two choices that are equal, are a corner solution (i.e., one of the two goods is not consumed), and come from different prices (which differ in the price of the good that is not consumed).

Remark 4. If $K = 2$, then condition (D2) in Sa-2) is vacuously satisfied, i.e., if $\mathcal{D}$ satisfies SARP and (D1), then it satisfies (D2).

In order to differentiate between SARP and Smooth SARP, we need two choices which are equal, interior solutions, and come from different prices. Similarly, in order to differentiate between

Smooth and Strong SARP, we need two choices which are equal, corner solutions, and come from different prices. In both cases, we need budget sets to intersect exactly where the choice was made. In such a sense, an ideal experiment to differentiate between these axioms would deliberate generate budget sets that intersect with previous choices, which the experimental design we rely on does not do. However, our data set still has several situations in which two choices are equal and come from different budget sets, implying that price vectors intersect where these choices were made.

5.1 Rationalizable Choices

We first focus exclusively on the subjects whose choices are rationalizable, which is the main focus of our paper. Table 1 presents summary statistics by data source. We observe that a vast majority (91.9%) of the subjects which satisfy SARP also satisfy Smooth SARP, i.e., have choices that a smooth and strictly concave utility can rationalize. On the other hand, the percentage of subjects which satisfy Strong SARP is significantly lower: 51.7%. Column (7) of the table measures the subjects which satisfy Smooth SARP and fail Strong SARP as a percentage of the subjects which satisfy SARP and fail Strong SARP. In other words, it measures, among the subjects whose choices are rationalizable but fail Strong SARP, the percentage of subjects whose choices can be rationalized by a differentiable utility. In total, 83% of rationalizable subjects who fail Strong SARP satisfy Smooth SARP. These results suggest that, in general, corner solutions are common, and hence Strong SARP might not be an adequate test for differentiability.

When comparing different samples, we observe that differences between Smooth and Strong SARP are wider for the Social environment than the Risk one, and for the Students sample than for the General one. The difference between social and risk choices is driven by the larger share of corner solutions in the first. Such corner solutions are mostly "selfish" choices in which the subject chooses the allocation that maximizes their consumption, giving zero to the other party. When conditioning on satisfying SARP, 78.2% of choices are "selfish" in the Social-Students sample, and such choices are 45.9% in the Social-General one.[21] On the other hand, all corner solutions (i.e., choices with one of the components being equal to zero) are a 39.4% of the choices in the Risk-Students data and a 9.2% in the Risk-General one (again, focusing only on subjects that satisfy

[21] In both data sets of social choices, there are also "altruistic" choices. In such choices, the subject assigns as much as possible to the second player. However, such choices are less common: 5.8% for the Students sample and 1.7% for the General one. Such choices can be explained by subjects whose behavior resembles a utilitarian preference *a la* Harsanyi (1955).

Table 1: Subjects Satisfying Axioms by Data Source

Source	(1) N	(2) SARP	(3) Smooth SARP	(4) Strong SARP	(5) (Smooth SARP) / (SARP)	(6) (Strong SARP) / (SARP)	(7) (Smooth - Strong) / (SARP - Strong)
Risk - Students	1020	258	235	169	91.1%	65.5%	74.2%
Risk - General	1182	231	212	204	91.8%	88.3%	29.9%
Risk - All	2202	489	447	373	91.4%	76.3%	63.7%
Social - Students	1058	346	335	87	96.8%	25.1%	95.7%
Social - General	1698	146	120	47	82.2%	32.2%	73.7%
Social - All	2756	492	455	134	92.5%	27.2%	89.7%
TOTAL	4958	981	902	507	91.9%	51.7%	83.2%

*N is number of subjects; SARP is number of subjects who satisfy SARP; Smooth SARP is number of subjects whose choices are smoothly rationalized by a strictly concave utility; Strong SARP is number of subjects who satisfy Strong SARP. Last column measures share of smoothly rationalizable subjects among rationalizable subjects who fail Strong SARP.

SARP). Since corner solutions drive differences between Smooth and Strong SARP, the prominence of "selfish" choices generates the empirical difference between both types of environments.

The difference between the Students and the General samples, with the Students samples presenting more corner solutions in both choice environments, is driven by different preference profiles in each case. For choices under risk, we can interpret the corner solutions as individuals behaving according to preferences with low risk aversion or risk neutrality. In the Students sample, the high share of corner solutions (39.4%, vs. 9.2% in the General sample) also translates into a higher share of subjects who present only corner solutions, i.e., who behave as risk neutral. Risk neutral subjects are 12.0% of the subjects satisfying SARP in the Risk-Students sample, but there are no such subjects in the Risk-General sample. Another possible explanation for the difference between the Students-Risk and General-Risk samples might be the different number of choices: the General sample has only twenty-five choices per subject. By having fewer choices, the General sample may have fewer situations in which budget sets intersect at the same point the choice is made. However, the difference in corner solutions, along with the fact that the share of subject satisfying Smooth SARP (column (5) in Table 1) is similar between both samples, make this explanation unlikely.

For the case of Social Choices, the main factor explaining the difference between the Students and the General sample is the presence of "selfish" choices, i.e., choices in which the subject assigns as much as possible for herself. Not only the share of corner solutions is higher for the Students sample (78.2% vs. 45.9%), but also the share of subjects who present only selfish choices, i.e., who behave according to "selfish" preferences is higher: 45.4% for the Students sample and 17.1% for the General one.

One difference between the Students and the General sample that we observe for Social choices and not for Risk choices is that, in the Social choices environment, the share of subjects that satisfy Smooth SARP is significantly lower for the General sample than for the Students one. This fact is partly driven by the presence of choices resembling a Leontief utility. The Leontief utility has different interpretations in both settings: in the Risk environment, it represents infinite risk aversion, and in the Social one, it represents egalitarian preferences. In both environments, these preferences predict that the subject will choose the same amount of both goods, i.e., that she will choose in the 45° line. Since choices are made graphically, we measure such choices with slack: we classify choices as egalitarian if the ratio x_1/x_2 is above $1/1.05$ and below 1.05. For both Risk and Social environments, the General sample presents a higher share of Leontief choices than the Students one; however, this difference is higher in the Social environment than in the Risk one. For the Risk environment, the General sample has 39.7% of Leontieff choices, and the Students sample has 21.1%. In the Social environment, such choices are 30.3% in the General sample and 3.0% in the Students sample. Furthermore, in the Risk environment, the difference between samples regarding subjects whose choices are all Leontieff is small: 4.3% for General and 4.7% for Students. On the other hand, this difference is considerably more prominent in the Social environment: 14.4% for General and 0.9% for Students.

We finish this section by presenting some examples of subjects whose choices follow a pattern similar to widely used families of preferences and explain how they either satisfy Smooth SARP and fail Strong SARP, or fail both axioms. These examples are presented in Figure 7; pairs of observations with the same choice coming from different prices are highlighted, and their prices are presented. Panels (a) and (b) present subjects whose choices are all corner solutions. Panel (a) presents a subject choosing in the Risk environment whose choices resemble risk neutral preferences, and panel (b) presents a subject choosing in the Social environment whose choices resemble "selfish" preferences. In both cases, there are pairs of observations with the same corner choice and different price vectors, which violates Strong SARP but not Smooth SARP. Panels (c) and (d) present subjects with pairs of observations with equal interior choices and different price vectors, which violates Smooth and Strong SARP. Both subjects make choices in the risk environment.

Panel (c) presents a subject choosing (approximately) according to a Leontief utility, i.e., all her choices are over the 45° degree line.[22] panel (d) presents a subject whose choices resemble safety-first preferences *a la* Roy (1952): this person assures a minimum consumption level, in this case roughly ten tokens, and for any level above that her behavior resembles risk-neutral preferences. As expected, the examples presented here are extreme cases in which choices are very close to a specific model. For most subjects, patterns are more complex than those presented in Figure 7, and it is unclear whether they fit a specific family of preferences.

5.2 Nonrationalizable Choices

We now include subjects who fail GARP, i.e., subjects whose choices cannot be rationalized, in our analysis. In order to do this, we recover choices using the Houtman and Maks (1985) Index (HM Index). In simple terms, the HM Index looks for the smallest subset of observations that needs to be removed such that the remaining observations satisfy the axiom in question (in our case, either SARP, Smooth SARP, or Strong SARP).[23] Usually, the HM Index, as other measures of distance from GARP, is interpreted as a measure of distance from economic rationality (see Kariv & Silverman, 2013, for a discussion of this topic).

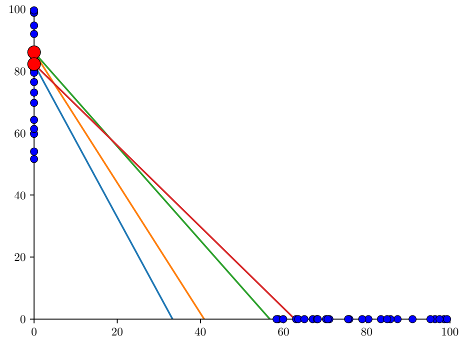
We choose the HM Index to compare different axioms for practical purposes. Although the question about how to recover preferences from choices that fail GARP is an open question in the literature, the most common methods all rely on the idea of *partial efficiency* introduced by Afriat (1973).[24] Along with the HM Index, the most popular partial efficiency methods are the Critical Cost Efficiency Index (Afriat, 1973) and the Varian (1990) Index. (Halevy et al., 2018) argue that among them, the Varian index is the most suitable to recover preferences, as it gives tighter bounds on the set of preferences recovered. However, as shown by Ugarte (2022b), under the Varian Index, strict convexity and differentiability cannot be falsified. In other words, we can always find preferences that achieve the value of the Varian Index and present these properties. On

[22] Although the example in panel (c) is taken from the Risk environment, representing infinite risk aversion, the same pattern is present in some choices in the Social environment, representing egalitarian preferences.

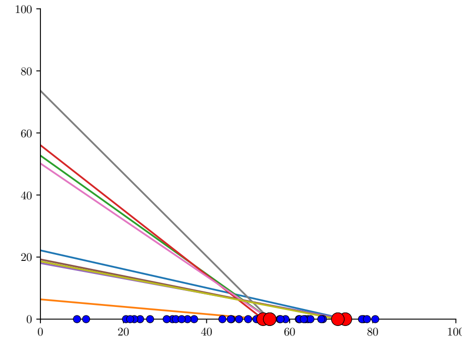
[23] Formally, for a data set $\mathcal{D}$ comprised of N observations and an Axiom $\mathcal{A}$, we have

$$HM(\mathcal{D}) = \min_{E \subset [N]} |E| \quad \text{s.t. } (p^i, x^i)_{i \in E} \text{ satisfies } \mathcal{A}.$$

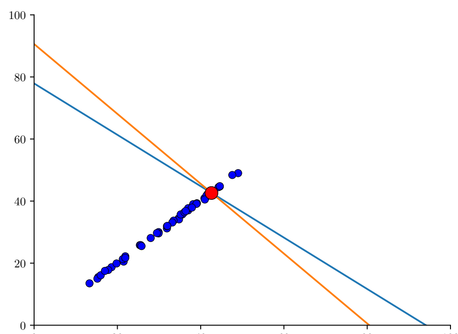
[24] de Clippel and Rozen (2021) and Ugarte (2022a) present methods that do not rely on partial efficiency.



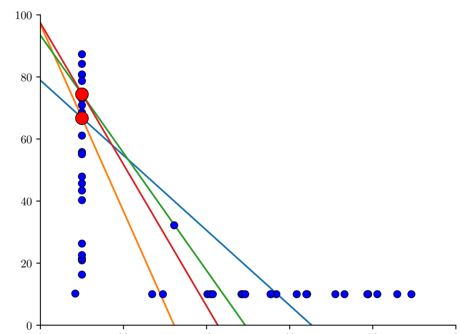
(a) Risk Neutral Preferences - Satisfies Smooth SARP, fails Strong SARP. Subject ID 221 from Dembo et al., 2021



(b) "Selfish" Preferences - Satisfies Smooth SARP, Fails Strong SARP. Subject ID 119 from Fisman et al. (2007)



(c) Leontief Preferences - Fails Smooth and Strong SARP. Subject ID 109 from Choi et al. (2007a)



(d) Safety-First Preferences - Fails Smooth and Strong SARP. Subject ID 39 from Cappelen et al. (2021)

Figure 7: Examples of Choice patterns that mimic common preferences in economic modelling. Panels (a) and (b) mimic differentiable preferences which fail Strong SARP. Panels (c) and (d) mimic nondifferentiable preferences, hence fail both Smooth and Strong SARP.

the other hand, by removing observations, the HM Index allows us to differentiate between SARP, Smooth SARP, and Strong SARP, even for choices that fail GARP.

To compute the HM Index, we rely on the techniques introduced by Demuynck and Rehbeck (2021). Furthermore, in the case of Smooth SARP, we leverage the fact that with two dimensions, SARP of the original data set and condition (D1) characterize Smooth SARP. Hence we first remove all violations of (D1) and then compute the HM Index.[25] We follow a similar procedure for Strong SARP but remove observations to make the data set invertible instead of only satisfying (D1).[26]

Table 2 presents the summary statistics of the HM Index for the different axioms and data sources. Across data sources, we observe that the difference between the average HM Index for Smooth SARP and SARP is smaller than the one between Smooth and Strong SARP. The only exception is the Risk-General sample, in which Smooth and Strong SARP have similar average levels. This pattern can be explained because, as explained before, the Risk-General data set presents a lower share of corner solutions than the other two. We also find that all differences between HM Indices for different axioms are higher for the Social than the Risk environment. Furthermore, the difference between Strong and Smooth SARP is the one that changes the most between Risk and Social choices, suggesting that corner solutions are still more relevant in the Social environment when we consider nonrationalizable subjects. When performing paired t-test for the different axioms, the difference between HM Indices for different axioms is statistically different from zero at the 0.1% significance level for all samples.[27]

Figure 8 presents the (reversed) cumulative distribution of the HM Index for the different samples. Since the HM Index measures distance from economic rationality, subjects with a lower index value (to the right of the graph) are closer to economic rationality (i.e., closer to satisfying GARP). Each bar represents the subjects with an HM Index that is less or equal to the given value, i.e., subjects at that distance or closer to rationalization. In particular, the bars at the zero level of the index represent the subjects consistent with the corresponding axiom, i.e., they coincide with the numbers in Table 1.

[25] Specifically, we need to compute candidates for the HM Index for each possible combination of observation removals to satisfy (D1). The HM Index is the minimum among such candidates.

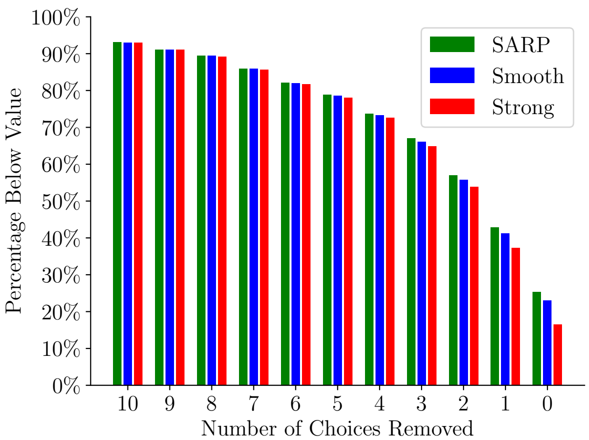
[26] For Strong SARP, in the case of several observations with the same corner solution, we always keep the observation with the higher price of the good with zero consumption. By doing it, we eliminate the higher number of revealed preferences, making it more likely for the remaining data set to satisfy this axiom.

[27] The difference in the average HM Index between the Students and General samples for both environments goes in line with previous evidence that the general population presents a lower measure of economic rationality than the sub-sample with a higher educational level. See for example Li et al. (2017) and Li et al. (2022).

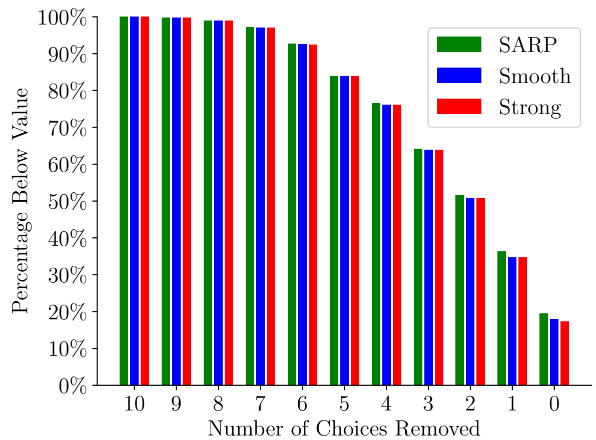
Table 2: Houtman Maks Index by Data Source

Source	(1) N	(2) SARP	(3) Smooth SARP	(4) Strong SARP
Risk - Students	1020	3.33 (0.121)	3.39 (0.121)	3.55 (0.119)
Risk - General	1182	2.79 (0.067)	2.84 (0.067)	2.85 (0.066)
Risk - All	2202	3.04 (0.067)	3.10 (0.067)	3.17 (0.066)
Social - Students	1058	3.37 (0.119)	3.46 (0.119)	5.75 (0.139)
Social - General	1698	6.81 (0.113)	6.90 (0.112)	7.85 (0.108)
Social - All	3073	5.04 (0.084)	5.12 (0.084)	6.43 (0.085)
TOTAL	4958	4.40 (0.060)	4.48 (0.060)	5.32 (0.063)

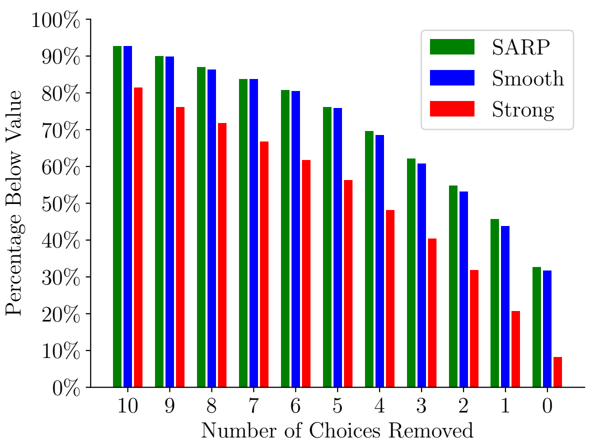
*N is number of subjects; SARP, Smooth SARP, and Strong SARP refer to the average of the Houtman-Maks Index for each type of Axiom. Standard errors in parenthesis.



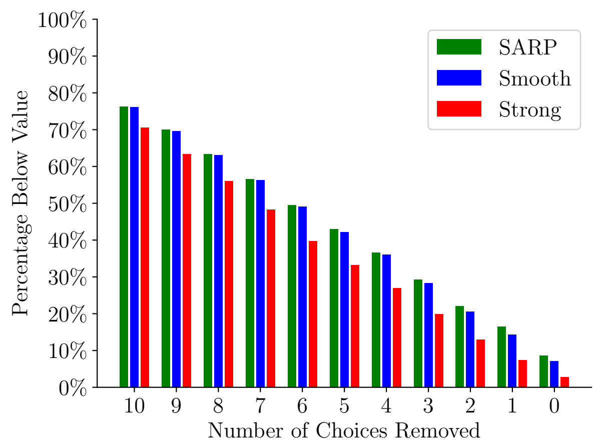
(a) Risk - Students (50 choices)



(b) Risk - General (25 choices)



(c) Social - Students (50 choices)



(d) Social - General (50 choices)

Figure 8: Percentage of Subjects with Houtman and Maks (1985) Index below value for each data set and different axioms.

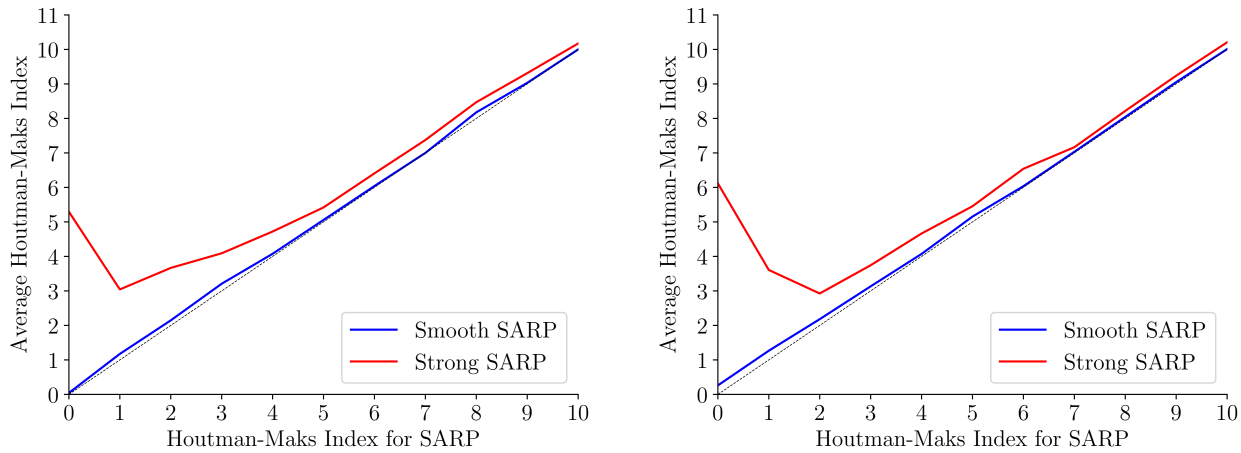
Our first conclusion from Figure 8 is that, for all samples, the HM Index for SARP and Smooth SARP behaves similarly. This result suggests that the cost of adding differentiability to the utility function is small in terms of the HM Index. Furthermore, we also observe that this difference decreases as we increase the cutoff value of the HM Index. One possible explanation is that subjects further away from GARP (i.e., with a higher HM Index) may present more erratic behavior. If this is the case, the probability of having two observations with the same choice decreases as the HM Index increases, making Smooth and Strong SARP harder to test.

