Online Appendix

to

Revealed preferences over risk and uncertainty

Matthew Polisson, John K.-H. Quah, and Ludovic Renou

This Online Appendix consists of the following sections.

(A1) We discuss the relationship between Theorem 1, Afriat's Theorem, and Fishburn's condition.

(A2) We extend the analysis of the rank dependent utility (RDU) model (Quiggin, 1982) to an arbitrary number of states.

(A3) This contains further applications of the GRID method which are not presented in the main text; in particular, we cover the choice acclimating personal equilibrium (CPE) model (Kőszegi and Rabin, 2007), the maxmin expected utility (MEU) model (Gilboa and Schmeidler, 1989), the variational preference (VP) model (Maccheroni, Marinacci, and Rustichini, 2006), a model with budget-dependent reference points, and a model of intertemporal consumption.

(A4) This extends the results on EU- and RDU-rationalizability with concave Bernoulli functions discussed in the main text.

(A5) We describe and explain the GARP and F-GARP tests, in their generalized form needed for calculating Afriat's or Varian's efficiency index.

(A6) This contains more analysis of the data from Choi *et al.* (2007).

(A7) More analysis of the data from Choi *et al.* (2014).

(A8) More analysis of the data from Halevy, Persitz, and Zrill (2018), including results on Varian's index.

(A9) Algorithms for calculating Varian's index for the locally nonsatiated, stochastically monotone, and expected utility models.

A1. Theorem 1, Afriat's Theorem, and Fishburn's condition

A1.1 Afriat's Theorem and Theorem 1

Returning to the context of Section I, we consider a data set of the form $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$, where B^t is a compact subset of $\mathbb{R}^{\bar{s}}_+$ and $\mathbf{x}^t \in B^t$. We assume that B^t is downward comprehensive, so $B^t = \underline{B}^t$. Afriat's (1967) Theorem characterizes data sets that are rationalizable by locally nonsatiated utility (LNU) functions. A utility function $U : \mathbb{R}^{\bar{s}}_+ \to \mathbb{R}$ is *locally nonsatiated* if at every open neighborhood N of $\mathbf{x} \in \mathbb{R}^{\bar{s}}_+$, there is $\mathbf{y} \in N$ such that $U(\mathbf{y}) > U(\mathbf{x})$. We present here a version of Afriat's Theorem in this environment. Note that our environment is actually more general than that of Afriat's (original) result since Afriat only considered data sets where B^t are linear budget sets. However, it is known that the result could be generalized to data sets with nonlinear budget constraints (see Forges and Minelli (2009) and Nishimura, Ok, and Quah (2017)), and we shall also refer to this result as Afriat's Theorem.[1]

Let $\mathcal{D} = \{\mathbf{x}^t : t = 1, 2, \ldots, T\}$; in other words, $\mathcal{D}$ consists of those bundles that were chosen by the subject at some observation in the data set. For bundles $\mathbf{x}^t$ and $\mathbf{x}^{t'}$ in $\mathcal{D}$, $\mathbf{x}^t$ is said to be *revealed preferred to* $\mathbf{x}^{t'}$ (we denote this by $\mathbf{x}^t \succcurlyeq^* \mathbf{x}^{t'}$) if $\mathbf{x}^{t'} \in B^t$;[2] $\mathbf{x}^t$ is said to be *strictly revealed preferred to* $\mathbf{x}^{t'}$ (and we denote this by $\mathbf{x}^t \succ^* \mathbf{x}^{t'}$) if $\mathbf{x}^{t'} \in B^t \backslash \partial B^t$. $\mathcal{O}$ obeys the Generalized Axiom of Revealed Preference (GARP) if, whenever there are observations $(\mathbf{p}^{t_i}, \mathbf{x}^{t_i})$ (for $i = 1, 2, \ldots, n$) in $\mathcal{O}$ satisfying

$$\mathbf{x}^{t_1} \succcurlyeq^* \mathbf{x}^{t_2},\ \mathbf{x}^{t_2} \succcurlyeq^* \mathbf{x}^{t_3},\ \ldots\ ,\ \mathbf{x}^{t_{n-1}} \succcurlyeq^* \mathbf{x}^{t_n}, \text{ and } \mathbf{x}^{t_n} \succcurlyeq^* \mathbf{x}^{t_1}, \tag{a.1}$$

then we cannot replace $\succcurlyeq^*$ with $\succ^*$ anywhere in this chain; in other words, while there can be revealed preference cycles in $\mathcal{O}$, they cannot contain a strict revealed preference. It is straightforward to show that if $\mathcal{O}$ is obtained from a subject maximizing a locally nonsatiated utility (LNU) function, then it must obey GARP. Less trivially, the converse is also true: *if $\mathcal{O}$ obeys GARP then there is a continuous and strictly increasing (in particular, locally*

[1] In Section A5.1, we present a closely related generalization of Afriat's Theorem to modified data sets.

[2] Our terminology differs a little from the standard, which refers to $\succcurlyeq^*$ as the *direct revealed preference* relation and uses *revealed preference* to refer to the transitive closure of this relation. Since our exposition avoids any discussion of the transitive closure, we have adopted the simpler terminology here.

nonsatiated) utility function that rationalizes $\mathcal{O}$ (see Forges and Minelli (2009) or Nishimura, Ok, and Quah (2017)).

In the original version of the theorem due to Afriat, the starting point is a classical data set $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^T$, in which case $\mathbf{x}^t$ is revealed preferred to $\mathbf{x}^{t'}$ if $\mathbf{p}^t \cdot \mathbf{x}^{t'} \leqslant \mathbf{p}^t \cdot \mathbf{x}^t$ (so $\mathbf{x}^{t'} \in B^t = \mathscr{B}(\mathbf{p}^t, \mathbf{p}^t \cdot \mathbf{x}^t)$), and $\mathbf{x}^t$ is strictly revealed preferred to $\mathbf{x}^{t'}$ if $\mathbf{p}^t \cdot \mathbf{x}^{t'} < \mathbf{p}^t \cdot \mathbf{x}^t$. When $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^T$ obeys GARP, it is possible to rationalize $\mathcal{O}$ with a utility function that is strictly increasing, continuous, and concave. The additional property of concavity is sensitive to the linearity of the budget sets; when budget sets are not linear, GARP does not guarantee rationalization with a concave utility function, or even a convex preference.