The main differences in Figure 8 are between choice environments. The Social environment generally presents a higher distance between Smooth and Strong SARP. Furthermore, although it decreases, such distance is still large when we include subjects with higher HM Index. While the share of subjects with HM Index below a cutoff in the Risk environment tends to equalize for both axioms as we increase the value of such cutoff, in the Social environment, the difference remains significant. For students, the HM Indices for SARP and Smooth SARP behave similarly (in terms of their cumulative distribution), and the main difference is in the value of the HM Index for Strong SARP. For the general population, the index is lower (i.e., a higher share of subjects are below a given value) in the Risk than in the General environment.[28]

To finish this section, we focus specifically on the Social environment, where differences between the HM Index for different axioms are larger. Specifically, we analyze the difference between SARP, Smooth, and Strong SARP by looking at the average HM Index for Smooth and Strong SARP conditional on a given value of the HM Index for SARP. This is shown in Figure 9. For both panels, the average values of the HM Index for Smooth and Strong SARP get closer to the average value for SARP as the value of the HM Index for SARP increases. In other words, differences between SARP, Smooth SARP, and Strong SARP are harder to differentiate for subjects with a higher HM Index.

A surprising feature of Figure 9 is that, for both samples, the average value of the HM Index for Strong SARP decreases when we increase the value of the index for SARP from zero to one (the same pattern follows from one to two for the General sample). This pattern is generated because "selfish" choices are more common in subjects with lower HM Index; this is, the share of selfish choices decreases as we increase the value of the HM Index. Furthermore, for both samples, all the subjects with "selfish" preferences (i.e., who present only "selfish" choices) are subjects whose HM

[28]Since in the Risk-General sample subjects make only 25 choices each, removing one observation is, as a share, twice as relevant as in the other samples.



(a) Social - Students

(b) Social - General

Figure 9: Average Houtman and Maks (1985) Index for Smooth and Strong SARP, conditional on value of the same index for SARP.

Index is equal to zero, i.e., who satisfy SARP.

Although our empirical analysis comes from an experimental design that is not ideal for testing differences between SARP, Smooth SARP, and Strong SARP, some patterns are consistent throughout our analysis. The first one is that the differences between Smooth SARP and SARP seem small but not zero. Hence, the assumption of differentiability is not general from an empirical standpoint but encloses a higher share of the subjects. Second, our results suggest that Strong SARP is too restrictive to test for a differentiable utility: many subjects which satisfy Smooth SARP fail Strong SARP. Equivalently, our results suggest that corner solutions are a relevant component of the choices, especially for Social choices. This fact is relevant as models in applied economics that rely on first-order conditions usually assume interior solutions. To assess if omitting corner solutions are relevant for any specific example is an area for future research.

6 Extensions

We finish this paper by presenting two extensions of our analysis to widely used families of utility functions: homothetic and quasilinear utilities. Homothetic utilities have the property that the marginal rate of substitutions depends only on the share between goods, not their level. Hence, the ratio of goods in the optimal consumption bundle depends only on the price ratio, not the income level. Quasilinear utilities have a numeraire good, and their indifference sets are parallel displacements along the axis of this good. In our model, the numeraire good is measured in the same income scale; therefore, the observed demand does not depend on wealth. Homothetic and quasilinear utilities are the only two cases in which the consumer surplus is a valid measure of welfare (Silberberg, 1990, see Section 11.5 in).

The rationalization tests have a cyclical structure for homothetic and quasilinear utilities. In each case, we first show that the cycles in the test can be used to infer indifferences in the data beyond the ones obtained by the revealed indifference relation. After that, we show that the same intuition used in the cases of SARP and GARP applies here. We use all the inferred indifferences to modify the data set by taking the meet of prices among indifferences (until finding a fixed point). As in the previous sections, the test for differentiability of the utility function is the original rationalization test applied to the modified data set.

6.1 Homothetic utility

A utility function u is *homothetic* if it is a monotonic transformation of a function that is homogeneous of degree one; this is, if $u(x) = f(g(x))$, where $f : \mathbb{R} \to \mathbb{R}$ is strictly increasing and $g : \mathbb{R}^K_+ \to \mathbb{R}$ is homogeneous of degree one. As utility representations are invariant to monotonic transformations (utility is an ordinal representation of preferences), we focus on utilities that are homogeneous of degree one. Hence we use the terms homothetic and homogeneous of degree one interchangeably.

Varian (1983a) shows that the Homothetic Axiom of Revealed Preferences, HARP, is a test for rationalization by a homothetic utility.[29]

Definition 9. $\mathcal{D}$ satisfies the *Homothetic Axiom of Revealed Preferences, HARP*, if for any sequence of different observations $(m_\ell)_{\ell\in[L]}$,

$$(p^{m_L} \cdot x^{m_1})(p^{m_1} \cdot x^{m_2})(p^{m_2} \cdot x^{m_3}) \dots (p^{m_{L-1}} \cdot x^{m_L}) \geq 1\,.$$

Furthermore, Varian (1983a) shows that if $\mathcal{D}$ is rationalizable by a homothetic utility, such utility can always be chosen to be strictly increasing, continuous, and convex, and that there are numbers $u^i > 0$ such that $u^i \leq u^j\, p^j \cdot x^i$ for all i, j.[30] Knoblauch (1993) characterizes the set of all

[29]When characterizing homothetic utility, Varian (1983a) points to several bibliographical remarks regarding previous related results. In particular, he states that Afriat (1981) already presented a version of rationalization by a homothetic utility.

[30]The motivation for these numbers is that if U is homothetic, then the marginal utility of income λ equals the utility $U(x)$. The numbers can be obtained by replacing this equality into the Afriat inequalities.

preferences that are homothetic and rationalize $\mathcal{D}$ for a given data set that satisfies HARP.

We present an example to understand the motivation for HARP and the restrictions it imposes to achieve differentiability. Suppose we have three observations i, j, m such that

$$(p^i \cdot x^j)(p^j \cdot x^m)(p^m \cdot x^i) = 1$$

and a homothetic utility U rationalizing such choices. As $p^i \cdot \left((p^j \cdot x^m)(p^m \cdot x^i)x^j\right) = 1$, revealed preferences imply $U(x^i) \geq U\left((p^j \cdot x^m)(p^m \cdot x^i)x^j\right)$. Moreover, since U is homothetic and x^j is the optimal choice from p^j, then $(p^j \cdot x^m)(p^m \cdot x^i)x^j$ is optimal from $\left[(p^j \cdot x^m)(p^m \cdot x^i)\right]^{-1} p^j$; in particular, as $\left[(p^j \cdot x^m)(p^m \cdot x^i)\right]^{-1} p^j \cdot (p^m \cdot x^i)x^m = 1$, we have $U\left((p^j \cdot x^m)(p^m \cdot x^i)x^j\right) \geq U\left((p^m \cdot x^i)x^m\right)$. Similarly, as x^m is optimal from p^m, homotheticity implies that $(p^m \cdot x^i)x^m$ is optimal from $(p^m \cdot x^i)^{-1} p^m$, thus $U\left((p^m \cdot x^i)x^m\right) \geq U(x^i)$. Collecting these inequalities yields

$$U(x^i) \geq U\left((p^j \cdot x^m)(p^m \cdot x^i)x^j\right) \geq U\left((p^m \cdot x^i)x^m\right) \geq U(x^i) \qquad (2)$$

If choices are rationalizable by a homothetic utility, then the previous (weak) inequalities must be equalities. Thus, if instead of starting from the product being equal to one, we start from $(p^i \cdot x^j)(p^j \cdot x^m)(p^m \cdot x^i) < 1$, then the first inequality in the above restriction would be strict, implying that rationalizing $\mathcal{D}$ with a homothetic utility is not possible. The exercise in the previous example is generalized in Lemma 7 in Appendix G.

Since rationalization by a homothetic utility implies that the inequalities in (2) have to be equalities, we can infer indifferences between x^i and projections of x^j and x^k. Furthermore, if we assume differentiability, we obtain that x^i, x^j, and x^m have the same MRS (for interior solutions). Since $(p^j \cdot x^m)(p^m \cdot x^i)x^j$ is affordable at price p^i, then it is optimal from such price. Hence the MRS at this bundle (and thus also at x^j) is equal to the price ratio in p^i, which is also equal to the MRS at x^i. Similarly, we can conclude equal MRS at x^i and $(p^m \cdot x^i)x^m$, and hence also at x^m. Figure 10 presents a simple example in which a homothetic rationalization exists but cannot be smooth.

As in the previous sections, the critical component to characterize rationalization by a differentiable utility is the existence of indifferences. For the reasons explained in the previous paragraph, we can infer indifferences by looking at sequences of observations like the ones analyzed in the definition of HARP. Also, the indifferences inferred from the revealed indifference relation in Definition 1 can also be inferred through the cycles analyzed in HARP.[31] Hence, it is sufficient to

[31] Suppose $(p^i \cdot x^j)(p^j \cdot x^m)(p^m \cdot x^i) = 1$ and we know $x^i \sim x^j \sim x^m$. GARP implies (without loss of generality)

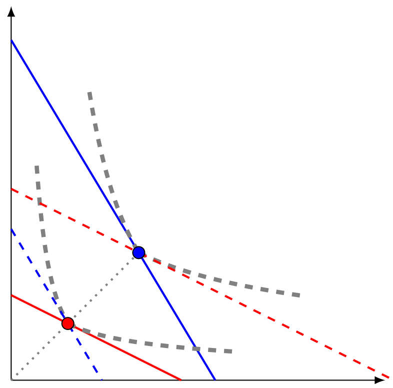


Figure 10: The data set can be rationalized by a homothetic utility. However, it cannot be differentiable as both choices cannot have the same MRS.

only focus on such cycles. We will refer to sequences that allow us to infer indifferences in the homothetic case as H1-sequences (i.e., HARP sequences that are less or equal to one).

Definition 10. An *H1-sequence* is a sequence of unique observations $(m_\ell)_{\ell\in[L]}$ such that

$$(p^{m_L}\cdot x^{m_1})(p^{m_1}\cdot x^{m_2})\dots(p^{m_{L-1}}\cdot x^{m_L})\leq 1$$

The reason to include a weak inequality in the previous definition (instead of equality) is analogous to the idea of not changing the definition of revealed indifferences in the creation of our modified data set (Definition 8). Although the inequality in the previous definition has to be an equality to satisfy HARP, defining the sequences in this form will ensure that creating the modified data set for homothetic rationalization stops in a finite number of steps. Furthermore, as in the case of GARP (Section 4), if at any step in the process of modifying the data set we have an H1-sequence for which the inequality is strict, we already know that such data set cannot be smoothly rationalized by a homothetic utility.

The motivation for creating a modified data set Section 4.1 is that indifferences between choices in the data imply equality of the MRSs. Along with concavity, this equality makes the indifference set linear. Since assuming a homothetic utility implies that we can infer indifferences between one choice and a scaled version of another (through H1-sequences), we need to add such indifferences in our analysis. To simplify our notation, we will focus directly on prices instead of choices. Following Varian (1982) (Section 3), we refer to a price as revealed indifferent to another if we can infer that the optimal choice from the former is indifferent to the optimal choice of the latter.

$p^i\cdot x^j = p^j\cdot x^m = p^m\cdot x^i = 1$. Therefore $U(x^i) = U((p^j\cdot x^m)(p^m\cdot x^i)x^j) = U((p^m\cdot x^i)x^m)$ is equivalent to $U(x^i) = U(x^j) = U(x^m)$.

The example motivating this is $(p^i \cdot x^j)(p^j \cdot x^m)(p^m \cdot x^i) = 1$. Focusing on observation i, this example generates the H1-sequence $s = (j, k, i)$. Since we infer that $(p^j \cdot x^m)(p^m \cdot x^i)x^j$ is both optimal from $(p^j \cdot x^m)(p^m \cdot x^i)p^j$ and indifferent to x^i, the agent is indifferent between prices p^i and $(p^j \cdot x^m)(p^m \cdot x^i)x^j$. Similarly, she is also indifferent between p^i and $(p^m \cdot x^i)p^m$. These indifferences imply that all these prices should be considered when modifying the price of observation i. The following definition formalizes such a notion.

Definition 11. For any H1-sequence $s = (m_\ell)_{\ell \in [L]}$ let

$$I^H(s) = \left\{ \left(\prod_{r=\ell}^{L-1} p^{m_r} \cdot x^{m_{r+1}} \right)^{-1} p^{m_\ell} \right\}_{\ell \in [L]}$$

be the set of prices that are revealed indifferent to p^{m_L}. Furthermore, denote by $\mathcal{S}^H(i)$ the set of all H1-sequences ending in i. The *one-step homothetic modification* of $\mathcal{D}$, $\Gamma^H(\mathcal{D})$, is $\Gamma^H(\mathcal{D}) = \left(\tilde{q}^{i,H}, x^i\right)_{i \in [N]}$, where

$$\tilde{q}^{i,H} = \bigwedge \bigcup_{s \in \mathcal{S}^H(i)} I^H(s).$$

Every H1-sequence s ending in i yields information about prices that are indifferent to p^i; the previous definition collects all such prices in $I^H(s)$. We take all the prices revealed indifferent to p^i by taking the union of all the sets $I^H(s)$ for sequences that end in i. Finally, the one-step homothetic modification of $\mathcal{D}$, $\Gamma^H(\mathcal{D})$ replaces each price by the meet of all prices revealed indifferent to p^i.

In the same spirit as the modified data set for GARP, the Homothetic Modified Data Set is obtained through the fixed point of the one-set homothetic modification.

Definition 12. For a data set $\mathcal{D}$, the *homothetic modified data set*, $D^H_\wedge$, is the fixed point of the one-step homothetic modification Γ^H, using $\mathcal{D}$ as the starting point.

Theorem 3 presents our characterization of rationalization by a homothetic and differentiable utility. Beyond the difference in the rationalization test and the method to infer indifferences, this result presents two differences from the ones in Theorems 1 and 2. The first difference is that we cannot obtain differentiability at the zero bundle in the case of a homothetic utility; this is because any function that is homogeneous of degree one and differentiable at $\mathbf{0}$ is linear, which is a stronger requirement than homotheticity.[32] The second difference regards the Afriat inequalities. In Theorems 1 and 2, we require the same numbers u^i for bundles that are revealed indifferent to

[32] To see that a function that is both homogeneous of degree one and differentiable at $\mathbf{0}$ is linear, take f satisfying both properties. Since for $\lambda > 0$ we have $f(\mathbf{0}) = f(\lambda \mathbf{0}) = \lambda f(\mathbf{0})$ we conclude $f(\mathbf{0}) = 0$. Fix $x > \mathbf{0}$; since f is

each other. Here, homotheticity implies that the same numbers are also required for two bundles if the first one is revealed indifferent (through H1-sequences) to a scaled version of the second.

Theorem 3. *The following are equivalent:*

Ho-1) *$\mathcal{D}$ is rationalizable by a strictly increasing, continuous, concave, homothetic utility function that is differentiable everywhere except at $\mathbf{0}$.*

Ho-2) *$\mathcal{D}^H_\wedge$ satisfies HARP*

Ho-3) *For two bundles x^i, x^j, write $x^i \approx^H_\wedge$ if there is an H1-sequence in $\mathcal{D}^H_\wedge$ containing both i and j, and $x^i \not\approx^H_\wedge x^j$ if there is no such sequence. There are numbers $u^i > 0$ and $\mu^i \geq \mathbf{0}$ such that*

$$u^i\, p^i \cdot x^j > u^j + \mu^i \cdot x^j \qquad \textit{whenever } x^i \not\approx^H x^j \tag{H1}$$

$$u^i\, p^i \cdot x^j = u^j + \mu^i \cdot x^j \qquad \textit{whenever } x^i \approx^H x^j \tag{H2}$$

$$u^i\, p^i - \mu^i = u^j\, p^j - \mu^j \qquad \textit{whenever } x^i \approx^H x^j \tag{H3}$$

$$u^i\, p^i - \mu^i \gg \mathbf{0} \qquad \textit{for all } i \in [N] \tag{H4}$$

$$\mu^i \cdot x^i = 0 \qquad \textit{for all } i \in [N] \tag{H5}$$

Ho-4) *$\mathcal{D}$ is rationalizable by a strictly increasing, continuous, concave, homothetic utility function that is infinitely differentiable everywhere except at $\mathbf{0}$.*

The proof of Theorem 3, presented in Appendix G, proceeds similarly to the previous results. We first show that smooth rationalization by a homothetic utility implies the classical test for homothetic preferences, HARP, applied to $\mathcal{D}^H_\wedge$. Then we construct the numbers in Ho-3). In his original construction of the Afriat inequalities for homothetic utility, Varian (1983a) presents an explicit construction, where each number comes from a minimization process; in our case, such construction is not sufficient as we need to achieve strict inequalities. To construct the numbers in Ho-3), we combine the construction in Varian (1983a) with a modified version of the algorithm proposed by Varian (1982) to construct the original Afriat inequalities for GARP.

To construct the rationalizing utility in Ho-4) requires a further step than the constructions in Theorems 1 and 2. The reason for it is that the convolution technique proposed by Chiappori

---

differentiable at zero

$$0 = \lim_{t \to 0^+} \frac{f(t\,x) - f(\mathbf{0}) - \nabla f(\mathbf{0}) \cdot (t\,x - \mathbf{0})}{||t\,x - \mathbf{0}||} = \lim_{t \to 0^+} \frac{t\,f(x) - t\,\nabla f(\mathbf{0}) \cdot x}{t\,||x||} = \frac{f(x) - \nabla f(\mathbf{0}) \cdot x}{||x||}.$$

Hence $f(x) = \nabla f(\mathbf{0}) \cdot x$. Since x is arbitrary, we conclude that f is linear.

and Rochet (1987) (and used in Theorems 1 and 2) does not preserve homotheticity. For this construction, we start by using the function proposed by Varian (1983a) (using the numbers in Ho-3) instead of the ones he proposed), which is homothetic. Then we project all the choices into one indifference curve on this function and use convolution to smooth that particular indifference curve. Then, we expand that smooth indifference curve to other utility levels, and when doing it, we choose a superset of $\mathbb{R}^K_+$ that assures such extension is well-defined. This expansion yields a (well-behaved) homothetic utility function. Finally, following Debreu (1972) and Neilson (1991), we conclude that this utility is also infinitely differentiable everywhere except at $\mathbf{0}$.

The fact that the main idea of smooth rationalization (to focus on the revealed indifferences to modify the data set, with the modification being taking the meet of prices) applies to more restrictive families of utility functions suggests the generality of this approach. It is area for future research to analyze whether it can be applied to other widely-used utility functions, for example expected utility for risk and uncertainty. However, the fact that all the cases analyzed in this paper follow the same motivation and use the same technical tools suggest that the answer this question should be positive.

6.2 Quasilinear Utility

A utility function is quasilinear if it takes the form $U(x) + y$, where $x \in \mathbb{R}^K_+$ is the commodities space, and y is a numeraire good. The numeraire good is measured in the same units as wealth, hence its price is one. Quasilinear utility functions are used in many areas of economics, including mechanism design, public economics, industrial organization, and international trade. In all such applications, quasilinear utilities are usually also assumed to be differential.

We can think of an agent as maximizing a quasilinear utility if the following terms

Definition 13. $\mathcal{D}$ is *quasilinear rationalizable* by $U(x) + y$ if there is a strictly increasing and concave utility U such that for all $i \in [N]$, x^i solves

$$\begin{aligned} \max_{x \in \mathbb{R}^K} \; & U(x) + y \\ \text{s.t. } \; & p^i \cdot x + y = 0 \\ & x \geq \mathbf{0} \end{aligned} \tag{3}$$

As in many applications, we assume that the numeraire good can be negative; hence the budget constraint is nonbinding in the maximization problem (see Chapter 4.2.3 in Chambers & Echenique, 2016). The previous definition implies that $\mathcal{D}$ is rationalizable by a quasilinear utility U if, and only if, $U(x^i) - 1 \geq U(x) - p^i \cdot x$ for all $i \in [N]$ and $x \in \mathbb{R}^K_+$.

Brown and Calsamiglia (2007) show that the rationalization by a quasilinear utility is equivalent to cyclical monotonicity, a condition to characterize the subgradient correspondence for convex real-valued functions (see Rockafellar, 2015).