To relate Afriat's Theorem more closely with Theorem 1, note that the following is easy to show: GARP is equivalent to the existence of a function $\bar{U} : \mathcal{D} \to \mathbb{R}$ that is *strictly increasing* that obeys the following revealed preference conditions:

$$\bar{U}(\mathbf{x}^t) \geqslant \bar{U}(\mathbf{x}) \text{ for all } \mathbf{x} \in B^t \cap \mathcal{D}, \text{ and} \tag{a.2}$$

$$\bar{U}(\mathbf{x}^t) > \bar{U}(\mathbf{x}) \text{ for all } \mathbf{x} \in (B^t \backslash \partial B^t) \cap \mathcal{D}. \tag{a.3}$$

Therefore, we could state Afriat's Theorem in the following manner: $\mathcal{O}$ *is rationalizable by a strictly increasing and continuous utility function if (and only if) there is a strictly increasing function* $\bar{U} : \mathcal{D} \to \mathbb{R}$ *such that (a.2) and (a.3) hold.* Stating Afriat's Theorem in this way highlights its similarity with, and its difference from, Theorem 1, in the case where the state probabilities are held fixed across observations. Clearly, the conditions (a.2) and (a.3) are analogous to (6) and (7) in Theorem 1; in both results, it suffices to find utilities on a particular finite subset of the true consumption space $\mathbb{R}_+^{\bar{s}}$, such that the chosen bundle $\mathbf{x}^t$ is superior to other bundles in B^t *that are within that subset.* But the results differ in two respects: (i) in the case of Afriat's Theorem, one is required to find any increasing function $\bar{U}$, whereas in Theorem 1, the function to be found has the form $\phi(\bar{\mathbf{u}}(\cdot))$ for some strictly increasing $\bar{u} : \mathcal{X} \to \mathbb{R}_+$; (ii) in the case of Afriat's Theorem, it suffices to compare $\mathbf{x}^t$ with feasible bundles in the finite set $\mathcal{D}$, whereas in Theorem 1, the comparison is made with feasible bundles in the finite grid $\mathcal{G}$.

Note that Theorem 1 cannot be improved by requiring (6) and (7) to hold only when comparing $\mathbf{x}^t$ with feasible bundles in the smaller set $\mathcal{D}$ (rather than $\mathcal{G}$). Indeed, consider the case of the single-observation data set in Example 1 (Section I.B of the main text). Conditions (8) and (9) will be satisfied trivially, since $\{\mathbf{x}^1\} = \{(1,2)\} = \mathcal{D}$, but that single observation is not EU-rationalizable.

A1.2 EU-rationalizability, Fishburn's condition, and GARP

We have shown that a data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is EU-rationalizable with probability weights $\{\boldsymbol{\pi}^t\}_{t=1}^T$ if and only if there is a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ obeying conditions (8) and (9). Conditions (8) and (9) generate a finite list of preference pairs between some chosen bundle $\mathbf{x}^t$ and another bundle $\mathbf{x}$ in $\underline{B}^t \cap \mathcal{G}$ or $\left(\underline{B}^t \backslash \partial B^t\right) \cap \mathcal{G}$. The strict increasing condition on $\bar{u}$ can also be reformulated as saying that the bundle $(r, r, \ldots, r)$ is strictly preferred to $(r', r', \ldots, r')$ whenever $r > r'$, for r, $r' \in \mathcal{X}$. We gather these together in a list $\{(\mathbf{a}^j, \mathbf{b}^j)\}_{j=1}^M$, where for all $j \leqslant N$ (with $N < M$), $\mathbf{a}^j$ is weakly preferred to $\mathbf{b}^j$ (so the pairs are drawn from (8)) and for $j > N$, $\mathbf{a}^j$ is strictly preferred to $\mathbf{b}^j$ (so the pairs are drawn from (9) and the strict increasing condition on $\bar{u}$). Each $\mathbf{a}^j$ (similarly $\mathbf{b}^j$) specifies the outcome in each state and the probability of that state. We can write $\mathbf{a}^j$ in its lottery form $g(\mathbf{a}^j)$, which is a vector with $|\mathcal{X}|$ entries, with the ith entry giving the probability of ith ranked number in $\mathcal{X}$; similarly, $\mathbf{b}^j$ can be written in its lottery form $g(\mathbf{b}^j)$. For example, in the example given in the Introduction, $\mathcal{X} = \{0, 1, 2, 3, 4, 6\}$ and the two states are equiprobable, so the bundle $(2, 4)$ chosen from B^1 has the lottery form $(0, 0, 1/2, 0, 1/2, 0)$.

We know from Fishburn (1975) that the list $\{(g(\mathbf{a}^j), g(\mathbf{b}^j))\}_{j=1}^M$ is consistent with expected utility (i.e., there is a strictly increasing $\bar{u}$ that solves (8) and (9)) if and only if it satisfies the following condition, which we shall refer to as *Fishburn's condition*:[3] there does not exist λ^j with $\sum_{j=1}^M \lambda^j = 1$, $\lambda^j \geqslant 0$ for all j, and $\lambda^j > 0$ for some $j > N$, such that

$$\sum_{j=1}^{M} \lambda^j g(\mathbf{a}^j) = \sum_{j=1}^{M} \lambda^j g(\mathbf{b}^j). \tag{a.4}$$

This condition is very intuitive: assuming that the agent has a preference over lotteries, the independence axiom says that the lottery $\sum_{j=1}^M \lambda^j g(\mathbf{a}^j)$ must be strictly preferred to

[3] Fishburn's (1975) result characterizes consistency with expected utility for a finite list of lottery preference pairs; it is not about portfolio choice as such.

$\sum_{j=1}^{M} \lambda^j g(\mathbf{b}^j)$, and therefore (a.4) is excluded.[4] Put another way, a violation of Fishburn's condition must imply a violation of the independence axiom.

Combining our characterization of EU-rationalizability with Fishburn's result gives the following: *a data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is EU-rationalizable with probability weights $\{\boldsymbol{\pi}^t\}_{t=1}^T$ if and only if the preference pairs on $\mathcal{G}$ (as revealed by the data) obey Fishburn's condition.*

We now specialize to the case where the probabilities are fixed across observations. In this case, a subject who is maximizing expected utility is maximizing the same expected utility function across observations, and EU-rationalizability is a special case of rationalizability by an LNU function. It follows that the conditions required of $\bar{u}$ for EU-rationalizability must also be stronger than GARP. Equivalently, it must be the case that Fishburn's condition implies GARP. Indeed, suppose there are observations $(\mathbf{p}^{t_i}, \mathbf{x}^{t_i})$ (for $i = 1, 2, \ldots, n$) in $\mathcal{O}$ satisfying (a.1) and we *can* replace $\succcurlyeq^*$ with $\succ^*$ somewhere in this chain, then there will be a violation of Fishburn's condition since

$$\frac{1}{n}[g(\mathbf{x}^{t_1}) - g(\mathbf{x}^{t_2})] + \frac{1}{n}[g(\mathbf{x}^{t_2}) - g(\mathbf{x}^{t_3})] + \cdots + \frac{1}{n}[g(\mathbf{x}^{t_{n-1}}) - g(\mathbf{x}^{t_n})] + \frac{1}{n}[g(\mathbf{x}^{t_n}) - g(\mathbf{x}^{t_1})] = 0.$$

(Note that since the probability of each state is the same across observations, there is no ambiguity in using $g(\mathbf{x}^t)$ to denote the lottey form of a bundle $\mathbf{x}^t$.)

A2. Testing the multi-state RDU model

We have already discussed the two-state case of the RDU model in Section I.D, and have also explained there how this model could be tested, which is the material most directly relevant to our empirical implementation in Section IV. In this section, we describe the RDU model more generally, i.e., in the case where there are more than two states, and we present the corresponding empirical test using the GRID method. Note that this test neither assumes nor requires the Bernoulli function to be concave; RDU-rationalizability with a concave Bernoulli function is covered in Section A4.2.