Definition 14. $\mathcal{D}$ is *cyclically monotone* if for any sequence of observations $(m_\ell)_{\ell \in [L]}$

$$(p^{m_L} \cdot x^{m_1} - 1) + (p^{m_1} \cdot x^{m_2} - 1) + (p^{m_2} \cdot x^{m_3} - 1) + \ldots + \cdot(p^{m_{L-1}} x^{m_L} - 1) \geq 0 \,. \tag{4}$$

If $\mathcal{D}$ satisfies cyclical monotonicity, the homothetic function $U(x) + y$ can always be chosen such that $U(\cdot)$ is continuous, strictly increasing, and convex. Furthermore, Brown and Calsamiglia (2007) show that this property is equivalent to the existence of numbers $u^i \in \mathbb{R}$ such that $u^i \geq u^j + 1 - p^i \cdot x^j$, which have the intuitive interpretation that a homothetic utility assures that the marginal utility of income, whose analogous in the Afriat inequalities is λ^i, is always equal to one.

As in the case of homothetic utilities and HARP, a quasilinear utility and cyclical monotonicity allow us to infer further indifferences than the ones in the revealed preferences. To see this, take the example $(p^i \cdot x^j - 1) + (p^j \cdot x^m - 1) + (p^m \cdot x^i - 1) = 0$, and suppose such choices are rationalizable by a quasilinear utility U. As x^i is optimal from p^i we have $p^i \cdot x^j - 1 \geq U(x^j) - U(x^i)$. Similarly $p^j \cdot x^m - 1 \geq U(x^m) - U(x^j)$, and $p^m \cdot x^i - 1 \geq U(x^i) - U(x^m)$. From the three inequalities, we obtain

$$0 = (p^i \cdot x^j - 1) + (p^j \cdot x^m - 1) + (p^m \cdot x^i - 1) \geq U(x^j) - U(x^i) + U(x^m) - U(x^j) + U(x^i) - U(x^m) = 0 \,.$$

Since the inequality in the previous equation has to be an equality, we conclude that $p^i \cdot x^j - 1 = U(x^j) - U(x^i)$, $p^j \cdot x^m - 1 = U(x^m) - U(x^j)$, and $p^m \cdot x^i - 1 = U(x^i) - U(x^m)$. Thus, x^j is optimal at prices p^i, x^m is optimal at price p^j, and x^i is optimal at price p^m. Furthermore, if U is differentiable, since both x^i and x^j are optimal from p^i, we can infer that (for interior solutions) the MRS at both bundles is given by the price ratios of p^i; similarly, we can conclude that x^j and x^m have the same MRS and that x^m and x^i also do (all for interior solutions). This argument is generalized in Lemma 14 in Appendix H; furthermore, it shows not only that the MRSs are the same but also that the vector of marginal utilities his the same for the three bundles.

As in the previous sections, we focus on indifferences to obtain differentiability of the utility function. As explained in the previous example, we infer such indifferences through the sequences of cyclical monotonicity that are equal to zero, which we call Q0-sequences.

Definition 15. A *Q0-sequence* is a sequence of unique observations $(m_\ell)_{\ell\in[L]}$ such that

$$(p^{m_L}\cdot x^{m_1}-1)+(p^{m_1}\cdot x^{m_2}-1)+(p^{m_2}\cdot x^{m_3}-1)+\ldots+\cdot(p^{m_{L-1}}x^{m_L}-1)\leq 0$$

As with H1-sequences, including a weak inequality instead of an equality ensures that the data set modification finishes in a finite number of steps. Furthermore, if the previous inequality is strict, we already know that the data set is not rationalizable by a quasilinear utility.

We define our one-step modification similarly to the case of homothetic utilities. The only difference is that prices do not need to be scaled in this case, as the budget constraint is nonbinding.

Definition 16. For any Q0-sequence $s=(m_\ell)_{\ell\in[L]}$ let $I^Q(s)=\{p^{m_\ell}\}_{\ell\in[L]}$. Furthermore, denote by $\mathcal{S}^Q(i)$ the set of all Q0-sequences ending in i. The *one-step quasilinear modification* of $\mathcal{D}$, $\Gamma^Q(\mathcal{D})$, is $\Gamma^Q(\mathcal{D})=\left(\tilde{q}^{i,Q},x^i\right)_{i\in[N]}$, where

$$\tilde{q}^{i,Q}=\bigwedge\bigcup_{s\in\mathcal{S}^Q(i)}I^Q(s)\,.$$

Every Q0-sequence s ending in i yields information about bundles that are indifferent to x^i; the previous definition collects all such bundles in $I^Q(s)$ and then collects all the bundles from the different sequences (by taking the union of the $I^Q(s)$s for different Q0-sequences finishing in i). This process yields all the bundles revealed indifferent to x^i.[33] We take the set of all observations whose choices we can infer are indifferent to x^i. Then, the one-step modification replaces the price p^i by the meet of prices in the such set. The quasilinear modified data set is obtained through the fixed point of Γ^Q.

Definition 17. For a data set $\mathcal{D}$, the *quasilinear modified data set*, $D^Q_\wedge$, is the fixed point of the one-step quasilinear modification Γ^Q, using $\mathcal{D}$ as the starting point.

The following result shows that cyclical monotonicity of the quasilinear modified data set is necessary and sufficient for rationalization by a quasilinear utility $U(x)+y$ that is differentiable. Furthermore, the existence of higher order derivatives of such utility has no empirical content. We also present a modified version of the Afriat inequalities, where the indifferences are obtained using the Q0-sequences of the modified data set.

Theorem 4. *The following are equivalent:*

Qu-1) *$\mathcal{D}$ is quasilinear rationalizable by $U(x)+y$ and U is differentiable.*

[33] Note that the indifferences inferred from Q0-sequences include the ones inferred by the revealed indifferent relation presented in Definition 1.

Qu-2) $\mathcal{D}_{\wedge}^{Q}$ *is cyclically monotone.*

Qu-3) For two bundles x^i, x^j*, write* $x^i \approx_{\wedge}^{Q}$ *if there is an Q0-sequence in* $\mathcal{D}_{\wedge}^{Q}$ *containing both* i *and* j*, and* $x^i \not\approx_{\wedge}^{Q} x^j$ *if there is no such sequence. There are numbers* $u^i > 0$ *and* $\mu^i \geq \mathbf{0}$ *such that*

$$u^i > u^j + 1 - p^i \cdot x^j + \mu^i \cdot x^j \qquad \text{whenever } x^i \not\approx^Q x^j \tag{Q0}$$

$$u^i = u^j + 1 - p^i \cdot x^j + \mu^i \cdot x^j \qquad \text{whenever } x^i \approx^Q x^j \tag{Q2}$$

$$p^i - \mu^i = p^j - \mu^j \qquad \text{whenever } x^i \approx^Q x^j \tag{Q3}$$

$$p^i - \mu^i \gg \mathbf{0} \qquad \text{for all } i \in [N] \tag{Q4}$$

$$\mu^i \cdot x^i = 0 \qquad \text{for all } i \in [N] \tag{Q5}$$

Qu-4) $\mathcal{D}$ *is quasilinear rationalizable by* $U(x) + y$ *and* U *is infinitely differentiable.*

7 Concluding Remarks

Applied models in economics often assume that consumers have a differentiable utility function. This assumption immensely satisfies the analysis, as optimal choices can be characterized through first-order conditions. Such characterization adds tractability to the model and simplifies comparative statics exercises. However, unlike other assumptions like monotonicity or concavity, the behavioral restrictions imposed by assuming a differentiable utility remain elusive. We provide a characterization of assuming a differentiable utility in terms of the requirements it imposes on behavior observed in a finite data set.

The revealed indifferent relation is the crucial component to rationalize a finite data set by a well-behaved (i.e., continuous, monotone, and concave) and differentiable utility. We call such property smooth rationalization. Revealed indifference situations arise if a chosen bundle is revealed preferred to another, and the latter is also revealed preferred to the former. The concavity of the utility function and linear prices imply that the indifference set between two choices that are revealed indifference to each other has to be flat (i.e., linear). Hence, all choices that are revealed indifferent to each other have the same marginal rate of substitution (MRS). This equality of MRSs imposes restrictions on the observed behavior: we need to be able to interpret each choice as optimal not only from the price it was chosen from but also from the meet of all prices of the choices it reveals indifferent to.

We propose a modification of the data set to include the requirements imposed by smooth rationalization on revealed indifferences. Specifically, we replace each price with the meet of prices within the set of revealed indifferences (and repeat the process if needed). We show that smooth rationalization is equivalent to being able to rationalize this modified data set. Hence, the behavioral requirements arising from the revealed indifferences are the only requirements imposed by assuming a differentiable utility. Furthermore, we show that the existence of second and higher order derivatives has no empirical content; whenever a data set can be rationalized by a differentiable utility, it can also be rationalized by an infinitely differentiable utility.

We apply our test of differentiable utility on experimental data sets from several sources. All the subjects choose from two-dimensional budget sets, and the choice is either between risky assets (risk choices) or between own consumption and consumption by a third party (social choices). Our results show that differentiability is a general property, in the sense that there is not a big difference between rationalization and rationalization by a differentiable utility. This result holds as well when we recover preferences of subjects whose choices are not rationalizable (using the Houtman & Maks, 1985, Index). However, we also find that some subjects present a choice pattern that resembles popular utility functions that are not differentiable, specifically Leontief utilities (both for risk and social choices) and safety-first utilities (Roy, 1952) (for risk choices). An additional finding of our analysis is that corner solutions are a common feature of choices. Corner solutions have two implications. First, they imply that the sufficient conditions for a differentiability utility proposed by Chiappori and Rochet (1987) are too strong, as many subjects who fail these conditions satisfy smooth rationalization. Second, it implies that, in order to model behavior through first-order conditions in economic models, it is necessary to include in such conditions the component coming from the nonnegativity restrictions of the chosen goods.

Conceptually, the main ideas of this paper are two. First, the behavioral restriction of assuming differentiability translates only into restrictions over the revealed indifferences. Second, such restrictions can be tested by modifying the data set in a specific way. Our results suggest that such an idea is general, as it can be applied to families of utility functions in which more structure is imposed. Specifically, we show that it applies to three cases: strictly concave, homothetic, and quasilinear utilities. In all cases, the methodology is the same: to modify the data set using revealed indifferences and the meet of prices among such indifferences, and then to apply the rationalization test to the modified data set. However, each case presents its specific characteristics. First, some families of utility functions allow us to infer more indifferences than the ones implied by revealed preferences; such indifferences must be included in our analysis. Second, each family of utility functions has a specific rationalization test, and that specific test must be applied to the modified data set. Despite these differences, all cases are motivated by the same logic and use the same technical tools in their proofs, suggesting that the method developed can be applied to a broad family of economic problems.

References

Afriat, S. N. (1967). The Construction of Utility Functions from Expenditure Data. *International Economic Review*, *8*(1).

Afriat, S. N. (1973). On a System of Inequalities in Demand Analysis: An Extension of the Classical Method. *International Economic Review*, *14*(2).

Afriat, S. N. (1981). On the constructability of consistent price indices between several periods simultaneously. In A. Deaton (Ed.), *Essays in applied demand analysis*. Cambridge University Press.

Aversa, V., Laczkovich, M., & Preiss, D. (1985). Extension of differentiable functions. *Commentationes Mathematicae Universitatis Carolinae*, *026*(3).

Beggs, A. (2021). Afriat and arbitrage. *Economic Theory Bulletin*, *9*(2).

Brown, D. J., & Calsamiglia, C. (2007). The nonparametric approach to applied welfare analysis. *Economic Theory*, *31*(1).

Cappelen, A. W., Kariv, S., Sørensen, E. Ø., & Tungodden, B. (2021). *The Development Gap in Economic Rationality of Future Elites*, Working Paper.

Chambers, C. P., & Echenique, F. (2016). *Revealed Preference Theory*. Cambridge University Press.

Chiappori, P.-A., & Rochet, J.-C. (1987). Revealed Preferences and Differentiable Demand. *Econometrica*, *55*(3).

Choi, S., Fisman, R., Gale, D., & Kariv, S. (2007a). Consistency and Heterogeneity of Individual Behavior under Uncertainty. *American Economic Review*, *97*(5).

Choi, S., Fisman, R., Gale, D., & Kariv, S. (2007b). Revealing Preferences Graphically: An Old Method Gets a New Tool Kit. *American Economic Review: papers and proceeding*, *97*(2).

Choi, S., Kariv, S., Müller, W., & Silverman, D. (2014). Who Is (More) Rational? *American Economic Review*, *104*(6).

Debreu, G. (1972). Smooth Preferences. *Econometrica*, *40*(4).

de Clippel, G., & Rozen, K. (2021). Relaxed Optimization: How Close is a Consumer to Satisfying First-Order Conditions? *Review of Economics and Statistics (forthcoming).*

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory*, *40*(2).

Dembo, A., Kariv, S., Polisson, M., & Quah, J. (2021). *Ever Since Allais*, Working Paper.

Demuynck, T., & Rehbeck, J. (2021). *Computing revealed preference goodness-of-fit measures with integer programming*, Working Paper.

Diewert, W. E. (1973). Afriat and revealed preference theory. *Review of Economic Studies*, *40*(3).

Diewert, W. E., & Parkan, C. (1985). Tests for the consistency of consumer data. *Journal of Econometrics*, *30*(1-2).

Echenique, F., & Saito, K. (2015). Savage in the Market. *Econometrica*, *83*(4).

Fisman, R., Jakiela, P., & Kariv, S. (2015). How did distributional preferences change during the Great Recession? *Journal of Public Economics*, *128*.

Fisman, R., Jakiela, P., & Kariv, S. (2017). Distributional preferences and political behavior. *Journal of Public Economics*, *155*.

Fisman, R., Jakiela, P., Kariv, S., & Markovits, D. (2015). The distributional preferences of an elite. *Science*, *349*(6254).

Fisman, R., Jakiela, P., Kariv, S., & Vannutelli, S. (2022). The Distributional Preferences of Americans, 2013-2016. *Experimental Economics (forthcoming)*.

Fisman, R., Kariv, S., & Markovits, D. (2007). Individual preferences for giving. *American Economic Review*, *97*(5).

Forges, F., & Minelli, E. (2009). Afriat's theorem for general budget sets. *Journal of Economic Theory*, *144*(1).

Fostel, A., Scarf, H. E., & Todd, M. J. (2004). Two new proofs of Afriat's theorem. *Economic Theory*, *24*(1).

Green, R. C., & Srivastava, S. (1986). Expected utility maximization and demand behavior. *Journal of Economic Theory*, *38*(2).

Halevy, Y., Persitz, D., & Zrill, L. (2018). Parametric Recoverability of Preferences. *Journal of Political Economy*, *126*(4).

Harsanyi, J. C. (1955). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy*, *63*(4).

Houthakker, H. S. (1950). Revealed Preference and the Utility Function. *Economica*, *17*(66).

Houtman, M., & Maks, J. (1985). Determining all maximal data subsets consistent with revealed preference. *Kwantitatieve methoden*, *19*.

Jeyakumar, V., Lee, G. M., & Dinh, N. (2004). Lagrange multiplier conditions characterizing the optimal solution sets of cone-constrained convex programs. *Journal of Optimization Theory and Applications*, *123*(1).

Kariv, S., & Silverman, D. (2013). An Old Measure of Decision-Making Quality Sheds New Light on Paternalism. *Journal of Institutional and Theoretical Economics*, *169*(1).

Knoblauch, V. (1993). Recovering homothetic preferences. *Economics Letters*, *43*(1).

Kubler, F., Selden, L., & Wei, X. (2014). Asset demand based tests of expected utility maximization. *American Economic Review*, *104*(11).

Lee, P. M., & Wong, K. C. (2005). Revealed preference and differentiable demand. *Economic Theory*, *25*(4).

Li, J., Casalino, L. P., Fisman, R., Kariv, S., & Markovits, D. (2022). Experimental evidence of physician social preferences. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(28).

Li, J., Dow, W. H., & Kariv, S. (2017). Social preferences of future physicians. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(48).

Matzkin, R. L., & Richter, M. K. (1991). Testing strictly concave rationality. *Journal of Economic Theory*.

Nakamura, Y. (2015). Differentiability of von Neumann-Morgenstern utility functions. *Journal of Mathematical Economics*, *60*.

Neilson, W. S. (1991). Smooth indifference sets. *Journal of Mathematical Economics*, *20*(2).

Nielsen, L. T. (1999). Differentiable von Neumann-Morgenstern utility. *Economic Theory*, *14*(2).

Nishimura, H., Ok, E. A., & Quah, J. (2017). A Comprehensive Approach to Revealed Preference Theory. *American Economic Review*, *107*(4).

Polisson, M., & Renou, L. (2016). Afriat's Theorem and Samuelson's 'Eternal Darkness'. *Journal of Mathematical Economics*, *65*.

Quah, J. (2014). *A test for weakly separable preferences*, University of Oxford, Department of Economics.

Rockafellar, R. (2015). *Convex analysis*.

Roy, A. D. (1952). Safety First and the Holding of Assets. *Econometrica*, *20*(3).

Rubinstein, A. (2012). *Lecture Notes in Microeconomic Theory: the Economic Agent*. Princeton University Press.

Samuelson, P. A. (1938). The Empirical Implications of Utility Analysis. *Econometrica*, *6*(4).

Silberberg, E. (1990). *The Structure of Economics: A Mathematical Analysis.* McGraw-Hill.

Ugarte, C. (2022a). *Preference Recoverability from Inconsistent Choices*, Working Paper.

Ugarte, C. (2022b). *The generality of the Strong Axiom*, Working Paper.

Varian, H. R. (1982). The Nonparametric Approach to Demand Analysis. *Econometrica*, *50*(4).

Varian, H. R. (1983a). Non-Parametric Tests of Consumer Behaviour. *The Review of Economic Studies*, *50*(1).

Varian, H. R. (1983b). Nonparametric Tests of Models of Investor Behavior. *The Journal of Financial and Quantitative Analysis*, *18*(3).

Varian, H. R. (1984). The Nonparametric Approach to Production Analysis. *Econometrica*, *52*(3).

Varian, H. R. (1990). Goodness-of-fit in optimizing models. *Journal of Econometrics*, *46*(1-2).

Zame, W. R., Tungodden, B., Sørensen, E. Ø., Kariv, S., & Cappelen, A. W. (2020). *Linking Social and Personal Preferences: Theory and Experiment*, Working Paper.

Proofs

A Remarks

Proof of Remark 1. For necessity suppose $\mathcal{D}$ satisfies SARP. Take any two choices x^i, x^j such that $x^i \succsim x^j$; if $x^i \neq x^j$ then $x^j \not\succsim^\star x^i$, which implies $x^j \not\succ^\star x^i$; if $x^i = x^j$ then $p^j \cdot x^i = 1$ and $x^j \not\succ^\star x^i$, hence $\mathcal{D}$ satisfies GARP. By contrapositive take $x^i \neq x^j$ such that $x^i \sim x^j$. Then there are sequences $(m_\ell)_{\ell=1}^L$ and $(n_s)_{s=1}^S$ such that

$$x^i \succsim^\star x^{m_1} \succsim^\star \ldots \succsim^\star x^{m_L} \succsim^\star x^j \succsim^\star x^{n_1} \succsim^\star \ldots \succsim^\star x^{n_S} \succsim^\star x^i \,.$$

Hence $x^i \succsim x^{n_S}$ and $x^{n_S} \succsim^\star x^i$, a violation of SARP. We conclude that $x^i \neq x^j$ implies $x^i \not\sim x^j$.

For sufficiency suppose $\mathcal{D}$ satisfies GARP and $x^i \not\sim x^j$ whenever $x^i \neq x^j$. Towards a contradiction suppose $\mathcal{D}$ fails SARP, i.e., there are x^m, x^n such that $x^m \neq x^n$, $x^m \succsim x^n$, and $x^n \succsim^\star x^m$. As $x^m \succsim^\star x^n$ implies $x^n \succsim x^m$ we have $x^m \sim x^n$, a contradiction. □

Proof of Remark 2. Take $x^i = x^j$. Since $p^i \cdot x^i = 1$ for all $i \in [N]$, (S3) and (S5) imply

$$\lambda^i = \lambda^i p^i \cdot x^i - \mu^i \cdot x^i = (\lambda^i p^i - \mu^i) \cdot x^i = (\lambda^j p^j - \mu^j) \cdot x^j = \lambda^j \,.$$

□

Proof of Remark 3. Suppose $\mathcal{D}$ is smoothly rationalized by U, which is strictly concave. Then $\mathcal{D}$ satisfies SARP, which implies that $\{j \in [N] : x^i \sim x^j\} = \{j \in [N] : x^i = x^j\}$ for every $i \in [N]$. Hence $\Gamma(\mathcal{D}) = \mathcal{D}^S_\wedge$. Furthermore, as U rationalizes $\mathcal{D}^S_\wedge$ then $\mathcal{D}^S_\wedge$ satisfies SARP, which implies that no different choices reveal indifferent to each other according to the prices in $\mathcal{D}^S_\wedge$ (Remark 1). As same choices have the same (modified) price in $\mathcal{D}^S_\wedge$, we have $\Gamma(\mathcal{D}^S_\wedge) = \mathcal{D}^S_\wedge$, hence $\Gamma(\Gamma(\mathcal{D})) = \Gamma(\mathcal{D})$. We conclude that $\mathcal{D}_\wedge = \Gamma(\Gamma(\mathcal{D})) = \Gamma(\mathcal{D}) = \mathcal{D}^S_\wedge$. □

Proof of Remark 4. Suppose $K = 2$. Since every choice x^i satisfies $x^i > 0$ (as $p^i \cdot x^i = 1$), then x^i has at most one dimension equal to zero. Hence by (D1) for any $j \in E(i)$, p^i and p^j differ in one dimension. This implies that there is $j \in E(i)$ such that $q^i = p^j$. As $x^i = x^j$ and $\mathcal{D}$ satisfies SARP, $\mathcal{D}^S_\wedge$ has to satisfy SARP as well. □

B A short review of smoothness by convolution

Before presenting the proofs of our main results, we review the method to smooth a function by convolution, which we use extensively through the paper. To apply this method in the revealed preference literature was proposed by Chiappori and Rochet (1987), and we follow their exposition. here we only state the results, see their paper and further references for proofs.