[4] To be precise, suppose that the agent has a preference over lotteries with prizes in $\mathcal{X}$. The independence axiom says that if lottery $g(\mathbf{a})$ is preferred (strictly preferred) to $g(\mathbf{b})$, then $\alpha g(\mathbf{a}) + (1-\alpha) g(\mathbf{c})$ is preferred (strictly preferred) to $\alpha g(\mathbf{b}) + (1-\alpha) g(\mathbf{c})$, where $g(\mathbf{c})$ is another lottery and $\alpha \in [0, 1]$. Repeated application of this property and the transitivity of the preference will guarantee that $\sum_{j=1}^{M} \lambda^j g(\mathbf{a}^j)$ is strictly preferred to $\sum_{j=1}^{M} \lambda^j g(\mathbf{b}^j)$.

We consider a setting where, at every observation, the probability of state s (for $s = 1, 2, \ldots, \bar{s}$) is $\pi_s > 0$.[5] Given a contingent consumption bundle $\mathbf{x} \in \mathbb{R}^{\bar{s}}_+$, we can rank the entries of $\mathbf{x}$ from the smallest to the largest, with any ties broken by the rank of the state. We denote the rank of x_s in $\mathbf{x}$ by $r(\mathbf{x}, s)$. For example, if there are five states and $\mathbf{x} = (1, 4, 4, 3, 5)$, we have $r(\mathbf{x}, 1) = 1$, $r(\mathbf{x}, 2) = 3$, $r(\mathbf{x}, 3) = 4$, $r(\mathbf{x}, 4) = 2$, and $r(\mathbf{x}, 5) = 5$. A rank dependent expected utility function gives to the bundle $\mathbf{x}$ the utility

$$V(\mathbf{x}) = \sum_{s=1}^{\bar{s}} \delta(\mathbf{x}, s) u(x_s), \tag{a.5}$$

where $u : \mathbb{R}_+ \to \mathbb{R}_+$ is a Bernoulli function,

$$\delta(\mathbf{x}, s) = \rho\left(\sum\nolimits_{\{s' : r(\mathbf{x}, s') \leqslant r(\mathbf{x}, s)\}} \pi_{s'}\right) - \rho\left(\sum\nolimits_{\{s' : r(\mathbf{x}, s') < r(\mathbf{x}, s)\}} \pi_{s'}\right), \tag{a.6}$$

and $\rho : [0, 1] \to [0, 1]$ is a continuous and strictly increasing function, with $\rho(0) = 0$ and $\rho(1) = 1$. (If $\{s' : r(\mathbf{x}, s') < r(\mathbf{x}, s)\}$ is empty, we let $\rho\left(\sum_{\{s' : r(\mathbf{x}, s') < r(\mathbf{x}, s)\}} \pi_{s'}\right) = 0$.) The function ρ, which we shall refer to as the *transformation function*, distorts the distribution of the bundle $\mathbf{x}$. An agent who maximizes rank dependent utility behaves like an expected-utility maximizer, except that he gives a weight of $\delta(\mathbf{x}, s)$ to the outcome in state s; note that this weight depends on the objective probability of state s but also on the rank of the outcome in that state. Since u is strictly increasing, $\delta(\mathbf{x}, s) = \delta(\mathbf{u}(\mathbf{x}), s)$ and we can write $V(\mathbf{x}) = \phi(\mathbf{u}(\mathbf{x}))$, where for any vector $\mathbf{u} = (u_1, u_2, \ldots, u_{\bar{s}})$,

$$\phi(\mathbf{u}) = \sum_{s=1}^{\bar{s}} \delta(\mathbf{u}, s) u_s. \tag{a.7}$$

It is straightforward to check that $\phi : \mathbb{R}^{\bar{s}}_+ \to \mathbb{R}$ is a strictly increasing and continuous function. Since V has the form assumed in Theorem 1, we can use that result to devise a test for RDU-rationalizability.

By definition, a data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is RDU-rationalizable if there is a transformation function $\rho : [0, 1] \to [0, 1]$ and a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that V as defined by (a.5) satisfies $V(\mathbf{x}^t) \geqslant V(\mathbf{x})$ for all $\mathbf{x} \in B^t$. The next proposition states the test for RDU-rationalizability; it generalizes to multiple states the test formulated in Section I.D.

[5] To keep the notation light, we confine ourselves to the case where $\boldsymbol{\pi}$ does not vary across observations. There is no conceptual difficulty in allowing for this variation if we choose to do so.

Proposition A.1. *Suppose that $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is RDU-rationalizable with the transformation function ρ and the Bernoulli function u. Let $\mathcal{X}$ be a finite set in $\mathbb{R}_+$ that contains $\mathcal{X}^*$, where the latter is defined by (5) (in the main text) and let*

$$\Gamma = \left\{ r \in \mathbb{R} : r = \sum_{s \in S'} \pi_s \text{ for some } S' \subseteq S = \{1, 2, \ldots, \bar{s}\} \right\}. \tag{a.8}$$

Then the restriction of ρ to Γ, $\bar{\rho} : \Gamma \to [0,1]$, and the restriction of u to $\mathcal{X}$, $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ are strictly increasing functions that satisfy

$$\sum_{s=1}^{\bar{s}} \delta(\mathbf{x}^t, s)\bar{u}(x_s) \geqslant \sum_{s=1}^{\bar{s}} \delta(\mathbf{x}, s)\bar{u}(x_s) \text{ for all } \mathbf{x} \in B^t \cap \mathcal{G} \tag{a.9}$$

and

$$\sum_{s=1}^{\bar{s}} \delta(\mathbf{x}^t, s)\bar{u}(x_s) > \sum_{s=1}^{\bar{s}} \delta(\mathbf{x}, s)\bar{u}(x_s) \text{ for all } \mathbf{x} \in \left(\underline{B}^t \backslash \partial B^t\right) \cap \mathcal{G}. \tag{a.10}$$

Conversely, if for a data set $\mathcal{O}$ there are strictly increasing functions $\bar{\rho} : \Gamma \to [0,1]$ (with $\bar{\rho}(0) = 0$ and $\bar{\rho}(1) = 1$) and $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ such that (a.9) and (a.10) are satisfied (with δ given by (a.6) (with $\bar{\rho}$ taking the place of ρ), then $\mathcal{O}$ is RDU-rationalizable with a transformation function ρ that extends $\bar{\rho}$ and a Bernoulli function u that extends $\bar{u}$.

Proof. The first part of this proposition follows immediately from the definition of RDU-rationalizability. For the converse, suppose there are $\bar{\rho}$ and $\bar{u}$ satisfying conditions (a.9) and (a.10); note that these are simply conditions (6) and (7) (in the main text), specialized to the RDU model, with ϕ given by (a.7). Let $\rho : [0,1] \to [0,1]$ be a transformation function extending $\bar{\rho}$; clearly ρ exists since Γ is finite and $\bar{\rho}$ is strictly increasing. Since (a.9) and (a.10) hold, Theorem 1 guarantees that $\mathcal{O}$ is RDU-rationalizable by the transformation function ρ and some Bernoulli function u extending $\bar{u}$. **QED**

The inequality conditions (a.9) and (a.10) are bilinear in $\{\bar{\rho}(\gamma)\}_{\gamma \in \Gamma}$ and $\{\bar{u}(r)\}_{r \in \mathcal{X}}$. So this result tells us that we can test for RDU-rationalizability by looking for a solution to a finite set of inequalities that are bilinear in a finite set of unknowns and, as we explained in Section I.D, problems of this type are decidable. Note that our treatment of the two-state case in Section I.D is a special case of Proposition A.1, with $\Gamma = \{0, \pi_2, \pi_1, 1\}$, $\rho_1 = \bar{\rho}(\pi_1)$ and $\rho_2 = \bar{\rho}(\pi_2)$.

A3. Further applications of the GRID method

For many models of choice under risk or under uncertainty, the GRID method is a useful approach to test rationalizability. As we explain in the main text (Section I), these tests often involve finding a Bernoulli function u and an aggregator function ϕ belonging to some family Φ (corresponding to a particular model) which together rationalize the data. In the subjective expected utility (SEU) case, the GRID test involves solving a system of inequalities that are bilinear in the utility levels $\{\bar{u}(r)\}_{r\in\mathcal{X}}$ and the subjective probabilities $\{\pi_s\}_{s=1}^{\bar{s}}$ (see Section I.D). Such a formulation seems natural enough in the SEU case; what is worth remarking (and perhaps not obvious *a priori*) is that the *same* pattern holds across many of the common models of choice under risk and under uncertainty: they can be tested by solving a system of inequalities that are bilinear in $\{\bar{u}(r)\}_{r\in\mathcal{X}}$ and a finite set of variables specific to the particular model in question. It is known that bilinear systems are decidable, in the sense that there is an algorithm that can determine in a finite number of steps whether or not a solution exists.

In the main text, we have already explained how the expected utility (EU), disappointment aversion (DA), and rank dependent utility (RDU) models can be tested using the GRID method. In this section, we further illustrate the flexibility of the GRID method by applying it to several prominent models of decision making under risk and under uncertainty (Sections A3.1-A3.4). We also explain how it can used to test models of discounted utility, which in formal terms, are very similar to the EU model and its generalizations (Section A3.5).

A3.1 Choice acclimating personal equilibrium

The choice acclimating personal equilibrium (CPE) model (Kőszegi and Rabin, 2007) (with a piecewise linear gain-loss function) specifies utility as $V(\mathbf{x}) = \phi(\mathbf{u}(\mathbf{x}), \boldsymbol{\pi})$, where

$$\phi((u_1, u_2, \ldots, u_{\bar{s}}), \boldsymbol{\pi}) = \sum_{s=1}^{\bar{s}} \pi_s u_s + \frac{1}{2}(1-\lambda) \sum_{r,s=1}^{\bar{s}} \pi_r \pi_s |u_r - u_s|, \tag{a.11}$$

$\boldsymbol{\pi} = \{\pi_s\}_{s=1}^{\bar{s}}$ are the objective probabilities, and $\lambda \in [0, 2]$ is the coefficient of loss aversion.[6]

[6] Our presentation of CPE follows Masatlioglu and Raymond (2016). The restriction of λ to $[0, 2]$ guarantees that V respects first order stochastic dominance but allows for loss-loving behavior (see Masatlioglu and Raymond (2016)).

We say that a data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is CPE-rationalizable with the probability weights $\boldsymbol{\pi}^t = (\pi_1^t, \pi_2^t, \ldots, \pi_{\bar{s}}^t) \gg 0$ if there is ϕ in the collection Φ_{CPE} of functions of the form (a.11), and a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that, for each t, $\phi(\mathbf{u}(\mathbf{x}^t), \boldsymbol{\pi}^t) \geqslant \phi(\mathbf{u}(\mathbf{x}), \boldsymbol{\pi}^t)$ for all $\mathbf{x} \in B^t$. Applying Theorem 1, *$\mathcal{O}$ is CPE-rationalizable if and only if there is $\lambda \in [0, 2]$ and a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ that solve (6) and (7) (in the main text).* It is notable that, irrespective of the number of states, this test is linear in the remaining variables for any given value of λ. Thus it is relatively straightforward to implement via a collection of linear tests (running over different values of $\lambda \in [0, 2]$).

A3.2 Maxmin expected utility

We again consider a setting where no objective probabilities can be attached to each state. An agent with maxmin expected utility (MEU) preferences (Gilboa and Schmeidler, 1989), evaluates each bundle $\mathbf{x} \in \mathbb{R}_+^{\bar{s}}$ using the formula $V(\mathbf{x}) = \phi(\mathbf{u}(\mathbf{x}))$, where

$$\phi(\mathbf{u}) = \min_{\boldsymbol{\pi} \in \Pi} \left\{ \sum_{s=1}^{\bar{s}} \pi_s u_s \right\}, \qquad \text{(a.12)}$$

where $\Pi \subset \Delta_{++} = \{\boldsymbol{\pi} \in \mathbb{R}_{++}^{\bar{s}} : \sum_{s=1}^{\bar{s}} \pi_s = 1\}$ is nonempty, compact in $\mathbb{R}^{\bar{s}}$, and convex. (Π can be interpreted as a set of probability weights.) Given these restrictions on Π, the minimization problem in (a.12) always has a solution and ϕ is strictly increasing.

A data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is said to be MEU-rationalizable if there is a function ϕ in the collection Φ_{MEU} of functions of the form (a.12), and a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that, for each t, $\phi(\mathbf{u}(\mathbf{x}^t), \boldsymbol{\pi}^t) \geqslant \phi(\mathbf{u}(\mathbf{x}), \boldsymbol{\pi}^t)$ for all $\mathbf{x} \in B^t$. By Theorem 1, this holds if and only if there exist Π and a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ that solve (6) and (7) (in the main text).

We claim that these conditions can be formulated in terms of the solvability of a set of bilinear inequalities. This is easy to see for the two-state case where we may assume, without loss of generality, that there is π_1^* and $\pi_1^{**} \in (0, 1)$ such that $\Pi = \{(\pi_1, 1 - \pi_1) : \pi_1^* \leqslant \pi_1 \leqslant \pi_1^{**}\}$. Then it is clear that $\phi(u_1, u_2) = \pi_1^* u_1 + (1 - \pi_1^*) u_2$ if $u_1 \geqslant u_2$ and $\phi(u_1, u_2) = \pi_1^{**} u_1 + (1 - \pi_1^{**}) u_2$ if $u_1 < u_2$. Consequently, for any $(x_1, x_2) \in \mathcal{G}$, we have $V(x_1, x_2) = \pi_1^* \bar{u}(x_1) + (1 - \pi_1^*) \bar{u}(x_2)$ if $x_1 \geqslant x_2$ and $V(x_1, x_2) = \pi_1^{**} \bar{u}(x_1) + (1 - \pi_1^{**}) \bar{u}(x_2)$ if $x_1 < x_2$ and this is independent of the precise choice of $\bar{u}$. Therefore, *$\mathcal{O}$ is MEU-rationalizable* *if and only if we can find* π_1^* *and* π_1^{**} *in* $(0, 1)$*, with* $\pi_1^* \leqslant \pi_1^{**}$*, and an increasing function* $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ *that solve (6) and (7) (in the main text).* The requirement takes the form of a system of bilinear inequalities that are linear in $\{\bar{u}(r)\}_{r \in \mathcal{X}}$ after conditioning on π_1^* and π_1^{**}.