Suppose we have a function $V : \mathbb{R}^K \to \mathbb{R}$ which has desirable properties but is not differentiable; in particular, it is strictly increasing and concave. Chiappori and Rochet (1987) propose to smooth such function taking a small number $\eta > 0$ and defining the following functions (where $B(\eta)$ is the ball centered at $\mathbf{0}$ or radius η)

$$\rho(x) = \begin{cases} \left[\int_{\mathbb{R}^K} \exp\left(-\frac{1}{||y||^2-1}\right) dy\right]^{-1} \exp\left(-\frac{1}{||x||^2-1}\right) & \text{if } ||x|| < 1 \\ 0 & \text{otherwise.} \end{cases}$$

$$\rho_\eta(x) = \frac{1}{\eta}\rho\left(\frac{x}{\eta}\right) \tag{5}$$

$$\begin{aligned} \tilde{U}(x) &= (V \star \rho_\eta)(x) \\ &= \int_{\mathbb{R}^K} V(x-\xi)\rho_\eta(\xi)\, d\xi \\ &= \int_{B(\eta)} V(x-\xi)\rho_\eta(\xi)\, d\xi \end{aligned} \tag{6}$$

The function $\rho_\eta(x)$ is infinitely differentiable,[34] symmetric ($\rho_\eta(x) = \rho_\eta(-x)$), weakly positive ($\rho_\eta(x) \geq 0$), and strictly positive whenever $||x|| < \eta$. Then $\tilde{U}$ is also infinitely differentiable.

[34] In particular, all the partial derivatives of any order of $\rho_\eta(x)$ equal zero whenever $||x|| = \eta$.

Furthermore, if $\tilde{U}$ inherits the following properties from V: strictly increasing, concavity, and strict concavity; this is, if V satisfies any of such properties then $\tilde{U}$ also does.

Two useful properties of ρ_η are

$$\int_{B(\eta)} \rho_\eta(\xi)\, d\xi = 1; \text{ and} \tag{7}$$

$$\int_{B(\eta)} \xi \rho_\eta(\xi)\, d\xi = \mathbf{0}\,. \tag{8}$$

C Theorem 1

This section presents the proof of Theorem 1. The direction of the proof is Sa-1) $\implies$ Sa-2) $\implies$ Sa-3) $\implies$ Sa-4) $\implies$ Sa-1). The proof that Sa-4) implies Sa-1) is immediate, so we omit it.

C.1 Sa-1) implies Sa-2)

Proof that Sa-1) implies Sa-2). (Although this proof could be made shorter by using Theorem 5, as in the proof of Theorem 2, we instead present an explicit proof that does not refer to the aforementioned result). Suppose $\mathcal{D}$ is smoothly rationalizable by a utility function $U : \mathbb{R}^K_+ \to \mathbb{R}$. Since U is strictly increasing, for any observation (p^i, x^i) there are Lagrange multipliers $\lambda^i > 0$ and $\mu^i \in \mathbb{R}^K_+$ such that for every $k \in [K]$

$$U_k(x^i) = \lambda^i\, p^i_k - \mu^i_k \tag{9}$$

$$\mu^i_k\, x^i_k = 0 \tag{10}$$

Furthermore, these conditions imply

$$\nabla U(x^i) \cdot x^i = \lambda^i\, p^i \cdot x^i - \mu^i \cdot x^i = \lambda^i \qquad \text{for all } i \in [N]. \tag{11}$$

If $x^i = x^j$ and $x^i_k > 0$, then (10) implies $\mu^i_k = \mu^j_k = 0$, which by (9) implies $\lambda^i\, p^i_k = \lambda^j\, p^j_k$. Since $\lambda^i > 0$,

$$1 = p^i \cdot x^i = p^i \cdot x^j = \sum_{\{k : x^j_k > 0\}} p^i_k\, x^j_k = \sum_{\{k : x^j_k > 0\}} \frac{\lambda^i\, p^i_k}{\lambda^i}\, x^j_k = \sum_{\{k : x^j_k > 0\}} \frac{\lambda^j\, p^j_k}{\lambda^i} x^j_k = \frac{\lambda^j}{\lambda^i} p^j \cdot x^j = \frac{\lambda^j}{\lambda^i}\,.$$

Hence $\lambda^i = \lambda^j$. Thus whenever $x^i = x^j$ and $x^i_k > 0$ we have $p^i_k = p^j_k$ (as $\lambda^i\, p^i_k = \lambda^j\, p^j_k$), i.e., (D1) holds.

From the previous paragraph and Remark 1 we have $\lambda^i = \lambda^j$ whenever $x^i \sim x^j$. Since $\mu_k^j \geq 0$, (9) implies

$$U_k(x^i) = \lambda^i p_k^j - \mu_k^j \leq \lambda^i p_k^j \qquad \text{for every } k \in [K], \text{ whenever } x^i = x^j\,.$$

Hence $U_k(x^i) \leq \lambda^i \min_{\{j:x^j=x^i\}} p_k^j$ for all k. Therefore, as U is strictly increasing

$$\mathbf{0} \ll \nabla U(x^i) \leq \lambda^i q^i\,. \tag{12}$$

As U is strictly concave and differentiable, for every $x \neq x^i$

$$\begin{aligned} U(x^i) - U(x) &> \nabla U(x^i) \cdot (x^i - x) \\ &= \nabla U(x^i) \cdot x^i - \nabla U(x^i) \cdot x \\ &= \lambda^i - \nabla U(x^i) \cdot x \\ &\geq \lambda^i (1 - q^i \cdot x) \end{aligned} \tag{13}$$

The second line distributes the product; the third one replaces (11); and the fourth one follows from (12) and $x \geq \mathbf{0}$. Equation (13) implies that the vector $\lambda^i q^i$ is a supergradient of U at x^i, although not necessarily the gradient. Even though U is differentiable, this difference can be made because x^i might be in the boundary of $\mathbb{R}_+^K$, i.e., it might have some components equal to zero. If $x^i \gg \mathbf{0}$ then $\lambda^i q^i = \lambda^i p^i$ is the gradient of U at x^i.

As $\lambda^i > 0$, (13) implies that for every $x \neq x^i$

$$q^i \cdot x \leq 1 \implies U(x^i) > U(x) \tag{14}$$

Hence $\mathcal{D}_\wedge^S$ is rationalized by U. As U is strictly concave, $\mathcal{D}_\wedge^S$ satisfies SARP. As $\mathcal{D}$ satisfies SARP, (D1) and (D2), it satisfies Sa-2). □

C.2 Sa-2) implies Sa-3)

The proof that Sa-2) implies Sa-3) uses three auxiliary results: Lemma 1 shows that SARP requires for a bundle not to be revealed preferred to any other bundle; Lemma 2 shows the existence of the classical Afriat inequalities used in the case of SARP (see Theorem 2 in Matzkin & Richter, 1991), but adds the restriction that the same bundle has always the same marginal utility of income; Lemma 3 shows that (D1) implies that there is no income expenditure.

Lemma 1. *If $\mathcal{D}$ satisfies SARP, then there is $i \in [N]$ such that $p^j \cdot x^m > 1$ whenever $x^j = x^i$ and $x^m \neq x^i$.*

Proof. By contrapositive suppose for every $i \in [N]$ there are j, m such that $x^j = x^i$, $x^m \neq x^i$, and $p^j \cdot x^i \leq 1$. We have $x^i \succsim^\star x^j \succsim^\star x^m$, hence $x^i \sim x^m$. Thus we can construct an infinite sequence $(x^{n_\ell})_{\ell \in \mathbb{R}}$ such that $x^{n_\ell} \neq x^{n_{\ell+1}}$ and $x^{n_\ell} \succsim x^{n_{\ell+1}}$. As the data set is finite ($N < \infty$) there is a choice that has to repeat in the sequence, i.e., there are ℓ', ℓ'' with $\ell'' > \ell' + 1$ such that $x^{m_{\ell'}} = x^{m_{\ell''}}$. Then we have $x^{m_{\ell'}} \succsim x^{m_{\ell'+1}}$ and $x^{m_{\ell'+1}} \succsim x^{m_{\ell''}} = x^{m_{\ell'}}$, therefore $x^{m_{\ell'}} \sim x^{m_{\ell'+1}}$. Since $x^{m_{\ell'}} \neq x^{m_{\ell'+1}}$, by Remark 1 we conclude that $\mathcal{D}$ satisfies SARP. □

Lemma 2. *If ((D1)) and ((D2)) hold then there are numbers $u^i \in \mathbb{R}$ and $\lambda^i > 0$ such that*

$$u^i > u^j + \lambda^i(1 - q^i \cdot x^j) \qquad \text{whenever } x^i \neq x^j$$

$$u^i = u^j \qquad \text{whenever } x^i = x^j$$

$$\lambda^i = \lambda^j \qquad \text{whenever } x^i = x^j$$

Proof. We proceed by induction on the number of observations. If $N = 1$, set $u^1 = \lambda^1 = 1$.

Suppose GARP holds for all databases comprised of $N - 1$ or less observations, and take $\mathcal{D}$ comprised of N observations. By Lemma 1, and without loss of generality, suppose N is such that $q^i \cdot x^j > 1$ whenever $x^i = x^N$ and $x^j \neq x^N$.

- If $x^i = x^N$ for all i, set $u^i = \lambda^i = 1$ for every $i \in [N]$. Then Lemma 3 implies $q^i \cdot x^j = 1$ for all i, j and the conditions hold.
- If there is j such that $x^j \neq x^N$, then the data set $(q^j, x^j)_{\{j : x^j \neq x^N\}}$ is a data set of $N - 1$ or less observations for which the conditions hold. Take $\varepsilon > 0$, and for every i such that $x^i \sim x^N$ set

  $$u^i = \min_{\{m : x^m = x^N\}} \min_{\{j : x^j = x^N\}} u^j - \lambda^j(1 - q^j \cdot x^m) - \varepsilon$$

  Then $u^i = u^j$ whenever $x^i = x^j$. Moreover, whenever $x^j \neq x^N$ and $x^i = x^N$

  $$u^i \leq u^j - \lambda^j(1 - p^j \cdot x^i) - \varepsilon < u^j - \lambda^j(1 - p^j \cdot x^i)$$

  Whenever $x^i = x^N$ set

  $$\lambda^i = \max\left\{ \max_{\{m : x^m = x^N\}} \max_{\{j : x^j \neq x^N\}} \frac{u^j - u^m}{p^m \cdot x^j - 1} + \varepsilon;\ 1 \right\}.$$

  Hence $\lambda^i = \lambda^N > 0$ whenever $x^i = x^N$. Finally, if $x^i = x^N$ and $x^j \neq x^N$ then

  $$u^j + \lambda^i(1 - p^i \cdot x^j) = u^j - \lambda^i(p^i \cdot x^j - 1)$$

$$
\begin{aligned}
&\leq u^j - \left(\frac{u^j - u^i}{p^i \cdot x^j - 1} + \varepsilon\right)(p^i \cdot x^j - 1) \\
&< u^j - \left(\frac{u^j - u^i}{p^i \cdot x^j - 1}\right)(p^i \cdot x^j - 1) \\
&= u^i
\end{aligned}
$$

The first inequality follows from the definition of λ, and the second from $p^i \cdot x^j > 1$. We conclude that the conditions holds for all $i, j \in [N]$.

$\square$

Lemma 3. *If $\mathcal{D}$ satisfies (D1) then $q^i \cdot x^i = 1$ for all $i \in [N]$.*

Proof. (D1) implies $q^i_k = p^i_k$ whenever $x^i_k > 0$. Hence

$$
q^i \cdot x^i = \sum_{\{k : x^i_k > 0\}} q^i_k \, x^i_k = \sum_{\{k : x^i_k > 0\}} p^i_k \, x^i_k = p^i \cdot x^i = 1 \,.
$$

$\square$

Proof that Sa-1) implies Sa-3). By Lemma 6 there are numbers $u^i \in \mathbb{R}$ and $\lambda^i > 0$ such that

$$
\begin{aligned}
&u^i > u^j + \lambda^i(1 - q^i \cdot x^j) && \text{whenever } x^i \neq x^j && (15) \\
&u^i = u^j + \lambda^i(1 - q^i \cdot x^j) && \text{whenever } x^i = x^j && (16) \\
&\lambda^i = \lambda^j && \text{whenever } x^i = x^j && (17)
\end{aligned}
$$

For every $i \in [N]$, set $\mu^i = \lambda^i(p^i - q^i)$. As $p^i \geq q^i$ and $\lambda^i > 0$ we have $\mu \geq \mathbf{0}$. Furthermore, replacing $\lambda \, q^i = \lambda \, p^i - \mu^i$ in (15) we conclude that (S1) holds, and (S2) holds by (16) and Lemma 3. As $q^i = q^j$ and $\lambda^i = \lambda^j$ whenever $x^i = x^j$,

$$
\lambda^i \, p^i - \mu^i = \lambda^i \, q^i = \lambda^j \, q^j = \lambda^j \, p^j - \mu^j \,.
$$

Therefore (S3). Similarly, as $q^i \gg \mathbf{0}$ then (17) implies $\lambda^i \, p^i - \mu^i = \lambda^i \, q^i \gg \mathbf{0}$ and (S4) holds. Finally, since (D1) holds Lemma 3 implies

$$
\mu^i \cdot x^i = \lambda^i(p^i - q^i) \cdot x^i = 0 \,,
$$

hence (S5) holds. This completes the proof. $\square$

C.3 Sa-3) implies Sa-4)

Proof that Sa-3) implies Sa-4). Set $M > 0$ and define the function $g : \mathbb{R}^K_+ \to \mathbb{R}_+$ by

$$\begin{aligned} g(x) &= \left(M + ||x||^2\right)^{1/2} - M^{1/2} \\ &= \left(M + \sum_{k\in[K]} (x_k)^2\right)^{1/2} - M^{1/2}. \end{aligned}$$

Then g is continuous, strictly convex, $g(\mathbf{0}) = 0$, and $g(x) > 0$ for $x \neq \mathbf{0}$ (see Matzkin & Richter, 1991). As $N < \infty$ and $\lambda^i p^i_k > 0$ for all $i \in [N]$ and $k \in [K]$, it follows from (S1) that there is $\varepsilon > 0$ small enough such that

$$u^i - \varepsilon g(x^i - x^j) > u^j + \lambda^i(1 - p^i x^j) + \mu^i x^j \qquad \text{whenever } x^i \neq x^j\text{; and} \tag{18}$$

$$\lambda^i p^i_k - \mu^i_k > \varepsilon \qquad \text{for all } i \in [N], k \in [K]. \tag{19}$$

For each $i \in [N]$ define the function $\phi_i : \mathbb{R}^K \to \mathbb{R}$ by

$$\phi^i(x) = u^i - \lambda^i(1 - p^i \cdot x) - \mu^i \cdot x - \varepsilon g(x - x^i),$$

which is continuous, strictly concave, and strictly increasing.[35] Furthermore, whenever $x^i = x^j$ it follows from (S2), Remark 2, and (S3), that $\phi^i(x) = \phi^j(x)$ for all x.

Define $V : \mathbb{R}^K \to \mathbb{R}$ by

$$V(x) = \min_{i\in[N]} \phi^i(x).$$

Since $V(\cdot)$ is the minimum of finitely many functions and all of them are continuous, strictly increasing, and strictly convex, it inherits these three properties. Furthermore, from (18) it follows that $\phi(x^i) < \phi^j(x^i)$ whenever $x^i \neq x^j$. Together with the fact that $\phi^i = \phi^j$ whenever $x^i = x^j$, this implies that $V(x^i) = u^i$ and $V(x) = \phi^i(x)$ in a neighborhood of x^i for every $i \in [N]$.

Since $V(x) = \phi^i(x)$ in a neighborhood of x^i, there is $\eta > 0$ small enough such that for all $i \in [N]$ and $\xi \in B(\eta)$ we have $V(x^i - \xi) = \phi^i(x^i - \xi)$. Define $\tilde{U} : \mathbb{R}^K \to \mathbb{R}$ following (6). Then $\tilde{U}$ is continuous, infinitely differentiable, strictly concave and strictly increasing (see Appendix B). Furthermore for all $i \in [N]$

$$\tilde{U}(x^i) = \int_{B(\eta)} V(x^i - \xi)\rho_\eta(\xi)\, d\xi$$

[35] $\phi^i(\cdot)$ is strictly increasing as for all $k \in [K]$ we have

$$\frac{\partial \phi^i(x)}{\partial x_k} = \lambda^i p^i_k - \mu^i_k - \varepsilon \frac{x_k}{\left(M + \sum_{\ell\in[K]}(x_\ell)^2\right)^{1/2}} = \lambda^i p^i_k - \mu^i_k - \varepsilon \left(\frac{(x_k)^2}{M + \sum_{\ell\in[K]}(x_\ell)^2}\right)^{1/2} > \lambda^i p^i_k - \mu^i_k - \varepsilon > 0.$$

The last inequality follows from (19).

$$
\begin{aligned}
&= \int_{B(\eta)} \left[\min_{j\in[N]} \phi^j(x^i-\xi)\right] \rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} \phi^i(x^i-\xi)\rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} \left[u^i - \lambda^i(1-p^i\cdot(x^i-\xi)) - \mu^i\cdot(x^i-\xi) - \varepsilon g(-\xi)\right] \rho_\eta(\xi)\, d\xi \\
&= \left[u^i - \lambda^i(1-p^i\cdot x^i) - \mu^i\cdot x^i\right] \int_{B(\eta)} \rho_\eta(\xi)\, d\xi - (\lambda^i p^i - \mu^i)\cdot \int_{B(\eta)} \xi\rho_\eta(\xi)\, d\xi - \varepsilon \int_{B(\eta)} g(-\xi)\rho_\eta(\xi)\, d\xi \\
&= u^i - \varepsilon \int_{B(\eta)} g(\xi)\rho_\eta(\xi)\, d\xi \qquad (20)
\end{aligned}
$$

The last equality follows from (7), (8), $g(x) = g(-x)$, $p^i \cdot x^i = 1$, and (S5).

Take $x \neq x^i$ satisfying $p^i \cdot x \leq 1$. Then

$$
\begin{aligned}
U(x) &= \int_{B(\eta)} V(x-\xi)\rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} \left[\min_{j\in[N]} \phi^j(x-\xi)\right] \rho_\eta(\xi)\, d\xi \\
&\leq \int_{B(\eta)} \left[\phi^i(x-\xi)\right] \rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} \left[u^i - \lambda^i(1-p^i\cdot(x-\xi)) - \mu^i\cdot(x-\xi) - \varepsilon g(x-\xi-x^i)\right] \rho_\eta(\xi)\, d\xi \\
&= \left[u^i - \lambda^i(1-p^i\cdot x) - \mu^i\cdot x\right] \int_{B(\eta)} \rho_\eta(\xi)\, d\xi - (\lambda^i p^i - \mu^i)\cdot \int_{B(\eta)} \xi\rho_\eta(\xi)\, d\xi - \\
&\qquad\qquad - \varepsilon \int_{B(\eta)} g(x-\xi-x^i)\rho_\eta(\xi)\, d\xi \\
&= u^i - \lambda^i(1-p^i\cdot x) - \mu^i\cdot x - \varepsilon \int_{B(\eta)} g((x-x^i)-\xi)\rho_\eta(\xi)\, d\xi \\
&\leq u^i - \varepsilon \int_{B(\eta)} g((x-x^i)-\xi)\rho_\eta(\xi)\, d\xi \\
&= u^i - \varepsilon \int_{B(\eta)} g(\xi+(x^i-x))\rho_\eta(\xi)\, d\xi \\
&< u^i - \varepsilon \int_{B(\eta)} \left[g(\xi) + \nabla g(\xi)\cdot(x-x^i)\right] \rho_\eta(\xi)\, d\xi \\
&= u^i - \varepsilon \int_{B(\eta)} g(\xi)\rho_\eta(\xi)\, d\xi - \varepsilon \int_{B(\eta)} \nabla g(\xi)\cdot(x-x^i)\, \rho_\eta(\xi)\, d\xi \\
&= U(x^i) - \varepsilon \int_{B(\eta)} \nabla g(\xi)\cdot(x-x^i)\, \rho_\eta(\xi)\, d\xi \\
&= U(x^i) - \varepsilon(x-x^i)\cdot \int_{B(\eta)} \xi \frac{\rho_\eta(\xi)}{(M+||\xi||^2)^{1/2}}\, d\xi \\
&= U(x^i)
\end{aligned}
$$

The second line follows from the definition of U; the third one from $i \in [N]$; the fourth one from the definition of ϕ; the fifth one rearranges terms; the sixth one from (7) and (8); the seventh one from $p^i \cdot x \leq 1$, $\lambda^i > 0$, $\mu^i \geq \mathbf{0}$, and $x > \mathbf{0}$; the eight one from symmetry of g; the ninth one from the strict convexity of g, where ∇g is the gradient of g;[36] the tenth one rearranges terms; the eleventh one from (20); the twelveth one by replacing ∇g; and the last one from the fact that $\frac{\rho_\eta(\xi)}{(M+||\xi||^2)^{1/2}}$ is symmetric around zero, thus the integral is equal to zero.