The result below covers the general case. The test involves solving a system of bilinear inequalities in the variables $\bar{\pi}_s(\mathbf{x})$ (for all s and $\mathbf{x} \in \mathcal{G}$) and $\bar{u}(r)$ (for all $r \in \mathcal{X}$). Note that $\bar{\boldsymbol{\pi}}(\mathbf{x}) = (\bar{\pi}_1(\mathbf{x}), \bar{\pi}_2(\mathbf{x}), \ldots, \bar{\pi}_{\bar{s}}(\mathbf{x}))$ is used to construct the set of priors Π (in (a.12)) and that $\bar{\boldsymbol{\pi}}(\mathbf{x})$ is the distribution in Π that minimizes the expected utility of the bundle $\mathbf{x}$ (see (a.16)).

Proposition A.2. *A data set* $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ *is MEU-rationalizable if and only if there is a function* $\bar{\boldsymbol{\pi}} : \mathcal{G} \to \Delta_{++}$ *and a strictly increasing function* $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ *such that*

$$\bar{\boldsymbol{\pi}}(\mathbf{x}^t) \cdot \bar{\mathbf{u}}(\mathbf{x}^t) \geqslant \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) \;\; \textit{for all } \mathbf{x} \in \mathcal{G} \cap \underline{B}^t, \tag{a.13}$$

$$\bar{\boldsymbol{\pi}}(\mathbf{x}^t) \cdot \bar{\mathbf{u}}(\mathbf{x}^t) > \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) \;\; \textit{for all } \mathbf{x} \in \mathcal{G} \cap (\underline{B}^t \backslash \partial \underline{B}^t), \textit{ and} \tag{a.14}$$

$$\bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) \leqslant \bar{\boldsymbol{\pi}}(\mathbf{x}') \cdot \bar{\mathbf{u}}(\mathbf{x}) \;\; \textit{for all } (\mathbf{x}, \mathbf{x}') \in \mathcal{G} \times \mathcal{G}. \tag{a.15}$$

If these conditions hold, $\mathcal{O}$ *admits an MEU-rationalization where* Π *(in (a.12)) is the convex hull of* $\{\bar{\boldsymbol{\pi}}(\mathbf{x})\}_{\mathbf{x} \in \mathcal{G}}$*, the Bernoulli function* $u : \mathbb{R}_+ \to \mathbb{R}$ *extends* $\bar{u}$*, and*

$$V(\mathbf{x}) = \min_{\boldsymbol{\pi} \in \Pi} \{\boldsymbol{\pi} \cdot \bar{\mathbf{u}}(\mathbf{x})\} = \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) \;\; \textit{for all } \mathbf{x} \in \mathcal{G}. \tag{a.16}$$

Proof: Suppose that $\mathcal{O}$ is rationalizable by ϕ as defined by (a.12). For any $\mathbf{x}$ in the finite grid $\mathcal{G}$, let $\bar{\boldsymbol{\pi}}(\mathbf{x})$ be an element in $\arg\min_{\boldsymbol{\pi} \in \Pi} \boldsymbol{\pi} \cdot \mathbf{u}(\mathbf{x})$ and let $\bar{u}$ be the restriction of u to $\mathcal{X}$. Then it is clear that the conditions (a.13)–(a.15) hold.

Conversely, suppose that there is a function $\bar{\boldsymbol{\pi}}$ and a strictly increasing function $\bar{u}$ obeying the conditions (a.13)–(a.15). Define Π as the convex hull of $\{\bar{\boldsymbol{\pi}}(\mathbf{x}) : \mathbf{x} \in \mathcal{G}\}$; Π is a nonempty and convex subset of Δ_{++} and it is compact in $\mathbb{R}^{\bar{s}}$ since $\mathcal{G}$ is finite. Suppose that there exists $\mathbf{x} \in \mathcal{G}$ and $\boldsymbol{\pi} \in \Pi$ such that $\boldsymbol{\pi} \cdot \bar{\mathbf{u}}(\mathbf{x}) < \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x})$. Since $\boldsymbol{\pi}$ is a convex combination of elements in $\{\bar{\boldsymbol{\pi}}(\mathbf{x}) : \mathbf{x} \in \mathcal{G}\}$, there must exist $\mathbf{x}' \in \mathcal{G}$ such that $\bar{\boldsymbol{\pi}}(\mathbf{x}') \cdot \bar{\mathbf{u}}(\mathbf{x}) < \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x})$, which contradicts (a.15). We conclude that $\bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) = \min_{\boldsymbol{\pi} \in \Pi} \boldsymbol{\pi} \cdot \bar{\mathbf{u}}(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{G}$. We define $\phi : \mathbb{R}_+^{\bar{s}} \to \mathbb{R}$ by $\phi(\mathbf{u}) = \min_{\boldsymbol{\pi} \in \Pi} \boldsymbol{\pi} \cdot \mathbf{u}$. Then the conditions (a.13) and (a.14) are just the conditions (6) and (7) (in the main text), and Theorem 1 guarantees that there is a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ extending $\bar{u}$ such that $\mathcal{O}$ is rationalizable by $V(\mathbf{x}) = \phi(\mathbf{u}(\mathbf{x}))$. **QED**

A3.3 Variational preferences

A popular model of decision making under uncertainty which generalizes maxmin expected utility is variational preferences (VP), introduced by Maccheroni, Marinacci, and Rustichini (2006). In this model, a bundle $\mathbf{x} \in \mathbb{R}_+^{\bar{s}}$ has utility $V(\mathbf{x}) = \phi(\mathbf{u}(\mathbf{x}))$, where

$$\phi(\mathbf{u}) = \min_{\boldsymbol{\pi} \in \Delta_{++}} \{\boldsymbol{\pi} \cdot \mathbf{u} + c(\boldsymbol{\pi})\} \tag{a.17}$$

and $c : \Delta_{++} \to \mathbb{R}_+$ is a continuous and convex function with the following boundary condition: for any sequence $\boldsymbol{\pi}^n \in \Delta_{++}$ tending to $\tilde{\boldsymbol{\pi}}$, with $\tilde{\pi}_s = 0$ for some s, we obtain $c(\boldsymbol{\pi}^n) \to \infty$. This boundary condition, together with the continuity of c, guarantee that there is $\boldsymbol{\pi}^* \in \Delta_{++}$ that solves the problem in (a.17).[7] Therefore, ϕ is well-defined and strictly increasing.

We say that $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is VP-rationalizable if there is a function ϕ in the collection Φ_{VP} of functions of the form (a.17), and a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that, for each t, $\phi(\mathbf{u}(\mathbf{x}^t), \boldsymbol{\pi}^t) \geqslant \phi(\mathbf{u}(\mathbf{x}), \boldsymbol{\pi}^t)$ for all $\mathbf{x} \in B^t$. By Theorem 1, this holds if and only if there exists a function $c : \Delta_{++} \to \mathbb{R}_+$ that is continuous, convex, and has the boundary property, and a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ that together solve (6) and (7) (in the main text), with ϕ defined by (a.17). The following result is a reformulation of this characterization that has a similar flavor to Proposition A.2; crucially, the necessary and sufficient conditions on $\mathcal{O}$ are formulated as a finite set of bilinear inequalities.

Proposition A.3. *A data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is VP-rationalizable if and only if there is a function $\bar{\boldsymbol{\pi}} : \mathcal{G} \to \Delta_{++}$, a function $\bar{c} : \mathcal{G} \to \mathbb{R}_+$, and a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ such that*

$$\bar{\boldsymbol{\pi}}(\mathbf{x}^t) \cdot \bar{\mathbf{u}}(\mathbf{x}^t) + \bar{c}(\mathbf{x}^t) \geqslant \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}) \;\text{ for all } \mathbf{x} \in \mathcal{G} \cap \underline{B}^t, \tag{a.18}$$

$$\bar{\boldsymbol{\pi}}(\mathbf{x}^t) \cdot \bar{\mathbf{u}}(\mathbf{x}^t) + \bar{c}(\mathbf{x}^t) > \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}) \;\text{ for all } \mathbf{x} \in \mathcal{G} \cap (\underline{B}^t \backslash \partial \underline{B}^t), \text{ and} \tag{a.19}$$

$$\bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}) \leqslant \bar{\boldsymbol{\pi}}(\mathbf{x}') \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}') \;\text{ for all } (\mathbf{x}, \mathbf{x}') \in \mathcal{G} \times \mathcal{G}. \tag{a.20}$$

If these conditions hold, then $\mathcal{O}$ can be rationalized by a variational preference V, with ϕ given by (a.17), such that the following holds:

---

[7] Indeed, pick any $\tilde{\boldsymbol{\pi}} \in \Delta_{++}$ and define $S = \{\boldsymbol{\pi} \in \Delta_{++} : \boldsymbol{\pi} \cdot \mathbf{u} + c(\boldsymbol{\pi}) \leqslant \tilde{\boldsymbol{\pi}} \cdot \mathbf{u} + c(\tilde{\boldsymbol{\pi}})\}$. The boundary condition and continuity of c guarantee that S is compact in $\mathbb{R}^{\bar{s}}$ and hence $\arg\min_{\boldsymbol{\pi} \in S}\{\boldsymbol{\pi} \cdot \mathbf{u} + c(\boldsymbol{\pi})\} = \arg\min_{\boldsymbol{\pi} \in \Delta_{++}}\{\boldsymbol{\pi} \cdot \mathbf{u} + c(\boldsymbol{\pi})\}$ is nonempty.

(i) $c : \Delta_{++} \to \mathbb{R}_+$ *satisfies* $c(\bar{\boldsymbol{\pi}}(\mathbf{x})) = \bar{c}(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{G}$*;*

(ii) *the Bernoulli function* $u : \mathbb{R}_+ \to \mathbb{R}$ *satisfies* $\bar{u}(r) = u(r)$ *for all* $r \in \mathcal{X}$*; and*

(iii) $\bar{\boldsymbol{\pi}}(\mathbf{x}) \in \arg\min_{\boldsymbol{\pi} \in \Delta_{++}} \{\boldsymbol{\pi} \cdot \mathbf{u}(\mathbf{x}) + c(\boldsymbol{\pi})\}$, *leading to* $V(\mathbf{x}) = \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x})$, *for all* $\mathbf{x} \in \mathcal{G}$.

Proof: Suppose $\mathcal{O}$ is rationalizable by ϕ as defined by (a.17). Let $\bar{u}$ be the restriction of u to $\mathcal{X}$. For any $\mathbf{x}$ in $\mathcal{G}$, let $\bar{\boldsymbol{\pi}}(\mathbf{x})$ be an element in $\arg\min_{\boldsymbol{\pi} \in \Delta_{++}} \{\boldsymbol{\pi} \cdot \mathbf{u}(\mathbf{x}) + c(\boldsymbol{\pi})\}$, and let $\bar{c}(\mathbf{x}) = c(\bar{\boldsymbol{\pi}}(\mathbf{x}))$. Then it is clear that the conditions (a.18)–(a.20) hold.

Conversely, suppose that there is a strictly increasing function $\bar{u}$ and functions $\bar{\boldsymbol{\pi}}$ and $\bar{c}$ obeying conditions (a.18)–(a.20). For every $\boldsymbol{\pi} \in \Delta_{++}$, define $\tilde{c}(\boldsymbol{\pi}) = \max_{\mathbf{x} \in \mathcal{G}} \{\bar{c}(\mathbf{x}) - (\boldsymbol{\pi} - \bar{\boldsymbol{\pi}}(\mathbf{x})) \cdot \bar{\mathbf{u}}(\mathbf{x})\}$. It follows from (a.20) that $\bar{c}(\mathbf{x}') \geqslant \bar{c}(\mathbf{x}) - (\bar{\boldsymbol{\pi}}(\mathbf{x}') - \bar{\boldsymbol{\pi}}(\mathbf{x})) \cdot \bar{\mathbf{u}}(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{G}$. Therefore, $\tilde{c}(\bar{\boldsymbol{\pi}}(\mathbf{x}')) = \bar{c}(\mathbf{x}')$ for any $\mathbf{x}' \in \mathcal{G}$. The function $\tilde{c}$ is convex and continuous but it need not obey the boundary condition. However, we know there is a function c defined on Δ_{++} that is convex, continuous, obeys the boundary condition, with $c(\boldsymbol{\pi}) \geqslant \tilde{c}(\boldsymbol{\pi})$ for all $\boldsymbol{\pi} \in \Delta_{++}$ and $c(\boldsymbol{\pi}) = \tilde{c}(\boldsymbol{\pi})$ for $\boldsymbol{\pi} \in \{\bar{\boldsymbol{\pi}}(\mathbf{x}) : \mathbf{x} \in \mathcal{G}\}$. We claim that, with c so defined, $\min_{\boldsymbol{\pi} \in \Delta_{++}} \{\boldsymbol{\pi} \cdot \bar{\mathbf{u}}(\mathbf{x}) + c(\boldsymbol{\pi})\} = \boldsymbol{\pi}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{G}$. Indeed, for any $\boldsymbol{\pi} \in \Delta_{++}$,

$$\boldsymbol{\pi} \cdot \bar{\mathbf{u}}(\mathbf{x}) + c(\boldsymbol{\pi}) \geqslant \boldsymbol{\pi} \cdot \bar{\mathbf{u}}(\mathbf{x}) + \tilde{c}(\boldsymbol{\pi}) \geqslant \boldsymbol{\pi} \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}) - (\boldsymbol{\pi} - \bar{\boldsymbol{\pi}}(\mathbf{x})) \cdot \bar{\mathbf{u}}(\mathbf{x}) = \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}).$$