Define U being the restriction of $\tilde{U}$ to $\mathbb{R}^K_+$. Then U is strictly increasing, strictly concave, infinitely differentiable, and rationalizes $\mathcal{D}$. □

D Proposition 1

Proof of Proposition 1. Sufficiency is given by Lemma 3. For necessity suppose $\mathcal{D}^S_\wedge$ satisfies SARP and $q^i \cdot x^i = 1$ for all i. Since $D^S_\wedge$ satisfies SARP, so does $\mathcal{D}$. Towards a contradiction suppose there are $i, j \in [N]$ and $k \in [K]$ such that $x^i = x^j$, $x^i_k > 0$ and $p^i_k \neq p^j_k$. Without loss of generality suppose $p^i_k > p^j_k$, which implies $p^i_k > q^i_k$. Then

$$1 = p^i \cdot x^i = p^i_k\, x^i_k + \sum_{k' \neq k} p^i_{k'}\, x^i_{k'} < q^i_k\, x^i_k + \sum_{k' \neq k} p^i_{k'}\, x^i_{k'} \leq q^i_k\, x^i_k + \sum_{k' \neq k} q^i_{k'}\, x^i_{k'} = q^i \cdot x^i\,,$$

a contradiction. □

E Proposition 2

To prove Proposition 2 we use the following result from optimization theory, due to Jeyakumar et al. (2004), which characterize the set of maximizers for a convex programming problem.

Theorem 5 (Jeyakumar et al. (2004), Theorem 2.2)**.** *Consider the convex programming problem*

$$\begin{aligned} &\min\ f(x) \\ &s.t.\ \ x \in C \\ &\quad -g(x) \in Z \end{aligned} \tag{P}$$

where X is a Banach space, Y is a locally convex (Hausdorff) space, C is a closed convex subset of X, Z is a closed convex cone in Y, $f : X \to \mathbb{R}$ is a continuous convex function, and $g : X \to Y$ is a continuous Z-convex mapping. Let $A = \{x \in C : g(x) \in -Z\}$ be the feasible set and $S = \{x \in$

[36] Strict convexity of g implies $g(x) + \nabla g(x) \cdot (x - y) > g(y)$ whenever $x \neq y$ (with equality if $x = y$). The inequality follows replacing $x = \xi$ and $y = \xi + (x^i - x)$.

$A : f(x) \leq f(y) \forall y \in A\}$ be the solution set. Let $a \in S$ and λ^a be a Lagrange multiplier associated with a. Then

$$S = \{x \in A : \lambda^a g(x) = 0,\, \exists u_a \in \partial f(x) \cap \partial f(a),\, u_a \cdot (x - a) = 0\}\,.$$

In particular, if f is differentiable on X, then

$$S = \{x \in A : \lambda^a g(x) = 0,\, \nabla f(x) = \nabla f(a),\, \nabla f(a) \cdot (x - a) = 0\}\,.$$

We can convert the problem (P) to the utility maximization problem (UM) by defining $f = -U$ and maximizing U instead of minimizing $-U$ (as U is concave, $-U$ is convex). The potential consumption space is $X = C = \mathbb{R}^K_+$. The budget constraint and the K nonnegativity restrictions (which need to be added for sufficiency of the first order conditions) total $K+1$ restrictions, hence $Y = \mathbb{R}^{K+1}$. The function $g : \mathbb{R}^K \to \mathbb{R}^{K+1}$, where

$$g_k(x) = \begin{cases} x_k & \text{if } k \leq K \\ 1 - p \cdot x & \text{if } k = K+1\,, \end{cases}$$

and $Z = \mathbb{R}^{K+1}_+$ (which is a closed convex cone in $\mathbb{R}^{K+1}$) define the restrictions of the problem.

We split the proof of Proposition 2 in two steps. In the first step we bound the marginal marginal utility in each choice by using the one-step modification Γ. In the second step we use the obtained bound to prove our desired result.

Lemma 4. *Suppose $\mathcal{D}$ is smoothly rationalized by U, and let λ^i be the marginal utility of money (i.e., the Lagrange multiplier of the budget set) when choosing x^i from price p^i. Denote by $\tilde{q}$ the prices in $\Gamma(\mathcal{D})$; this is, $\Gamma(\mathcal{D}) = (\tilde{q}^i, x^i)$. Then*

$$\nabla U(x^i) \leq \lambda^i\, \tilde{q}^i\,.$$

Proof. Suppose $\mathcal{D}$ is rationalizable by a strictly increasing, concave, and differentiable utility $U : \mathbb{R}^K_+ \to \mathbb{R}$. Since $\mathcal{D}$ is rationalizable, by Afriat's theorem it satisfies GARP. Also, as U is strictly increasing, for any observation (p^i, x^i) there are Lagrange multipliers $\lambda^i > 0$ and $\mu^i \in \mathbb{R}^K_+$ such that $\nabla U(x^i) = \lambda^i p^i - \mu^i$, $\mu^i \cdot x^i = 0$, and (as a consequence of the previous two equations), $\nabla U(x^i) \cdot x^i = \lambda^i$.

Take $i, j \in [N]$ such that $x^i \sim x^j$. As $x^i \succsim x^j$ and $x^j \succsim x^i$ there are sequences of observations $(m_\ell)_{\ell \in [L]}$ and $(n_s)_{s \in [S]}$ such that

$$x^i \succsim^\star x^{m_1} \succsim^\star \ldots \succsim^\star x^{m_L} \succsim^\star x^j \succsim^\star x^{n_1} \succsim^\star \ldots \succsim^\star x^{n_S} \succsim^\star x^i\,. \tag{21}$$

Revealed preferences imply $U(x^i) = U(x^j) = U(x^{m_\ell}) = U(x^n_s)$ for all $\ell \in [L]$, $s \in [S]$.

As GARP holds, $x^i \succsim^\star x^{m_1}$ and $x^{m_1} \succsim x^i$ imply $p^i \cdot x^{m_1} = 1$, i.e., x^{m_1} is affordable from price p^i. As $U(x^i) = U(x^{m_1})$, Theorem 5 implies $\nabla U(x^i) = \nabla U(x^{m_1})$. The same argument applied to (21) implies

$$\nabla U(x^i) = \nabla U(x^{m_1}) = \nabla U(x^{m_2}) = \ldots = \nabla U(x^{m_L}) = \nabla U(x^j)\,.$$

From $\nabla U(x^i) = \nabla U(x^{m_1})$ we have

$$\lambda^i\, p^i + \mu^i = \nabla U(x^i) = \nabla U(x^{m_1}) = \lambda^{m_1}\, p^{m_1} - \mu^{m_1}\,. \tag{22}$$

Moreover, the characterization of the set of optimal solutions in Theorem 5 implies $\mu^{m_1} \cdot x^i = 0$. Taking dot product of (E) and x^i we obtain $\lambda^i = \lambda^{m_1}$. By the same argument we have

$$\lambda^i = \lambda^{m_1} = \lambda^{m_2} = \ldots = \lambda^{m_L} = \lambda^j\,.$$

Combining (E) and (E) we obtain

$$\nabla U(x^i) = \nabla U(x^j) = \lambda^j\, p^j - \mu^j \leq \lambda^j\, p^j = \lambda^i\, p^j \qquad\qquad \text{whenever } x^i \sim x^j\,.$$

As j is arbitrary and U is strictly increasing, the definition of $\tilde{q}^i$ implies

$$\mathbf{0} \ll \nabla U(x^i) \leq \lambda^i \tilde{q}^i\,.$$

□

Proof of Proposition 2. Suppose $\mathcal{D}$ is smoothly rationalizable by U. By Lemma 4 we have $\nabla U(x^i) \leq \lambda^i\, \tilde{q}^i$ for all $i \in [N]$. As U is concave

$$\begin{aligned} U(x^i) - U(x) &\geq \nabla U(x^i) \cdot (x^i - x) \\ &= \nabla U(x^i) \cdot x^i - \nabla U(x^i) \cdot x \\ &= \lambda^i - \nabla U(x^i) \cdot x \\ &\geq \lambda^i (1 - \tilde{q}^i \cdot x) \end{aligned} \tag{23}$$

The second line distributes the product, the third one replaces (11), and the fourth one follows from Lemma 4 and $x \geq \mathbf{0}$. As $\lambda^i > 0$, equation 23 implies that $U(x^i) \geq U(x)$ whenever $\tilde{q} \cdot x \leq 1$. Therefore U rationalizes $\Gamma(\mathcal{D})$.

Finally, since $\Gamma(\mathcal{D})$ is rationalized by U, repeating the previous argument we obtain that $\Gamma(\Gamma(\mathcal{D}))$ also is. Furthermore, the same holds for any number of iterations of Γ to $\mathcal{D}$. Since $\mathcal{D}_\wedge$ is obtained by iterating Γ on $\mathcal{D}$ a finite number of times, $\mathcal{D}_\wedge$ rationalized by U.

□

Proof of Corollary 1. Since $\mathcal{D}$ is smoothly rationalizable by U, $\mathcal{D}_\wedge$ also is (Proposition 2). Hence each choice x^i is the optimal choice from the price q^i. It follows from the first-order conditions of the maximization problem that there are $\lambda^i > 0$ and $\mu^i \in \mathbb{R}^K_+$ such that

$$\nabla U(x^i) = \lambda^i q^i - \mu^i \leq \lambda^i q^i \,.$$

Furthermore, the complementary slackness condition implies $\mu^i_k = 0$, therefore $U_k(x^i) = \lambda^i q^i_k$, whenever $x^i_k > 0$. □

F Theorem 2

The direction of the proof is Ga-1) $\Longrightarrow$ Ga-2) $\Longrightarrow$ Ga-3) $\Longrightarrow$ Ga-4) $\Longrightarrow$ Ga-1). The proof that Ga-4) implies Ga-1) is immediate, so we omit it.

F.1 Ga-1) implies Ga-2)

Proof that Ga-1) implies Ga-2). If $\mathcal{D}$ is smoothly rationalizable, then $\mathcal{D}_\wedge$ also is (Proposition 2). By Afriat's theorem, $\mathcal{D}_\wedge$ satisfies GARP. □

F.2 Ga-2) implies Ga-3)

Lemma 5. *There is $i \in [N]$ such that $p^j \cdot x^m > 1$ whenever $x^j \sim x^i$ and $x^m \not\sim x^i$.*

Proof. By contrapositive, suppose for each $i \in [N]$ there are $x^j \sim x^i$ and $x^m \not\sim x^i$ such that $p^j \cdot x^m \leq 1$. As $x^m \not\sim x^i$, then $m \neq i$. Moreover, transitivity of $\succsim$ and $x^j \succsim x^m$ imply $x^i \succsim x^m$. Hence, we can construct an infinite sequence $(n_\ell)_{\ell=1}^\infty$ such that $x^{n_\ell} \succsim x^{n_{\ell+1}}$ and $x^{n_{\ell+1}} \not\sim x^{n_\ell}$, hence $x^{n_{\ell+1}} \not\succsim x^{n_\ell}$, for every ℓ. As $\mathcal{D}$ is finite, there is an observation that repeats in the sequence, i.e., there are $r, s \in \mathbb{N}$ such that $r + 1 < s$, and $x^{n_r} = x^{n_s}$. But then $x^{n_{r=1}} \succsim x^{n_s} = x^{n_r}$, a contradiction. □

Lemma 6. *If $\mathcal{D}$ satisfies GARP there are numbers $u^i \in \mathbb{R}$ and $\lambda^i > 0$ such that*

$$u^i > u^j + \lambda^i(1 - p^i \cdot x^j) \qquad \textit{whenever } x^i \not\sim x^j$$

$$u^i = u^j \qquad \textit{whenever } x^i \sim x^j$$

$$\lambda^i = \lambda^j \qquad \textit{whenever } x^i \sim x^j$$

Proof. We proceed by induction on the number of observations. If $N = 1$, set $u^1 = \lambda^1 = 1$.

Suppose GARP holds for all databases comprised of $N-1$ or less observations, and take $\mathcal{D}$ comprised of N observations. By Lemma 5, and without loss of generality, suppose N is such that $p^i \cdot x^j > 1$ whenever $x^i \sim x^N$ and $x^j \not\sim x^N$.

- If $x^i \sim x^N$ for all i, set $u^i = \lambda^i = 1$ for every $i \in [N]$. Then the conditions hold.
- If there is j such that $x^j \not\sim x^N$, then the data set $(p^j, x^j)_{\{j : x^j \not\sim x^N\}}$ is a data set of $N-1$ or less observations for which the conditions hold. Take $\varepsilon > 0$, and for every i such that $x^i \sim x^N$ set

$$u^i = \min_{\{m : x^m \sim x^N\}} \min_{\{j : x^j \not\sim x^N\}} u^j - \lambda^j(1 - p^j \cdot x^m) - \varepsilon$$

Then, as $\sim$ is an equivalence relation, $u^i = u^j$ whenever $x^i \sim x^j$. Moreover, whenever $x^j \not\sim x^N$ and $x^m \sim x^N$

$$u^m \leq u^j - \lambda^j(1 - p^j \cdot x^m) - \varepsilon < u^j - \lambda^j(1 - p^j \cdot x^m)$$

Finally, whenever $x^i \sim x^N$ set

$$\lambda^i = \max\left\{ \max_{\{m : x^m \sim x^N\}} \max_{\{j : x^j \not\sim x^N\}} \frac{u^j - u^m}{p^m \cdot x^j - 1} + \varepsilon;\ 1 \right\}.$$

Hence $\lambda^i = \lambda^N > 0$ whenever $x^i \sim x^N$. Finally, if $x^i \sim x^N$ and $x^j \not\sim x^N$ then

$$\begin{aligned}
u^j + \lambda^i(1 - p^i \cdot x^j) &= u^j - \lambda^i(p^i \cdot x^j - 1) \\
&\leq u^j - \left(\frac{u^j - u^i}{p^i \cdot x^j - 1} + \varepsilon \right) (p^i \cdot x^j - 1) \\
&< u^j - \left(\frac{u^j - u^i}{p^i \cdot x^j - 1} \right) (p^i \cdot x^j - 1) \\
&= u^i
\end{aligned}$$

The first inequality follows from the definition of λ, and the second from $p^i \cdot x^j > 1$. We conclude that the conditions holds for all $i, j \in [N]$.

$\square$

Proof that Ga-1) implies Ga-3). If $\mathcal{D}_\wedge$ satisfies GARP, by Lemma 6 there are numbers $u^i \in \mathbb{R}$ and $\lambda^i > 0$ such that

$$u^i > u^j + \lambda^i(1 - q^i \cdot x^j) \qquad \text{whenever } x^i \not\sim_\wedge x^j \tag{24}$$

$$u^i = u^j \qquad \text{whenever } x^i \sim_\wedge x^j \tag{25}$$

For every $i \in [N]$, set $\mu^i = \lambda^i(p^i - q^i)$. The proof can be completed with the same reasoning shown in the proof that Sa-1) implies Sa-3) in Theorem 1. $\square$

F.3 Ga-3) implies Ga-4)

Proof that Ga-3) implies Ga-4). Suppose there are $u^i \in \mathbb{R}$, $\lambda^i > 0$ and $\mu^i \geq \mathbf{0}$ such that (G1) to (G5) hold. For each $i \in [N]$ define the functions $\phi^i : \mathbb{R}^K_+ \to \mathbb{R}$ and $V : \mathbb{R}^K_+ \to \mathbb{R}$ by

$$\phi^i(x) = u^i - \lambda^i(1 - p^i \cdot x) - \mu^i \cdot x$$
$$V(x) = \min_{i \in [N]} \phi^i(x)$$

Each ϕ^i is continuous, concave, and (by (G4)) strictly increasing. As the minimum of finitely many functions, V inherits these properties. Moreover, from (G2), (G3), and an argument analogous to Remark 2, we have

$$\phi^i(x) = \phi^j(x) \text{ for all } x \text{ whenever } x^i \sim x^j \,. \tag{26}$$

From (G1) and (26) we have $V(x^i) = \phi^i(x^i)$. Furthermore, as V is continuous there is $\eta > 0$ such that

$$V(x^i - x) = \phi^i(x^i - x) \quad \text{ for all } i \in [N] \text{ and } x \in B(\eta)\,. \tag{27}$$

Use such η to define $\tilde{U} : \mathbb{R}^K \to \mathbb{R}$ as in (6). Then $\tilde{U}$ is continuous, infinitely differentiable, and inherits both concavity and being strictly increasing from V. Furthermore for all $i \in [N]$

$$\begin{aligned}
\tilde{U}(x^i) &= \int_{B(\eta)} V(x^i - \xi)\rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} \phi^i(x^i - \xi)\rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} \left[u^i - \lambda^i(1 - p^i \cdot (x^i - \xi)) - \mu^i \cdot (x^i - \xi)\right] \rho_\eta(\xi)\, d\xi \\
&= \left[u^i - \lambda^i(1 - p^i \cdot x^i) - \mu^i \cdot x^i\right] \int_{B(\eta)} \rho_\eta(\xi)\, d\xi - (\lambda^i p^i - \mu^i) \cdot \int_{B(\eta)} \xi \rho_\eta(\xi)\, d\xi \\
&= u^i \,.
\end{aligned} \tag{28}$$

The second line follows from (27); the third one replaces the definition of ϕ^i; the fourth splits terms; and the last one follows from $p^i \cdot x^i = 1$, (G5), (7), and (8).

Finally take x such that $p^i \cdot x \leq 1$. Then

$$\begin{aligned}
\tilde{U}(x) &= \int_{B(\eta)} V(x - \xi)\rho_\eta(\xi)\, d\xi \\
&\leq \int_{B(\eta)} \phi^i(x - \xi)\rho_\eta(\xi)\, d\xi
\end{aligned}$$

$$
\begin{aligned}
&= \left[u^i - \lambda^i(1 - p^i \cdot x) - \mu^i \cdot x\right] \int_{B(\eta)} \rho_\eta(\xi)\, d\xi - (\lambda^i p^i - \mu^i) \cdot \int_{B(\eta)} \xi \rho_\eta(\xi)\, d\xi \\
&\leq u^i \\
&= \tilde{U}(x^i)\,.
\end{aligned}
$$

The second line follows from $i \in [N]$; the third one replaces the definition of ϕ^i and splits terms; the fourth one follows from $p^i \cdot x \leq 1$, $\mu^i \cdot x \geq 0$, (7), and (8); and the last one from (28).

Let $U : \mathbb{R}^K_+ \to \mathbb{R}$ be the restriction of $\tilde{U}$ to $\mathbb{R}^K_+$. Then U smoothly rationalizes $\mathcal{D}$. □

G Theorem 3

G.1 Ho-1) implies Ho-2)

Lemma 7. *Suppose $\mathcal{D}$ is rationalizable by a homothetic utility U and take $(x^{m_\ell})_{\ell \in [L]}$ such that*

$$
(p^{m_L} \cdot x^{m_1})(p^{m_1} \cdot x^{m_2})(p^{m_2} \cdot x^{m_3}) \ldots (p^{m_{L-1}} \cdot x^{m_L}) = 1\,.
$$

For each $\ell \in [L]$ define

$$
y^\ell = \left(\prod_{r=\ell}^{L-1} p^{m_r} \cdot x^{m_{r+1}} \right) x^{m_\ell}\,.
$$

Then $U(y^1) = U(y^2) = \ldots = U(y^L) = U(x^{m_L})$.

Proof. First, note that $y^L = x^{m_L}$. Hence, it is sufficient to show that $U(y^1) = U(y^2) = \ldots = U(y^L)$. We start by showing show that $U(y^\ell) \geq U(y^{\ell+1})$ for all $\ell \in [L-1]$, note that

$$
\begin{aligned}
p^{m_\ell} \cdot y^\ell &= p^{m_\ell} \cdot \left(\prod_{r=\ell}^{L-1} p^{m_r} \cdot x^{m_{r+1}} \right) x^{m_\ell} \\
&= \left(\prod_{r=\ell}^{L-1} p^{m_r} \cdot x^{m_{r+1}} \right) (p^{m_\ell} \cdot x^{m_\ell}) \\
&= \left(\prod_{r=\ell}^{L-1} p^{m_r} \cdot x^{m_{r+1}} \right) \\
&= (p^{m_\ell} \cdot x^{m_{\ell+1}}) \left(\prod_{r=\ell+1}^{L-1} p^{m_r} \cdot x^{m_{r+1}} \right) \\
&= p^{m_\ell} \cdot \left(\prod_{r=\ell+1}^{L-1} p^{m_r} \cdot x^{m_{r+1}} \right) x^{m_{\ell+1}} \\
&= p^{m_\ell} y^{\ell+1}
\end{aligned}
$$

We know that x^{m_ℓ} is optimal (according to U) when the price is p^{m_ℓ}. Hence, as U is homothetic, y^ℓ is optimal (according to U) when the price is $(p^{m_\ell} \cdot y^\ell)^{-1} p^{m_\ell}$. As $(p^{m_\ell} \cdot y^\ell)^{-1} p^{m_\ell} \cdot y^{\ell+1} = (p^{m_\ell} \cdot y^\ell)^{-1} p^{m_\ell} \cdot y^\ell = 1$, revealed preferences imply $U(y^\ell) \geq U(y^\ell + 1)$.