On the other hand, $\bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) + c(\bar{\boldsymbol{\pi}}(\mathbf{x})) = \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) + \bar{c}(\mathbf{x})$, which establishes the claim. We define $\phi : \mathbb{R}^{\bar{s}}_+ \to \mathbb{R}$ by (a.17); then (a.18) and (a.19) are just versions of (6) and (7) (in the main text), and so Theorem 1 guarantees that there is a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ extending $\bar{u}$ such that $\mathcal{O}$ is rationalizable by $V(\mathbf{x}) = \phi(\mathbf{u}(\mathbf{x}))$. **QED**

A3.4 Models with budget-dependent reference points

So far in our discussion we have assumed that the agent has a preference over different contingent outcomes, without being too specific as to what actually constitutes an outcome in the agent's mind. On the other hand, models such as prospect theory have often emphasized the impact of reference points, and *changing* reference points, on decision making. Some of these phenomena can be easily accommodated within our framework.

For example, imagine a portfolio choice experiment where, at observation t, the subject chooses a state contingent bundle $\mathbf{x}^t$ from a constraint set $B^t \in \mathbb{R}_+^{\bar{s}}$, so the data set is $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$. The standard way of thinking about the subject's behavior is to assume that his choice from B^t is governed by a preference defined on $\mathbb{R}_+^{\bar{s}}$, which implicitly means that the situation where he receives nothing from the experiment in every state (formally the vector 0) is the subject's constant reference point. But a researcher may well be interested in whether the subject has a different reference point or multiple reference points that vary with the budget (and perhaps manipulable by the researcher). Most obviously, suppose that the subject has an endowment point $\boldsymbol{\omega}^t \in \mathbb{R}_+^{\bar{s}}$ and a classical budget set $B^t = \{\mathbf{x} \in \mathbb{R}_+^{\bar{s}} : \mathbf{p}^t \cdot \mathbf{x} \leqslant \mathbf{p}^t \cdot \boldsymbol{\omega}^t\}$. In this case, a possible hypothesis is that the subject will evaluate different bundles in B^t based on a utility function defined on the deviation from the endowment; in other words, the endowment is the subject's reference point. Another possible reference point is that bundle in B^t which gives the same payoff in every state.

Whatever it may be, suppose the researcher has a hypothesis about the possible reference point at observation t, which we shall denote by $\mathbf{q}^t \in \mathbb{R}_+^{\bar{s}}$, and that the subject chooses according to some utility function $V : [-K, \infty)^{\bar{s}} \to \mathbb{R}_+$ where $K > 0$ is sufficiently large so that $[-K, \infty)^{\bar{s}} \subset \mathbb{R}^{\bar{s}}$ contains all the possible reference point-dependent outcomes in the data, i.e., the set $\bigcup_{t=1}^T \tilde{B}^t$, where

$$\tilde{B}^t = \{\mathbf{x}' \in \mathbb{R}^{\bar{s}} : \ \mathbf{x}' = \mathbf{x} - \mathbf{q}^t \text{ for some } \mathbf{x} \in B^t\}.$$

Let $\{\phi(\cdot, t)\}_{t=1}^T$ be a collection of functions, where $\phi(\cdot, t) : [-K, \infty)^{\bar{s}} \to \mathbb{R}$ is increasing in all of its arguments. We say that $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is *rationalizable by* $\{\phi(\cdot, t)\}_{t=1}^T$ *and the reference points* $\{\mathbf{q}^t\}_{t=1}^T$ if there exists a Bernoulli function $u : [-K, \infty) \to \mathbb{R}_+$ such that $\phi(\mathbf{u}(\mathbf{x}^t - \mathbf{q}^t), t) \geqslant \phi(\mathbf{u}(\mathbf{x} - \mathbf{q}^t), t)$ for all $\mathbf{x} \in B^t$. This is formally equivalent to saying that the modified data set $\mathcal{O}' = \{(\mathbf{x}^t - \mathbf{q}^t, \tilde{B}^t)\}_{t=1}^T$ is rationalizable by $\{\phi(\cdot, t)\}_{t=1}^T$. Applying Theorem 1, rationalizability holds if and only if there is a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ that obeys (6) and (7) (in the main text), where

$$\mathcal{X} = \{r \in \mathbb{R} : r = x_s^t - q_s^t \text{ for some } t, \ s\} \cup \{-K\}.$$

Therefore, we may test whether $\mathcal{O}$ is rationalizable by expected utility, or by any of the models described so far, in conjunction with budget dependent reference points. Note that a test of rank dependent utility in this context is sufficiently flexible to accommodate phenomena emphasized by cumulative prospect theory (see Tversky and Kahneman (1992)), such as a Bernoulli function $u : [-K, \infty) \to \mathbb{R}$ that is S-shaped (and hence nonconcave) around 0 and probabilities distorted by a weighting function.

A3.5 Models of intertemporal consumption

Models of intertemporal consumption are formally very similar to models of risky or uncertain consumption, so the GRID method could be applied to study these models as well. To give a sense of how our method could be applied, we shall consider a data set collected from a budgetary choice experiment of the type performed by Andreoni and Sprenger (2012).

There is a finite number of observations of each subject. At observation t, the subject is asked to divide a budget of m^t between consumption at date $d(t)$ and a later date $D(t)$, with the interest rate being $r^t > -1$. In formal terms, the budget set at observation t is

$$B^t = \{(x_{d(t)}, x_{D(t)}) \in \mathbb{R}^2_+ : (1 + r^t)x_{d(t)} + x_{D(t)} \leqslant m^t\}. \qquad \text{(a.21)}$$

As in Andreoni and Sprenger (2012) we are interested in checking if the data is consistent with a discounted utility model, possibly with a present bias. In other words, given a data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ we are asking if there is a discount factor $\delta \in (0, 1)$, a present bias coefficient $\beta > 0$, and a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that the choice $\mathbf{x}^t = (x^t_{d(t)}, x^t_{D(t)})$ at observation t maximizes

$$u(x_{d(t)}) + \delta^{D(t)-d(t)} u(x_{D(t)})$$

among all $(x_{d(t)}, x_{D(t)}) \in B^t$, in the case where $d(t) > 0$; in the case where $d(t) = 0$ (so the earlier payment is made at the current date), $\mathbf{x}^t = (x^t_0, x^t_{D(t)})$ should maximize

$$u(x_0) + \beta\delta^{D(t)} u(x_{D(t)})$$

among all $(x_0, x_{D(t)}) \in B^t$. In the event that the data is not exactly consistent with this model, we would like to find u, β and δ that gives the best fit in the sense of maximizing Afriat's efficiency index.