Since $U(y^\ell) \geq U(y^{\ell+1})$ for all $\ell \in [L-1]$, a sufficient condition to achieve equality is $U(y_L) \geq U(y^1)$. As $y^L = x^{m_L}$, then y^L is optimal when the price is p^{m_L}. Also

$$(p^{m_L} \cdot y^1) = p^{m_L} \cdot \left(\prod_{r=1}^{L-1} p^{m_r} \cdot x^{m_{r+1}} \right) x^{m_1} = (p^{m_L} \cdot x^{m_1})(p^{m_1} \cdot x^{m_2})(p^{m_2} \cdot x^{m_3}) \ldots (p^{m_{L-1}} \cdot x^{m_L}) = 1\,.$$

As y^1 is affordable at price p^{m_L}, we conclude that $U(y^L) \geq U(y^1)$. This completes the proof. □

Lemma 8. *Suppose $\mathcal{D}$ is smoothly rationalizable by U and take $\Gamma^H(\mathcal{D}) = (\tilde{q}^{i,H}, x^i)_{i \in [N]}$ then for every $i \in [N]$*

$$\nabla U(x^i) \leq \lambda^i \, \tilde{q}^{i,H}$$

where λ^i is the Lagrange multiplier of the budget constraint when the price is p^i.

Proof. Take an H1-sequence $s = (m_\ell)_{\ell \in [L]}$ and, without loss of generality, assume $m_L = i$. For each $\ell \in [L]$ define y^ℓ as in Lemma 7. By the same Lemma we have $U(x^i) = U(y^1) = u(y^2) = \ldots = U(y^L)$, where $y^L = x^{m_L}$. For $\ell \in [L]$ define

$$t^\ell = \left(\prod_{r=\ell}^{[L-1]} p^{m_r} \cdot x^{m_{r+1}} \right)^{-1} p^\ell\,.$$

Then we have $\{t^\ell\}_{\ell \in [L]} = I^H(s)$ Note that homotheticity of U implies that, for every ℓ, y^ℓ is the optimal choice, according to U, from price t^ℓ.

Define the data set $\hat{\mathcal{D}} = (t^\ell, y^\ell)_{\ell \in [L]}$. Since all the choices in $\hat{\mathcal{D}}$ are optimal according to U, then U rationalizes $\hat{\mathcal{D}}$. Furthermore, from the proof in Lemma 7 we conclude that all the choices in $\hat{\mathcal{D}}$ revealed indifferent to each other. Hence, Lemma 4 applied to $\hat{\mathcal{D}}$ implies $\nabla U(x^i) \leq \lambda^i \bigwedge_{\ell \in [L]} t^\ell = \lambda^i \bigwedge_{q \in I^H(s)} q$.

Since s is an arbitrary H1-sequence ending in i, we conclude that

$$\nabla U(x^i) \leq \bigwedge_{s \in \mathcal{S}^H(i)} \left(\lambda^i \bigwedge_{q \in I^H(s)} q \right) = \lambda^i \bigwedge_{s \in \mathcal{S}^H(i)} \left(\bigwedge_{q \in I^H(s)} q \right) = \lambda^i \bigwedge \bigcup_{s \in \mathcal{S}(i)^H} I^H(s) = \lambda^i \tilde{q}^{i,H}\,.$$

□

Lemma 9. *If $\mathcal{D}$ is smoothly rationalizable by a homothetic utility U, then so is $\Gamma^H(\mathcal{D})$.*

Proof. Suppose $\mathcal{D}$ is rationalizable by a homothetic utility U. Take $i \in [N]$ and x such that $\tilde{q}^{i,H} \cdot x \leq 1$; then

$$\begin{aligned} U(x^i) - U(x) &\geq \nabla U(x^i) \cdot (x^i - x) \\ &= \nabla U(x^i) \cdot x^i - \nabla U(x^i) \cdot x \\ &\geq \lambda^i - \lambda^i \, \tilde{q}^{i,H} \cdot x \\ &\geq 0 \, . \end{aligned}$$

The first line follows form concavity; the second distributes terms; the third from (11), Lemma 8 and $x \geq 0$; and the last one from $\tilde{q}^{i,H} \cdot x \leq 1$. We conclude that U, which is strictly increasing, concave, differentiable, and homothetic, rationalizes $\Gamma^H(\mathcal{D})$. □

Proof that Ho-1) implies Ho-2). Suppose $\mathcal{D}$ is smoothly rationalizable by a homothetic utility U. By Lemma 9 then $\Gamma^H(\mathcal{D})$ is also rationalized by U. Applying the same result iteratively implies that U rationalizes any finite iteration of Γ^H applied to $\mathcal{D}$. As $\mathcal{D}^H_\wedge$ is the result of iteratively applying Γ^H to $\mathcal{D}$ a finite number of times, then U rationalizes $\mathcal{D}^H_\wedge$. Finally, as U is homothetic this implies that $\mathcal{D}^H_\wedge$ satisfies HARP. □

G.2 Ho-2) implies Ho-3)

Lemma 10. *Suppose $\mathcal{D}^H_\wedge$ satisfies HARP. Then $x^i \approx^H x^j$ if, and only if, $(q^{j,H} \cdot x^i)(q^{i,H} \cdot x^j) = 1$.*

Proof. Sufficiency follows form the definition of $\approx^H$. For necessity suppose $x^i \approx^H x^j$; then there is an H1-sequence $(m_\ell)_{\ell \in [L]}$ such that $\ell' = i$ and (without loss of generality) $L = j$. By definition of $\mathcal{D}^H_\wedge$ we have

$$q^{j,H} \leq \left(\prod_{\ell=\ell'}^{L-1} q^{m_\ell,H} \cdot x^{m_{\ell+1}} \right)^{-1} q^{i,H} \, .$$

Furthermore, as $\mathcal{D}^H_\wedge$ satisfies HARP, it also satisfies GARP, hence $q^{i,H} \cdot x^i = 1$ and

$$q^{j,H} \cdot x^i \leq \left(\prod_{\ell=\ell'}^{L-1} q^{m_\ell,H} \cdot x^{m_{\ell+1}} \right)^{-1} \tag{29}$$

As $(m_\ell)_{\ell \in [L]}$ is an H1-sequence we have

$$(q^{i,H} \cdot x^{m_{\ell'+1}})(q^{m_{\ell'+1},H} \cdot x^{m_{\ell'+2}}) \ldots (q^{m_{L-1},H} \cdot x^j)(q^{j,H} \cdot x^{m_1})(q^{m_1,H} \cdot x^{m_2}) \ldots (q^{m_{\ell'-1},H} \cdot x^i) = 1 \, .$$

By definition of $\mathcal{D}^H_\wedge$ this implies

$$q^{i,H} \leq \left((q^{j,H} \cdot x^{m_1}) \prod_{\ell=1}^{\ell'-1} q^{m_\ell,H} \cdot x^{m_{\ell+1}} \right)^{-1} q^{j,H} \, .$$

which, as $q^{j,H} \cdot x^j = 1$, implies

$$q^{i,H} \cdot x^j \leq \left((q^{j,H} \cdot x^{m_1}) \prod_{\ell=1}^{\ell'-1} q^{m_\ell,H} \cdot x^{m_{\ell+1}} \right)^{-1} .$$

Multiplying the previous inequality and (29) we obtain

$$\begin{aligned}
(q^{j,H} \cdot x^i)(q^{i,H} \cdot x^{j,H}) &\leq \left[(q^{j,H} \cdot x^{m_1}) \left(\prod_{\ell=1}^{\ell'-1} q^{m_\ell,H} \cdot x^{m_{\ell+1}} \right) \left(\prod_{\ell=\ell'}^{L-1} q^{m_\ell,H} \cdot x^{m_{\ell+1}} \right) \right]^{-1} \\
&= \left[(q^{m_L,H} \cdot x^{m_1}) \left(\prod_{\ell=1}^{L} q^{m_\ell,H} \cdot x^{m_{\ell+1}} \right) \right]^{-1} \\
&= 1 .
\end{aligned}$$

The second line follows from $m_L = j$, and the third one because $(m_\ell)_{\ell\in[L]}$ is an H1-sequence. Since $\mathcal{D}^H_\wedge$ satisfies HARP we have $(q^{j,H} \cdot x^i)(q^{i,H} \cdot x^j) \geq 1$, which, together with the previous inequality, implies the desired result. $\square$

Lemma 11. *Let $\mathcal{Z}(i)$ be the set of all finite sequences of observations $(m_\ell)_{\ell\in[L]}$ satisfying $x^{m_L} = x^i$. Define*

$$v^i = \min_{\mathcal{Z}(i)} (p^{m_1} \cdot x^{m_2})(p^{m_2,H} \cdot x^{m_3}) \ldots (q^{m_{L-1},H} \cdot x^i)$$

If $\mathcal{D}^H_\wedge$ satisfies HARP, then $v^i \leq v^j\, q^{j,H} \cdot x^i$ for all $i, j \in [N]$, $v^i = v^j\, q^{j,H} \cdot x^i$ whenever $x^i \approx^H x^j$, and there is $i' \in [N]$ such that $v^j < v^m q^{m,H} \cdot x^j$ whenever $x^{i'} \approx^H x^j$ and $x^{i'} \not\approx^H x^m$.

Proof. As $\mathcal{D}^H_\wedge$ satisfies HARP, by the proof of Theorem 2 in Varian (1983a) we know that $\tilde{u}^i$ is well defined; this is, that the minimum is achieved. Furthermore, it follows from the same proof that $v^i \leq v^j\, q^{j,H} \cdot x^i$ for all i, j. Since $q^i_H \gg \mathbf{0}$ and $x^i > \mathbf{0}$ for all i, we have $v^i > 0$.

First we show that $v^i = v^j p^j \cdot x^i$ whenever $x^i \approx^H x^j$. By Lemma 10, $x^i \approx^H x^j$ implies $(q^{j,H} \cdot x^i)(q^{i,H} \cdot x^j) = 1$. As $v^i \leq v^j\, q^{j,H} \cdot x^i$ and $v^j \leq v^i\, q^{i,H} \cdot x^j$ we have

$$\begin{aligned}
v^i &\leq v^j\, (q^{j,H} \cdot x^i) \left(\frac{q^{i,H} \cdot x^j}{q^{i,H} \cdot x^j} \right) \left(\frac{q^{j,H} \cdot x^i}{q^{j,H} \cdot x^i} \right) \\
&= v^j\, \frac{(q^{i,H} \cdot x^j)(q^{j,H} \cdot x^i)}{q^{i,H} \cdot x^j} \\
&= v^j\, \frac{1}{q^{i,H} \cdot x^j} \\
&\leq v^i .
\end{aligned}$$

Therefore all the inequalities have to be equalities; in particular $v^i\, q^{i,H} \cdot x^j = v^j$.

We show that there is $i' \in [N]$ such that $i' \in [N]$ such that $v^j < v^m q^{m,H} \cdot x^j$ whenever $x^{i'} \approx^H x^j$ and $x^{i'} \not\approx^Q x^m$ by contradiction. Suppose for every $i \in [N]$ there are $x^j \approx^Q x^i$ and $x^m \not\approx x^i$ such that $v^j = v^m\, q^{m,H} \cdot x^j$. Since $x^j \approx^Q x^i$, Lemma 10 implies $(q^{i,H} \cdot x^j)(q^{j,H} \cdot x^i) = 1$. Hence

$$v^m\,(q^{m,H} \cdot x^j)(q^{j,H} \cdot x^i) = v^j(q^{j,H} \cdot x^i) = v^i(q^{i,H} \cdot x^j)(q^{j,H} \cdot x^i) = v^i.$$

In particular, the previous equality also holds if $i = j$. Thus we can construct an infinite sequence $(x^{m_\ell})_{\ell \in [L]}$ such that, for all ℓ odd, $x^{m_\ell} \approx^H x^{m_{\ell+1}}$ and $x^{m_\ell} \not\approx^H x^{m_{\ell+2}}$, which satisfies

$$v^{m_1} = v^{m_2} q^{m_2,H} \cdot x^{m_1} = v^{m_3}(q^{m_2,H} \cdot x^{m_1}) = \ldots = v^{m_\ell} \prod_{s=1}^{\ell-1} q^{m_{s+1},H} \cdot x^{m_s} = \ldots$$

Since the sequence is infinite and there are finitely many observations, there have to be ℓ', ℓ'', with $\ell' < \ell'' + 3$ such that $m_{\ell'} = m_{\ell''}$. Then

$$v^{m_{\ell'}} \prod_{\ell=1}^{\ell'-1} q^{m_{\ell+1},H} \cdot x^{m_\ell} = v^{m_{\ell''}} \prod_{\ell=1}^{\ell''-1} q^{m_{\ell+1},H} \cdot x^{m_\ell} = v^{m_{\ell''}} \left(\prod_{\ell=1}^{\ell'-1} q^{m_{\ell+1},H} \cdot x^{m_\ell} \right) \left(\prod_{\ell=\ell'}^{\ell''-1} q^{m_{\ell+1},H} \cdot x^{m_\ell} \right)$$

Since $v^{m_{\ell'}} = v^{m_{\ell''}}$ and $q^{m_{\ell'},H} = q^{m_{\ell''},H}$ we have

$$1 = \prod_{\ell=\ell'}^{\ell''-1} q^{m_{\ell+1},H} \cdot x^{m_\ell} = (q^{m_{\ell'+1},H} \cdot x^{m_{\ell'}})(q^{m_{\ell'+2},H} \cdot x^{m_{\ell'+1}}) \ldots (q^{m_{\ell'},H} \cdot x^{m_{\ell''-1}})\,.$$

Since $x^{m_{\ell'}}$, $x^{m_{\ell'+1}}$, $x^{m_{\ell'+2}}$ and $x^{m_{\ell'+3}}$ are in the same H1-sequence, we have $x^{m_{\ell'}} \approx^H x^{m_{\ell'+2}}$ and $x^{m_{\ell'+1}} \approx^H x^{m_{\ell'+3}}$. This contradicts the fact that $x^{m_\ell} \not\approx^H x^{m_{\ell+2}}$ whenever ℓ is odd. ☐

Lemma 12. *If $\mathcal{D}^H_\wedge$ satisfies HARP then there are numbers $u^i > 0$ such that $u^i\, q^{i,H} \cdot x^j > u^j$ whenever $x^i \not\approx^Q x^j$ and $u^i\, q^{i,H} \cdot x^j = u^j$ whenever $x^i \approx^H x^j$.*

Proof. We show the result by induction of N. If $N = 1$ then $u^1 = 1$ satisfies the conditions.

Suppose the conditions hold for any database of $N-1$ or less elements, and take $\mathcal{D}^H_\wedge$ comprised of N observations. Take the numbers v^i defined in Lemma 11 and i' such that $v^j < v^m q^{m,H} \cdot x^j$ whenever $x^{i'} \approx^H x^j$ and $x^{i'} \not\approx^H x^m$. Denote $E = \{i \in [N] : x^i \approx^H x^{i'}\}$ and $D = [N] \setminus E$.

- If $D = \emptyset$, Lemma 11 imply that the numbers v^i satisfy the conditions.

- If $D \neq \emptyset$, then $(q^{i,H}, x^i)_{i \in D}$ is a data set comprised of $N-1$ or less observations. By induction hypothesis there are numbers $\tilde{u}^m > 0$ for $m \in D$ such that the conditions hold. Take the numbers v^i defined in Lemma 11 and define

$$\alpha = \min_{i \in I} \min_{m \in D} \frac{v^m\, p^m \cdot x^i}{v^i} - 1 > 0$$

$$u^i = \left(1 + \frac{\alpha}{2}\right) v^i \qquad \text{for every } i \in E\,.$$

Then

$$u^i\, p^i \cdot x^m > v^m \qquad \text{whenever } i \in E \text{ and } m \in D, \tag{30}$$

$$v^m\, p^m \cdot x^i > u^i \qquad \text{whenever } i \in E \text{ and } m \in D, \tag{31}$$

$$u^i\, p^i \cdot x^j = u^j \qquad \text{whenever } i, j \in E\,. \tag{32}$$

Take a sequence $\beta^n \to 1$, where $\beta^n \in (0,1)$ for all n, and for each $m \in D$ and $n \in \mathbb{N}$ define

$$w^m(n) = (v^m)^{\beta^n}\, (\tilde{u}^m)^{1-\beta^n}\,.$$

Take any $m, m' \in D$ and $n \in \mathbb{N}$. If $x^m \approx^Q x^{m'}$

$$w^m(n)\, p^m \cdot x^{m'} = (v^m\, p^m \cdot x^{m'})^{\beta^n}\, (\tilde{u}^m\, p^m \cdot x^{m'})^{1-\beta^n} = (v^{m'})^{\beta^n}\, (\tilde{u}^{m'})^{1-\beta^n} = w^{m'}(n)\,;$$

and if $x^m \not\approx^Q x^{m'}$

$$w^m(n)\, p^m \cdot x^{m'} = (v^m\, p^m \cdot x^{m'})^{\beta^n}\, (\tilde{u}^m\, p^m \cdot x^{m'})^{1-\beta^n} > (v^{m'})^{\beta^n}\, (\tilde{u}^{m'})^{1-\beta^n} = w^{m'}(n)\,.$$

Since $w^m(n) \to v^m$ for every $m \in \mathcal{D}$, for n_0 large enough we have

$$\left.\begin{aligned} u^i\, p^i \cdot x^m &> w^m(n_0) \\ w^m(n_0)\, p^m \cdot x^i &> u^i \end{aligned}\right\} \qquad \text{whenever } i \in E \text{ and } m \in D\,.$$

Setting $u^m = w^m(n_0)$ for all $m \in D$ assures that the numbers u^i satisfy the desired properties.

□

Proof that Ho-2) implies Ho-3). Let $\mathcal{D}^H_\wedge$ satisfy HARP, and take numbers $u^i > 0$ such that $u^i\, q^{i,H} \cdot x^j > u^j$ whenever $x^i \not\approx^Q x^j$ and $u^i\, q^{i,H} \cdot x^j = u^j$ whenever $x^i \approx^H x^j$ (Lemma 12). For every $i \in [N]$

$$\mu^i = u^i (p^i - q^{i,H})\,.$$

Since $u^i > 0$ and $p^i \geq q^{i,H}$ we have $\mu^i \geq \mathbf{0}$.

Take $x^i \not\approx^H x^j$. We have $u^i\, p^i \cdot x^j = u^i\, q^{i,H} \cdot x^j + u^i (p^i - q^{i,H}) \cdot x^j = u^i\, q^{i,H} \cdot x^j + \mu^i \cdot x^j > u^j + \mu^i \cdot x^j$. hence (H1) holds. Similarly if $x^i \approx^H x^j$ then $u^i\, p^i \cdot x^j = u^i\, q^{i,H} \cdot x^j + u^i (p^i - q^{i,H}) \cdot x^j = u^i\, q^{i,H} \cdot x^j + \mu^i \cdot x^j = u^j + \mu^i \cdot x^j$. and (H2) holds.

To see that (H3) holds take $x^i \approx^H x^j$. By Lemma 10 we have $(q^{i,H} \cdot x^j)(q^{j,H} \cdot x^i) = 1$. Then by definition of $\mathcal{D}^H_\wedge$ we have $q^{i,H} \leq (q^{j,H} \cdot x^i)q^{j,H}$ and $q^{j,H} \leq (q^{i,H} \cdot x^i)^{-1}q^{j,H}$. By (H2) we have

$$u^i\, p^i - \mu^i = u^i\, q^{i,H} = u^j(q^{j,H} \cdot x^i)q^{i,H} \leq u^j(q^{j,H} \cdot x^i)\frac{q^{j,H}}{q^{j,H} \cdot x^i} = u^j q^{j,H} = u^j p^j - \mu^j$$

and similarly $u^j\, p^j - \mu^j \leq u^i\, p^i - \mu^i$. The two inequalities imply (H3).

Since $u^i > 0$ and $q^{i,H} \gg \mathbf{0}$ we have $u^i\, p^i - \mu^i = u^i - q^{i,H} \gg \mathbf{0}$, therefore (H4) holds. Finally, as $D^H_\wedge$ satisfies HARP it also satisfies GARP; therefore $q^{i,H} \cdot x^i = 1$ and $\mu^i \cdot x^i = u^i(p^i \cdot x^i - \mu^{i,H} \cdot x^i) = \mathbf{0}$, i.e., (H5) holds. ◻

G.3 Ho-3) implies Ho-4)

We split this proof on two steps. The first step finds a strictly increasing, concave, and infinitely differentiable function f (along with a suitable domain S) and develops some of its properties. The second step uses f to construct a homothetic U that smoothly rationalizes the data.

Lemma 13. *There exist an open set S and a strictly increasing, concave, and infinitely differentiable function $f : S \to \mathbb{R}$, satisfying the following properties:*

1. $\left(\mathbb{R}^K_+ \setminus \{\mathbf{0}\}\right) \subset S$*;*
2. *For every $x \in S$ there is a unique value $\alpha(x) > 0$ satisfying $f(\alpha(x)^{-1}x) = 1$.*
3. *For every $y > \mathbf{0}$ we have*

$$f\left(\frac{x^i}{f(x^i)}\right) \geq f(y) \text{ whenever } y^i \geq \mathbf{0} \text{ and } p^i \cdot y \leq \frac{1}{f(x^i)}\, p^i \cdot x^i$$

Proof. Take $u^i > 0$ and $\mu^i \geq \mathbf{0}$ satisfying (H1)-(H5). Define

$$\phi^i(x) = (u^i p^i - \mu^i) \cdot x$$
$$V(x) = \min_{i \in [N]} \phi^i(x)$$

By (H4), ϕ^i is strictly increasing, and as a linear function is concave and homogeneous of degree one. Hence V inherits all these properties. Furthermore, (H4) implies $V(x) > 0$ whenever $x > \mathbf{0}$.

By (H3) we have $\phi^i(x) = \phi^j(x)$ for all x whenever $x^i \approx^H x^j$. Hence, (H1) and (H2) imply $V(x^i) = u^i$. Furthermore, by (H1) there is κ small enough such that, for every $\xi \in B(\kappa)$, $\phi^i(x^i - \xi) = \phi^j(x^i - \xi)$ whenever $x^i \approx^H x^j$ and $\phi^i(x^i - \xi) < \phi^j(x^i - \xi)$ whenever $x^i \not\approx^H x^j$.

For every $i \in [N]$ define $y^i = V(x^i)^{-1}x^i$, and note that $V(y^i) = \phi^i(y^i) = 1$. Let $\overline{v} = \max_{i\in[N]} V(x^i) > 0$ and let $\eta = \kappa \min\{1, 1/\overline{v}\} > 0$. Using the value of η defined above, define the function $\tilde{U} : \mathbb{R}^K \to \mathbb{R}$ as in (6). Then $\tilde{U}$ is strictly increasing, concave, and infinitely differentiable.

Let

$$S = \bigcap_{i\in[N]} \left\{x \in \mathbb{R}^K_+ : \phi^i(x) > 0\right\} .$$

Since each function ϕ^i is affine, the sets $\left\{x \in \mathbb{R}^K_+ : \phi^i(x) > 0\right\}$ are open, hence S also is.

Let $f : S \to \mathbb{R}$ be the restriction of $\tilde{U}$ to S. We now show the required properties, following the same enumeration as the statement of the Lemma.

1. From (H4), for every $x > \mathbf{0}$ we have $\phi^i(x) > 0$ for all $i \in [N]$. Therefore $\left(\mathbb{R}^K_+ \setminus \{\mathbf{0}\}\right) \subset S$.

2. Take an arbitrary $x \in S$; we show that there is a unique $\alpha(x)$ such that $f(\alpha(x)^1 x) = 1$. Since $x \in S$, we have $\phi^i(x) = (u^i\, p^i - \mu^i) \cdot x > 0$ for all $i \in [N]$, therefore (as $[N]$ is finite) $\min_{i\in[N]}(u^i\, p^i - \mu^i) \cdot x > 0$. Hence for $\beta > 0$ small enough

$$\beta^{-1} \min_{i\in[N]}(u^i\, p^i - \mu^i) \cdot x > 0 > 1 + \max_{j\in[N]}(u^j\, p^j - \mu^j) \cdot (\eta\mathbf{1})$$

Thus

$$\begin{aligned}
f(\beta^{-1}x) &= \int_{B(\eta)} \min_{i\in[N]}(u^i\, p^i - \mu^i) \cdot (\beta^{-1}x - \xi)\rho_\eta(\xi)\, d\xi \\
&\geq \int_{B(\eta)} \left(\beta^{-1} \min_{i\in[N]}(u^i\, p^i - \mu^i) \cdot x - \max_{j\in[N]}(u^j\, p^j - \mu^j) \cdot \xi\right) \rho_\eta(\xi)\, d\xi \\
&\geq \int_{B(\eta)} \left(\beta^{-1} \min_{i\in[N]}(u^i\, p^i - \mu^i) \cdot x - \max_{j\in[N]}(u^j\, p^j - \mu^j) \cdot (\eta\mathbf{1})\right) \rho_\eta(\xi)\, d\xi \\
&> \int_{B(\eta)} \rho_\eta(\xi)\, d\xi \\
&= 1
\end{aligned}$$

Fix $i \in [N]$. Since $\tilde{U}$ is continuous

$$\begin{aligned}
\lim_{\beta\to\infty} f(\beta^{-1}x) &= \lim_{\beta\to\infty} \tilde{U}(\beta^{-1}y) \\
&= \tilde{U}(\mathbf{0}) \\
&= \int_{B(\eta)} \min_{j\in[N]} \phi^j(\mathbf{0} - \xi)\rho_\eta(\xi)\, d\xi \\
&\leq \int_{B(\eta)} \phi^i(\mathbf{0} - \xi)\rho_\eta(\xi)\, d\xi
\end{aligned}$$

$$= (u^i p^i - \mu^i) \cdot \left(\mathbf{0} \int_{B(\eta)} \rho_\eta(\xi)\, d\xi - \int_{B(\eta)} \xi \rho_\eta(\xi)\, d\xi \right)$$
$$= 0$$

Since f inherits continuity from $\tilde{U}$ and $\lim_{\beta\to\infty} \beta^{-1}x = \mathbf{0}$, the Intermediate Value Theorem implies that there is $\alpha(x) \in (\beta, \infty)$ satisfying $f(\alpha(x)^{-1}x) = 1$. We show that $\alpha(x)$ is unique by showing that $f(\alpha x)$ is strictly increasing in α. Take $\gamma > \alpha > 0$; then

$$\begin{aligned}
f(\gamma x) - f(\alpha x) &= \int_{B(\eta)} \min_{j\in[N]} \phi^j(\gamma x - \xi)\rho_\eta(\xi)\, d\xi - \int_{B(\eta)} \min_{j\in[N]} \phi^j(\alpha x - \xi)\rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} \left(\min_{j\in[N]} \phi^j(\gamma x - \xi) - \min_{j\in[N]} \phi^j(\alpha x - \xi) \right) \rho_\eta(\xi)\, d\xi \\
&\geq \int_{B(\eta)} \left(\min_{j\in[N]} \phi^j(\gamma x - \xi) - \phi^j(\alpha x - \xi) \right) \rho_\eta(\xi)\, d\xi \\
&= (\gamma - \alpha) \int_{B(\eta)} \min_{j\in[N]} (u^i p^i - \mu^i) \cdot x \rho_\eta(\xi)\, d\xi \\
&> 0\,.
\end{aligned}$$

where the first inequality follows from $\min f(x) - \min g(x) \geq \min(f(x) - g(x))$,[37] and the last one from $\gamma > \alpha$ and $x \in S$.

3. As $\eta \leq \kappa$, $V(x^i - \xi) = \phi^i(x^i - \xi)$ whenever $\xi \in B(\eta)$. Hence

$$\begin{aligned}
f(x^i) &= \int_{B(\eta)} V(x^i - \xi)\rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} \phi^i(x^i - \xi)\rho_\eta(\xi)\, d\xi \\
&= \int_{B(\eta)} (u^i p^i - \mu^i) \cdot (x^i - \xi)\rho_\eta(\xi)\, d\xi \\
&= (u^i p^i - \mu^i) \cdot \left(x^i \int_{B(\eta)} \rho_\eta(\xi)\, d\xi - \int_{B(\eta)} \xi \rho_\eta(\xi)\, d\xi \right) \\
&= (u^i p^i - \mu^i) \cdot x^i \\
&= u^i
\end{aligned}$$

Therefore $x^i/\tilde{U}(x^i) = y^i$. In a similar way, as $\eta \leq \kappa/\bar{v} \leq \kappa/V(x^i)$, whenever $\xi \in B(\eta)$ we have $V(x^i)\xi \in B(V(x^i)\eta) \subset B(\kappa)$, and

$$V(y^i - \zeta) = \frac{1}{V(x^i)} V(V(x^i)y^i - V(x^i)\zeta)$$

[37] Let x_0 be the minimizer of $f(x)$. Then $\min f(x) - \min g(x) \geq \min f(x) - g(x_0) = f(x_0) - g(x_0) \geq \min(f(x) - g(x))$.

$$
\begin{aligned}
&= \frac{1}{V(x^i)} V(x^i - V(x^i)\zeta) \\
&= \frac{1}{V(x^i)} \phi^i(x^i - V(x^i)\zeta) \\
&= \frac{1}{V(x^i)} \phi^i(V(x^i)y^i - V(x^i)\zeta) \\
&= \phi^i(y^i - \zeta)\,.
\end{aligned}
$$

Hence

$$
\begin{aligned}
f(y^i) &= \int_{B(\eta)} V(y^i - \xi)\rho_\eta(\xi)\,d\xi \\
&= \int_{B(\eta)} \phi^i(y^i - \xi)\rho_\eta(\xi)\,d\xi \\
&= \int_{B(\eta)} (u^i\,p^i - \mu^i)\cdot(y^i - \xi)\rho_\eta(\xi)\,d\xi \\
&= (u^i\,p^i - \mu^i)\cdot\left(y^i \int_{B(\eta)} \rho_\eta(\xi)\,d\xi - \int_{B(\eta)} \xi\rho_\eta(\xi)\,d\xi\right) \\
&= (u^i\,p^i - \mu^i)\cdot y^i \\
&= u^i\,p^i \cdot \frac{x^i}{u^i} \\
&= 1\,.
\end{aligned}
$$

Finally, take $y \geq \mathbf{0}$ such that $p^i \cdot y \leq p^i \cdot y^i$. We have

$$
\begin{aligned}
\tilde{U}(y) &= \int_{B(\eta)} V(y - \xi)\rho_\eta(\xi)\,d\xi \\
&= \int_{B(\eta)} \min_{j\in[N]} \phi^j(y - \xi)\rho_\eta(\xi)\,d\xi \\
&\leq \int_{B(\eta)} \phi^i(y - \xi)\rho_\eta(\xi)\,d\xi \\
&= \int_{B(\eta)} (u^i\,p^i - \mu^i)\cdot(y - \xi)\,\rho_\eta(\xi)\,d\xi \\
&= (u^i\,p^i - \mu^i)\cdot\left(y \int_{B(\eta)} \rho_\eta(\xi)\,d\xi - \int_{B(\eta)} \xi\rho_\eta(\xi)\,d\xi\right) \\
&= u^i\,p^i\cdot y - \mu^i\cdot y \\
&\leq u^i\,p^i\cdot y^i \\
&= 1 \\
&= f(y^i)
\end{aligned}
$$

The second inequality follows from $\mu^i \cdot y \geq 0$ and $p^i \cdot y \leq p^i \cdot y^i$. This completes the proof.

□

Proof that Ho-3) implies Ho-4). Take the open set S and the function $f : S \to \mathbb{R}$ from Lemma 13. For every $x \in S$ take the unique value $\alpha(x) > 0$ such that $f(\alpha(x)^{-1}x) = 1$ and define $W : S \cup \{\mathbf{0}\} \to \mathbb{R}$ by

$$W(x) = \begin{cases} \alpha(x) & \text{if } x \geq \mathbf{0} \\ 0 & \text{if } x = \mathbf{0} \end{cases} \tag{33}$$

From the previous definition we have $W(x) > 0$ whenever $x \in S$. Furthermore, from the proof of Lemma 13 we have that $f(\alpha x)$ is strictly increasing in α. Hence both f and W have the same upper and lower contour sets around one; this is, $W(x) \geq 1 \iff f(x) \geq 1$, and $W(x) > 1 \iff f(x) > 1$.

Now we show that W is continuous, strictly increasing, concave, infinitely differentiable in S, homothetic, and rationalizes the data.

- W is homothetic: Take $\lambda > 0$. Since $W(\mathbf{0}) = 0$ we have $\lambda W(\mathbf{0}) = W(\lambda \mathbf{0})$. If $x \neq \mathbf{0}$ we have $f(W(x)^{-1}x) = 1$, and $1 = f(W(\lambda x)^{-1}\lambda x)$. Since there is a unique value $\alpha(x) > 0$ such that $f(\alpha(x)^{-1}x) = 1$, we have $\alpha(x) = W(x)^{-1} = W(\lambda x)^{-1}\lambda$, therefore $\lambda W(x) = W(\lambda x)$.

- W is strictly increasing: Take x, y such that $x > y$. If $y = \mathbf{0}$ by definition of W we have $W(x) > 0 = W(y)$. If $y \neq \mathbf{0}$, towards a contradiction suppose $U(y) \geq U(x) > 0$, which implies $W(x)^{-1}x > W(y)^{-1}y$. As f is strictly increasing, we have $f(W(y)^{-1}y) < f(W(x)^{-1}x)$. But the definition of W implies $f(W(y)^{-1}y) = f(W(x)^{-1}x) = 1$, a contradiction.

- W is concave: Denote by $P_a^f = \{x \in \mathbb{R}_+^K : f(x) \geq a\}$ and $P_a^W = \{x \in \mathbb{R}_+^K : W(x) \geq a\}$ the superlevel sets of f and W at a, respectively. Since both W and f have the same contour sets around one, we have $P_1^f = P_1^W$. Since f is concave, P_1^f is convex, and P_1^W also is. Furthermore, homotheticity of W implies that for any $a \geq 0$ we have $P_a^W = \{a\,x \in \mathbb{R}_+^K : x \in P_1^W\}$, hence P_a^W is convex as well. This implies that the epigraph of W is convex, therefore W is concave.

- W is continuous: For continuity at $\mathbf{0}$, take $x^n \to \mathbf{0}$ ($x^n \in S$) and $\delta > 0$, and denote by $\mathbf{1}$ the K-dimensional vector of ones. Since $x^n \to \mathbf{0}$, for n large enough we have $x_k^m < W(\mathbf{1})^{-1}\delta$ for all $k \in [K]$ and $m \geq n$. Thus, as f is strictly increasing, $f(\delta^{-1}x^m) < f(W(\mathbf{1})^{-1}\mathbf{1}) = 1$. Homogeneity of degree one implies $W(x^m) < \delta$. Since $W(x^n) \geq 0$ for all n and δ is arbitrary, we have $\lim_{n\to\infty} W(x^n) = W(\mathbf{0}) = 0$.

To see that W is continuous on S take a, b such that $0 < a < b$. Since $W(x) > 0$ for all $x \in S$,

$$W^{-1}((a,b)) = W^{-1}((a,\infty) \cap (0,b)) = W^{-1}((a,\infty)) \cap W^{-1}((0,b))$$

Furthermore

$$W^{-1}((a,\infty)) = \{x \in S : W(x) > a\} = \left\{x \in S : f(a^{-1}x) < 1\right\} = \left\{a^{-1}y \in S : f(y) < 1\right\} .$$

Since f is continuous, the set $\left\{y \in \mathbb{R}^K_+ : f(y) < 1\right\}$ is open, hence $U^{-1}((a,\infty))$ also is. Similarly we can show that $U^{-1}((0,b))$ is open. As the intersection of two open sets is open, we conclude that W is continuous in S. Therefore W is continuous.

- W is infinitely differentiable in S: Denote by $I^f_a = \{x \in \mathbb{R}^K_+ : Z(x) = a\}$ and $I^W_a = \{x \in \mathbb{R}^K_+ : W(x) = a\}$ the level sets of f and W at level a, respectively. Since f and W have the same contour sets around one, we have $I^f_1 = I^U_1$. Since f is infinitely differentiable, I^f_1 is a $K-1$ dimensional C^∞ manifold (see Debreu, 1972), and I^W_1 also is. Furthermore, as W is homogeneous of degree one, for every $a > 0$ we have $I^W_a = \{ax \in \mathbb{R}^K_+ : x \in I^W_1\}$, hence I^W_a is also a $K-1$ dimensional C^∞ manifold. Since W is continuous, strictly increasing, homogeneous of degree one, and all its indifference sets in S are $K-1$ dimensional C^∞ manifolds, Theorem 1 in Neilson (1991) implies that W is infinitely differentiable in S.[38]

- $W(x)$ rationalizes $\mathcal{D}$: Take $i \in [N]$ and x satisfying $p^i \cdot x \leq 1$. Recall that $y^i = f(x^i)^{-1}x^i$. Since $f(y^i) = 1$, then $f(y) \leq 1$ whenever $p^i \cdot y \leq p^i \cdot y^i$. As f and W have the same contour sets around 1 and $f(y^i) = 1$, then $W(y) \leq 1$ whenever $p^i \cdot y \leq p^i \cdot y^i$. Furthermore, as $f(y^i) = 1$, we have $W(y^i) = 1$ as well. Hence as W is homothetic $W(x^i) = W(f(x^i)y^i) = f(x^i)W(y^i) = f(x^i)$ and whenever $p^i \cdot x \leq 1$

$$\begin{aligned} W(x) &= f(x^i)W\left(\frac{p^i \cdot x^i}{f(x^i)}x\right) \\ &= f(x^i)W((p^i \cdot y^i)\,x) \\ &\leq f(x^i) \\ &= W(x^i)\,. \end{aligned}$$

The first equality follows from W being homogeneous of degree one and $p^i \cdot x^i = 1$, the second from the definition of y^i, and the inequality from $p^i \cdot (p^i \cdot y^i)\,x \leq p^i \cdot y^i$, which implies $W((p^i \cdot y^i)\,x) \leq W(y^i) = 1$. Therefore W rationalizes $\mathcal{D}$.

[38] Although Theorem 1 in Neilson (1991) is developed for a function whose domain is $\mathbb{R}^K_{++}$, the proof does not uses anything particular about that domain and hence applies to every open domain.

Let $U : \mathbb{R}^K_+ \to \mathbb{R}$ the restriction of W to $\mathbb{R}^K_+$. Then U is continuous, strictly increasing, infinitely differentiable in $\mathbb{R}^K_+ \setminus \{\mathbf{0}\}$, homothetic, and rationalizes the data. ☐

H Theorem 4

H.1 Qu-1) implies Qu-2)

Lemma 14. *Suppose $\mathcal{D}$ is rationalizable by a quasilinear utility $U(x)+y$, where U is differentiable, and $(m_\ell)_{\ell\in[L]}$ is such that*

$$p^{m_L} \cdot (x^{m_1} - x^{m_L}) + \sum_{\ell=1}^{L-1} p^{m_\ell} \cdot (x^{m_\ell} - x^{m_{\ell+1}}) = 0\,. \tag{34}$$

Then $\nabla U(x^{m_1}) = \nabla U(x^{m_2}) = \ldots = \nabla U(x^{m_L})$.

Proof. Since $\mathcal{D}$ is rationalizable by a quasilinear utility $U(x) + y$, we have $p^{m_\ell} \cdot (x^{m_{\ell+1}} - x^{m_\ell}) \geq U(x^{m_{\ell+1}}) - U(x^{m_\ell})$ for all $\ell \in [L-1]$, and $p^{m_L} \cdot (x^{m_1} - x^{m_L}) \geq U(x^{m_1}) - U(x^{m_L})$. Replacing into (34) we obtain

$$0 = p^{m_L} \cdot (x^{m_1} - x^{m_L}) + \sum_{\ell=1}^{L-1} p^{m_\ell} \cdot (x^{m_\ell+1} - x^{m_\ell}) \geq U(x^{m_1}) - U(x^{m_L}) + \sum_{\ell=1}^{L-1} U(x^{m_\ell+1}) - U(x^{m_\ell}) = 0$$

Since the inequality of the previous equation has to be an equality, we conclude that $p^{m_\ell} \cdot (x^{m_{\ell+1}} - x^{m_\ell}) = U(x^{m_{\ell+1}}) - U(x^{m_\ell})$, i.e. $U(x^{m_{\ell+1}})$ is optimal at price p^ℓ, for all $\ell \in [L-1]$. Similarly, x^{m_1} is optimal from price p^{m_L}. Since U is differentiable, Theorem 5 implies $\nabla U(x^{m_\ell}) = \nabla U(x^{m_{\ell+1}})$ for all $\ell \in [L-1]$ and $\nabla U(x^{m_L}) = \nabla U(x^{m_1})$. Therefore $\nabla U(x^{m_1}) = \nabla U(x^{m_2}) = \ldots = \nabla U(x^{m_L})$. ☐

Lemma 15. *If $\mathcal{D}$ is rationalizable by a quasilinear utility $U(x) + y$, then it is rationalizable by $U(x)$.*

Proof. If $\mathcal{D}$ is rationalizable by a quasilinear utility $U(x) + y$ then $U(x^i) - p^i \cdot x^i \geq U(x) - p^i \cdot x$ for every $i \in [N]$ and $x \in \mathbb{R}^k_+$. In particular, if $p^i \cdot x \leq 1$ then $U(x^i) \geq U(x) + p^i \cdot x^i - p^i \cdot x = U(x) + 1 - p^i \cdot x \geq U(x)$. Therefore $\mathcal{D}$ is rationalizable by U. ☐

Lemma 16. *If $\mathcal{D}$ is rationalizable by the quasilinear $U(x)+y$ where U is strictly increasing, concave, and differentiable, then $\Gamma^Q(\mathcal{D}) = \left(\tilde{q}^{i,Q}, x^i\right)_{i\in[N]}$ also is. Furthermore $\tilde{q}^{i,Q} \cdot x^i = 1$ for all i.*

Proof. Suppose $\mathcal{D}$ is quasilinear rationalizable by $U(x) + y$. From the first order conditions of the maximization problem, for every $i \in [N]$ we have $\nabla U(x^i) = p^i - \mu^i \leq p^i$. Furthermore, the complementary slackness conditions on μ^i imply $\mu^i \cdot x^i = 0$; therefore $\nabla U(x^i) \cdot x^i = p^i \cdot x^i$.

Take $i \in [N]$ and $s = (m_\ell)_{\ell \in [L]} \in \mathcal{S}^Q(i)$ a Q0-sequence satisfying $m_L = i$. Since $\mathcal{D}$ is quasilinear rationalizable by U, then $\nabla U(x^{m_\ell}) = p^{m_\ell} + \mu^{m_\ell} \leq p^{m_\ell}$ for all $\ell \in [L]$. Furthermore, Lemma 14 implies $\nabla U(x^i) = \nabla U(x^{m_\ell})$ for every $\ell \in [L]$; hence $\nabla U(x^i) \leq p$ for every $p \in I^Q(s)$. Thus $\nabla U(x^i) \leq \bigwedge I^Q(s)$. As s is an arbitrary sequence in $\mathcal{S}^Q(i)$ we conclude that

$$\nabla U(x^i) \leq \bigwedge_{s \in \mathcal{S}^Q(i)} \bigwedge I^Q(s) = \bigwedge \bigcup_{s \in \mathcal{S}^Q(i)} I^Q(s) = \tilde{q}^{i,Q} \,.$$

Define $f : \mathbb{R}^K_+ \times \mathbb{R} \to \mathbb{R}$ by $f((x, y)) = U(x) - y$. Since U is concave, f also is. Hence for every $i \in [N]$ and $x \in \mathbb{R}^K_+$

$$\begin{aligned}
U(x^i) - \tilde{q}^{i,Q} \cdot x^i - \big(U(x) - \tilde{q}^{i,Q} \cdot x\big) &= f(x^i,\, \tilde{q}^{i,Q} \cdot x^i) - f(x,\, \tilde{q}^{i,Q} \cdot x) \\
&\geq \nabla f(x^i,\, \tilde{q}^{i,Q} \cdot x^i) \cdot \big((x^i,\, \tilde{q}^{i,Q} \cdot x^i) - (x,\, \tilde{q}^{i,Q} \cdot x)\big) \\
&= \nabla U(x^i) \cdot (x^i - x) - \big(\tilde{q}^{i,Q} \cdot x^i - \tilde{q}^{i,Q} \cdot x\big) \\
&= p^i \cdot x^i - \nabla U(x^i) \cdot x - \tilde{q}^{i,Q} \cdot x^i + \tilde{q}^{i,Q} \cdot x \\
&\geq p^i \cdot x^i - \tilde{q}^{i,Q} \cdot x - \tilde{q}^{i,Q} \cdot x^i + \tilde{q}^{i,Q} \cdot x \\
&= (p^i - \tilde{q}^{i,Q}) \cdot x^i \\
&\geq 0 \,.
\end{aligned}$$

The first inequality follows form the concavity of f, the second from $\nabla U(x^i) \leq \tilde{q}^{i,Q}$ and $x \geq \mathbf{0}$, and the last one from $p^i \geq \tilde{q}^{i,Q}$ and $x^i > \mathbf{0}$. We conclude that $\Gamma^Q(\mathcal{D})$ is quasilinear rationalizable by $U(x) + y$.

To see that $\tilde{q}^{i,Q} \cdot x^i = 1$, note that since $\Gamma^Q(\mathcal{D})$ is quasilinear rationalizable by $U(x)+y$, the first order conditions imply $\nabla U(x^i) = \tilde{q}^{i,Q} - \tilde{\mu}^{i,Q}$, where $\tilde{\mu}^{i,Q}$ is the Lagrange multiplier associated with the nonnegativity condition $x \geq \mathbf{0}$ in the optimization problem from price $\tilde{\mu}^{i,Q}$. The complementary slackness condition $\tilde{\mu}^{i,Q} \cdot x^i = 0$ implies $\nabla U(x^i) \cdot x^i = \tilde{q}^{i,Q} \cdot x^i$. Since from the first order conditions of the maximization problem in $\mathcal{D}$ we have $\nabla U(x^i) \cdot x^i = p^i \cdot x^i = 1$, we conclude that $\tilde{q}^{i,Q} \cdot x^i = 1$. □

Proof that Qu-1) implies Qu-2). Suppose $\mathcal{D}$ is quasilinear rationalizable by $U(x) + y$ and U is differentiable; then by Lemma 16 $\Gamma^Q(\mathcal{D})$ is also quasilinear rationalizable by $U(x) + y$. Applying the same argument iteratively, we conclude that $\mathcal{D}^Q_\wedge$, which is a finite composition of Γ^Q onto $\mathcal{D}$, is also quasilinear rationalizable by $U(x) + y$. By Theorem 2.2 in Brown and Calsamiglia (2007), $\mathcal{D}^Q_\wedge$ is cyclically monotone. □

H.2 Qu-2) implies Qu-3)

Lemma 17. *Suppose $\mathcal{D}^Q_\wedge$ is cyclically monotone. Then $x^i \approx^Q x^j$ if, and only if, $(q^{i,Q} \cdot x^j - 1) + (q^{j,Q} x^i - 1) = 0$ whenever $x^i \approx^Q x^j$.*

Proof. Necessity follows by definition of $\approx^Q$, since if $q^{i,Q} \cdot (x^j - x^i) + q^{j,Q} \cdot (x^i - x^j) = 0$ then (j, i) is a Q0-sequence. For sufficiency suppose $\mathcal{D}^Q_\wedge$ is cyclically monotone and $x^i \approx^Q x^j$. Since $x^i \approx^Q x^j$, by definition of $\mathcal{D}^Q_\wedge$ we have both $q^{i,Q} \leq q^{j,Q}$ and $q^{j,Q} \leq q^{i,Q}$, therefore $q^{i,Q} = q^{j,Q}$. From Lemma 16 we have that, denoting $\Gamma^Q(\mathcal{D}) - \left(\tilde{q}^{i,Q}, x^i\right)_{i \in [N]}$, $\tilde{q}^{i,Q} \cdot x^i = p^i \cdot x^i = 1$. Applying the same argument iteratively, since $\mathcal{D}^Q_\wedge$ is a finite iteration of $\Gamma^Q()$ on $\mathcal{D}$, we conclude that $q^{i,H} \cdot x^i = q^{j,H} \cdot x^j = 1$. Thus

$$(q^{i,Q} \cdot x^j - 1) + (q^{j,Q} x^i - 1) = (q^{j,Q} \cdot x^j - 1) + (q^{i,Q} x^i - 1) = 0 .$$

□

Lemma 18. *Let $\mathcal{Z}(i)$ be the set of all finite sequences of observations $(m_\ell)_{\ell \in [L]}$ satisfying $x^{m_L} = x^i$. Define*

$$v^i = \min_{\mathcal{Z}(i)} (q^{m_1,Q} \cdot x^{m_2} - 1) + (q^{m_2,Q} \cdot x^{m_3} - 1) + \ldots + (q^{m_{L-1},Q} \cdot x^i - 1) .$$

If $\mathcal{D}^Q_\wedge$ is cyclically monotone then $v^i \geq v^j + 1 - q^{i,Q} \cdot x^j$ for every $i, j \in [N]$, $v^i = v^j + 1 - q^{i,Q} \cdot x^j$ whenever $x^i \approx^Q x^j$, and there is $i' \in [N]$ such that $v^m > v^j + 1 - q^{m,Q} \cdot x^j$ whenever $x^{i'} \approx^Q x^j$ and $x^{i'} \not\approx^Q x^m$.

Proof. That v^i are well defined, is assured by cyclical monotonicity of $\mathcal{D}^Q_\wedge$, as whenever an observation is repeated in the sequence, removing the cycle cannot increase the value. Hence there is a minimizer that has no cycles, and since the number of sequences with no cycles is finite, the minimum exists. Let $(x^{m_\ell})_{\ell \in [L]} \in \mathcal{Z}(j)$ such that $v^j = (q^{m_1,Q} \cdot x^{m_2} - 1) + (q^{m_2,Q} \cdot x^{m_3} - 1) + \ldots + (q^{m_{L-1},Q} \cdot x^j - 1)$. Then by definition of v_i we have

$$\begin{aligned} v^j + 1 - q^{i,Q} \cdot x^j &= (q^{m_1,Q} \cdot x^{m_2} - 1) + (q^{m_2,Q} \cdot x^{m_3} - 1) + \ldots + (q^{m_{L-1},Q} \cdot x^j - 1) + (1 - q^{i,Q} \cdot x^j) \\ &\geq v^i . \end{aligned}$$

Take $x^i \approx^Q x^j$. By Lemma 17 we have $1 - q^{i,Q} \cdot x^j = -(1 - q^{j,Q} \cdot x^i)$. As $v^i \geq v^j + 1 - q^{i,Q} \cdot x^j$ and $v^j \geq v^i + 1 - q^{j,Q} \cdot x^i$ we have

$$\begin{aligned} v^i &\geq v^j + (1 - q^{i,Q} \cdot x^j) \\ &= v^j - (1 - q^{j,Q} \cdot x^i) \end{aligned}$$

$$\geq v^i\,.$$

Hence all the inequalities have to be equalities, and in particular $v^i = v^j + (1 - q^{i,Q} \cdot x^j)$.

Finally, towards a contradiction suppose for every $i \in [N]$ there are $j, m \in [N]$ such that $x^i \approx^Q x^j$, $x^i \not\approx x^m$, and $v^m = v^j + 1 - q^{m,Q} \cdot x^j$. Since $x^i \approx^Q x^j$ we have

$$v^m = v^j + 1 - q^{m,Q} \cdot x^j = v^i + 1 - q^{i,Q} \cdot x^j + 1 - q^{m,Q} \cdot x^j$$

Thus we can construct an infinite sequence $(x^{m_\ell})_{\ell=1}^{\infty}$ such that for all ℓ odd we have $x^{m_\ell} \approx^Q x^{m_{\ell+1}}$ and $x^{m_\ell} \not\approx^Q x^{m_{\ell+2}}$. Furthemore

$$\begin{aligned} v^{m_1} &= v^{m_2} + (1 - q^{m_1,Q} \cdot x^{m_2}) \\ &= v^{m_3} + (1 - q^{m_2,Q} \cdot x^{m_3}) + (1 - q^{m_1,Q} \cdot x^{m_2}) \\ &\vdots \\ &= v^{m_\ell} + \sum_{r=1}^{\ell-1} (1 - q^{m_r,Q} \cdot x^{m_{r+1}}) \\ &\vdots \end{aligned}$$

As the sequence is infinite, there are $\ell', \ell'' \in \mathbb{N}$ such that $\ell' + 3 < \ell''$ and $m_{\ell'} = m_{\ell''}$. We have

$$\begin{aligned} v^{m_{\ell'}} + \sum_{\ell=1}^{\ell'-1} (1 - q^{m_\ell,Q} \cdot x^{m_{r+1}}) &= v^{m_{\ell''}} + \sum_{\ell=1}^{\ell''-1} (1 - q^{m_\ell,Q} \cdot x^{m_{\ell+1}}) \\ &= v^{m_{\ell''}} + \sum_{\ell=1}^{\ell'-1} (1 - q^{m_\ell,Q} \cdot x^{m_{\ell+1}}) + \sum_{\ell=\ell'}^{\ell''-1} (1 - q^{m_\ell,Q} \cdot x^{m_{\ell+1}}) \end{aligned}$$

Since $v^{m_{\ell'}} = v^{m_{\ell''}}$ and $x^{m_{\ell'}} = x^{m_{\ell''}}$

$$0 = (1 - q^{m_{\ell''-1},Q} \cdot x^{m_{\ell'}}) + \sum_{\ell=\ell'}^{\ell''-2} (1 - q^{m_\ell,Q} \cdot x^{m_{\ell+1}})$$

Since $x^{m_{\ell'}}$, $x^{m_{\ell'+1}}$, $x^{m_{\ell'+2}}$ and $x^{m_{\ell'+3}}$ are in the same Q0-sequence, we have $x^{m_{\ell'}} \approx^Q x^{m_{\ell'+2}}$ and $x^{m_{\ell'+1}} \approx^Q x^{m_{\ell'+3}}$. This contradicts the fact that $x^{m_\ell} \not\approx^Q x^{m_{\ell+2}}$ whenever ℓ is odd. $\square$

Lemma 19. *If $\mathcal{D}^Q_\wedge$ is cyclically monotone, then there are numbers $u^i \in \mathbb{R}$ such that $u^i > u^j + 1 - q^{i,Q} \cdot x^j$ whenever $x^i \not\approx^Q x^j$, and $u^i = u^j + 1 - q^{i,Q} \cdot x^j$ whenever $x^i \approx^Q x^j$.*

Proof. We proceed by induction on the numbers of observations. If $N = 1$ then $u^1 = 1$ satisfies the conditions.

Suppose the conditions hold for any data set comprised of $N-1$ or less observations, and take $\mathcal{D}$ comprised of N observations. Take the numbers v^i defined in Lemma 18 and $i' \in [N]$ such that $u^m > u^i + 1 - q^{m,Q} \cdot x^i$ whenever $x^{i'} \approx^Q x^i$ and $x^{i'} \approx^Q x^m$. Denote $E = \{i \in [N] : x^i \approx^Q x^{i'}\}$ and $D = [N] \setminus E$.

- If $D = \emptyset$ then the condition is assured by Lemma 18.
- If $D \neq \emptyset$, then $(q^{i,Q}, x^i)_{i \in D}$ is a data set comprised of $N-1$ or less observations. By induction hypothesis there are numbers $\tilde{u}^m$ for $m \in D$ such that

$$u^m = u^{m'} + 1 - q^{m,Q} \cdot x^{m'} \qquad \text{for all } m, m' \in D\,.$$

the conditions hold. Take the numbers v^i defined in Lemma 18 and define

$$\alpha = \min_{\substack{i \in E \\ m \in D}} v^m - v^j + q^{m,Q} \cdot x^j - 1$$

$$u^i = v^i + \frac{\alpha}{2} \qquad \text{for all } i \in E\,.$$

Hence

$$u^i > v^m + 1 - q^{i,Q} \cdot x^m \qquad \text{for all } i \in E \text{ and } m \in D$$

$$v^m > u^i + 1 - q^{m,Q} \cdot x^i \qquad \text{for all } i \in E \text{ and } m \in D$$

$$u^i = u^j + q^{i,Q} \cdot x^j \qquad \text{for all } i, j \in E\,.$$

Take a sequence $\beta^n \to 1$, where $\beta^n \in (0,1)$ for all n, and for every $m \in D$ define

$$w^m(n) = \beta^n\, v^m + (1-\beta^n)\tilde{u}^m\,.$$

Take any $m, m' \in D$ and $n \in \mathbb{N}$. If $x^m \approx^Q x^{m'}$ then

$$\begin{aligned} w^m(n) &= \beta^n\, v^m + (1-\beta^n)\tilde{u}^m \\ &= \beta^n(v^{m'} + 1 - q^{m,Q} \cdot x^{m'}) + (1-\beta^n)(\tilde{u}^{m'} + 1 - q^{m,Q} \cdot x^{m'}) \\ &= w^{m'}(n) + 1 - q^{m,Q} \cdot x^{m'} \end{aligned}$$

and if $x^m \not\approx^Q x^{m'}$ then

$$\begin{aligned} w^m(n) &= \beta^n\, v^m + (1-\beta^n)\tilde{u}^m \\ &> \beta^n(v^{m'} + 1 - q^{m,Q} \cdot x^{m'}) + (1-\beta^n)(\tilde{u}^{m'} + 1 - q^{m,Q} \cdot x^{m'}) \end{aligned}$$

$$= w^{m'}(n) + 1 - q^{m,Q} \cdot x^{m'}$$

Since $w^m(n) \to v^m$ for every $m \in D$ for n_0 large enough we have

$$\begin{aligned} u^i &> w^m(n_0) + 1 - q^{i,Q} \cdot x^m \\ w^m(n_0) &> u^i + 1 - q^{m,Q} \cdot x^i \end{aligned} \qquad \text{for all } i \in E \text{ and } m \in D\,.$$

Setting $u^m = w^m(n_0)$ for every $m \in D$ assures that the numbers u^i satisfy the desired properties.

□

Proof that Qu-2) implies Qu-3). By Lemma 19 there are numbers u^i satisfying $u^i = u^j + 1 - q^{i,Q} \cdot x^j$ whenever $x^i \approx^Q x^j$ and $u^i > u^j + 1 - q^{i,Q} \cdot x^j$ whenever $x^i \not\approx^Q x^j$. Define

$$\mu^i = p^i - q^{i,Q}\,.$$

Whenever $x^i \not\approx^Q x^j$ we have $u^i > u^j + 1 - q^{i,Q} \cdot x^j = u^j + 1 - p^i \cdot x^j + \mu^i \cdot x^j$; hence (Q0) holds. Similarly, whenever $x^i \approx^Q x^j$ we have $u^i = u^j + 1 - q^{i,Q} \cdot x^j = u^j + 1 - p^i \cdot x^j + \mu^i \cdot x^j$, and (Q2) holds.

Whenever $x^i \approx^Q x^j$, Lemma 17 implies $(q^{i,Q} \cdot x^j - 1) + (q^{j,Q} \cdot x^j - 1) = 0$. Since by definition of $\mathcal{D}^Q_\wedge$ we have $q^{i,Q} \le q^{j,Q}$ and $q^{j,Q} \le q^{i,Q}$, therefore $q^{i,Q} = q^{j,Q}$. Hence $p^i - \mu^i = q^{i,Q} = q^{j,Q} = p^j - \mu^j$, and (Q3) holds.

Since $q^{i,Q} \gg \mathbf{0}$ we have $p^i - \mu^i = q^{i,Q} \gg \mathbf{0}$, i.e., (Q4) holds. Finally since $D^Q_\wedge$ is cyclically monotone it satisfies GARP, which implies $q^i \cdot x^i = 1$, and therefore $\mu^i \cdot x^i = p^i \cdot x^i - q^i \cdot x^i = 0$, hence (Q5) holds.

□

H.3 Qu-3) implies Qu-4)

Proof that Qu-3) implies Qu-4). Suppose there are numbers $u^i \in \mathbb{R}$ and $\mu^i \geq \mathbf{0}$ such that (Q0)-(Q5) hold. Define the functions $\phi^i : \mathbb{R}^K \to \mathbb{R}$ and $V : \mathbb{R}^K \to \mathbb{R}$ by

$$\begin{aligned} \phi^i(x) &= u^i + p^i \cdot x - 1 - \mu^i \cdot x \\ V(x) &= \min_{i \in [N]} \phi^i(x) \end{aligned}$$

By an argument similar to the proof in Appendix F.3 the function V is continuous, concave, and strictly increasing, if $x^i \approx^Q x^j$ then $\phi^i(x) = \phi^j(x)$ for all x, and there is $\eta > 0$ such that

$V(x^i - \xi) = \phi^i(x^i - \xi)$ whenever $\xi \in B(\eta)$. Use such η in the definition of $\rho_\eta(x)$ in (5) and define $\tilde{U} : \mathbb{R}^K \to \mathbb{R}$

$$\tilde{U}(x) = (V \star \rho_\eta)(x) \,.$$

$\tilde{U}$ is strictly increasing, concave, and infinitely differentiable. Furthermore, from an argument similar to (28) we have $\tilde{U}(x^i) = u^i$ for all $i \in [N]$.

Let U be the restriction of $\tilde{U}$ to $\mathbb{R}^K_+$. U is strictly increasing, concave, and infinitely differentiable. Finally, for every $x \in \mathbb{R}^K_+$ we have

$$\begin{aligned}
U(x) &= \int_{B(\eta)} V(x-\xi)\rho_\eta(\xi)\, d\xi \\
&\leq \int_{B(\eta)} \phi^i(x-\xi)\rho_\eta(\xi)\, d\xi \\
&= [u^i + p^i \cdot x - 1 - \mu^i \cdot x] \int_{B(\eta)} \rho_\eta(\xi)\, d\xi - (\lambda^i p^i - \mu^i) \cdot \int_{B(\eta)} \xi \rho_\eta(\xi)\, d\xi \\
&= u^i + p^i \cdot x - 1 - \mu^i \cdot x \\
&\leq u^i - p^i \cdot x^i + p^i \cdot x \\
&= U(x^i) - p^i \cdot x^i + p^i \cdot x \,.
\end{aligned}$$

The second line follows from $i \in [N]$; the third one splits terms; the fourth one from (7) and (8); the fifth one from $\mu^i \cdot x \geq 0$ and $p^i \cdot x^i = 1$; and the last one from $U(x^i) = u^i$. We conclude that $\mathcal{D}$ is quasilinear rationalizable by $U(x) + y$. □