For a *given* β and δ, we can check if there is Bernoulli function that rationalizes the data by using Theorem 1. Simply set

$$\mathcal{X} = \{0\} \cup \{r : r = x^t_{d(t)} \text{ or } r = x^t_{D(t)} \text{ for some observation } t\}$$

and let $\mathcal{G} = \mathcal{X}^2$. Such a Bernoulli function exists if and only if we can find an increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ with the revealed preference conditions (6) and (7) holding between the observed choice $\mathbf{x}^t$ and alternatives in $B^t \cap \mathcal{G}$. If $d(t) > 0$, this requires

$$\bar{u}(x^t_{d(t)}) + \delta^{D(t)-d(t)}\bar{u}(x^t_{D(t)}) \geqslant \bar{u}(a) + \delta^{D(t)-d(t)}\bar{u}(b) \text{ for all } (a,b) \in \underline{B}^t \cap \mathcal{G} \text{ and}$$

$$\bar{u}(x^t_{d(t)}) + \delta^{D(t)-d(t)}\bar{u}(x^t_{D(t)}) > \bar{u}(a) + \delta^{D(t)-d(t)}\bar{u}(b) \text{ for all } (a,b) \in \left(\underline{B}^t \backslash \partial B^t\right) \cap \mathcal{G}.$$

and in the case where $d(t) = 0$, this requires

$$\bar{u}(x^t_0) + \beta\delta^{D(t)}\bar{u}(x^t_{D(t)}) \geqslant \bar{u}(a) + \beta\delta^{D(t)}\bar{u}(b) \text{ for all } (a,b) \in \underline{B}^t \cap \mathcal{G} \text{ and}$$

$$u(x^t_0) + \beta\delta^{D(t)}u(x^t_{D(t)}) > u(a) + \beta\delta^{D(t)}u(b) \text{ for all } (a,b) \in \left(\underline{B}^t \backslash \partial B^t\right) \cap \mathcal{G}.$$[8]

So to test (or estimate) such a model involves two stages: first, we fix (δ, β) and perform a linear test to check if there is an increasing $\bar{u}$ obeying the revealed preference conditions stated above or, failing that, to calculate Afriat's efficiency index; and second, do a search through different values of (β, δ), in order to find a value at which the data set passes exactly, or comes as close to passing as possible (in the sense of maximizing Afriat's efficiency index). Since this only involves searching two parameters (β and δ) for a pass, or a good fit, this is computationally feasible.[9] We could also test the model with concavity imposed on u. In that case, in stage 1, we perform for a given (δ, β), the linear test specified by Proposition 1 to determine the associated Afriat's efficiency index;[10] in second stage, search for the (β, δ) at which the data passes exactly or where the efficiency index is maximized.

A4. More on Concave EU- and RDU-rationalizability

We begin this section with an intuitive example of a classical data set with two observations that is EU-rationalizable but only with a nonconcave Bernoulli function. We then explain how we could test for RDU-rationalizability with a concave Bernoulli function in

[9] Note that the test for the rank dependent utility model we performed in the empirical section of the paper also involves searching through two parameters; in that case, they are the distorted probabilities of the two states (see Section I.D and Section IV in the main text).

[10] This corresponds to an application of the proposition where, at observation t, $\pi^t_1 = 1$ and $\pi^t_2 = \delta^{D(t)-d(t)}$ if $d(t) > 0$ and $\pi^t_2 = \beta\delta^{D(t)}$ if $d(t) = 0$. While the probability of each state π_s is held fixed across observations in the proposition, it is clear that that is not crucial to the result, nor do we really need the weights to add up to 1 (though they must be positive).

the two-state case (again using the GRID method); this is essentially a continuation of the discussion in Section III, which describes the test for concave EU-rationalizability. The final subsection gives the general GRID testing procedure for EU- or RDU-rationalizability with concave Bernoulli functions when there are multiple states.

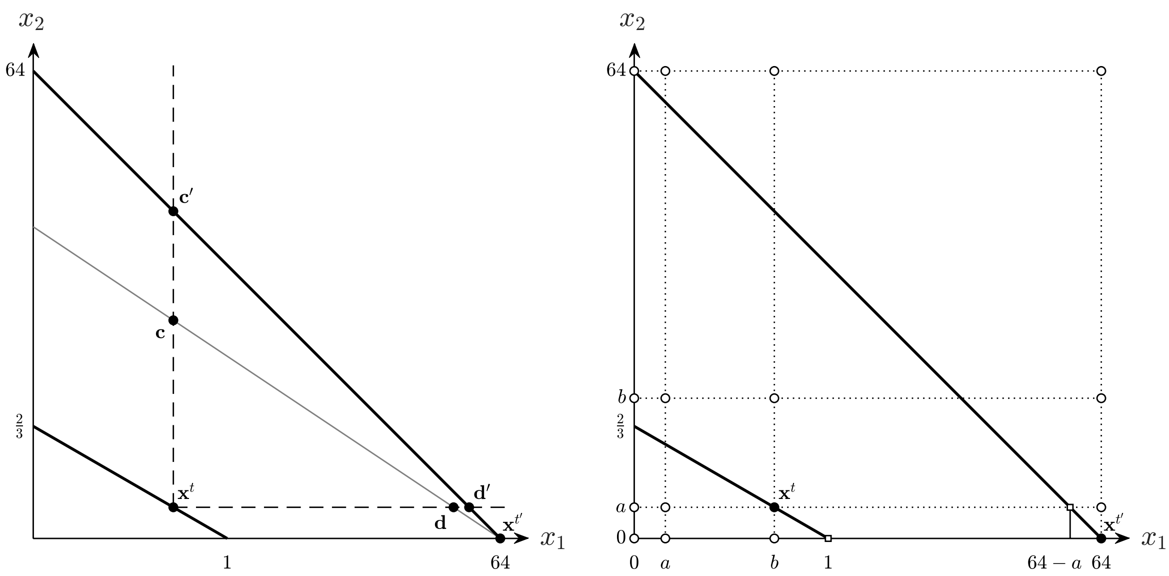
A4.1 Example of EU-rationalizability with a nonconcave Bernoulli function

Suppose an agent maximizes expected utility and has the Bernoulli function $\hat{u}(y) = (y-4)^3$, which is strictly concave for $y < 4$ and strictly convex for $y > 4$. Note that Bernoulli functions of this type are not a novel contrivance: indeed, they were used by Friedman and Savage (1948, Figure 2) to explain why an agent can simultaneously buy insurance and accept risky gambles. We assume that there are two states of the world, which occur with equal probability. At $\mathbf{p}^t = (1, 3/2)$ and with wealth equal to 1, the agent chooses $x_1 \in [0,1]$ to maximize $f(x_1) = (x_1 - 4)^3 + (2(1-x_1)/3 - 4)^3$. Over this range, the Bernoulli function is strictly concave and so is f; one could check that $f'(1) < 0$ so that there is unique interior solution which we denote $\mathbf{x}^t = (b, a)$ (see Figure A.1). (Solving the (quadratic) first order condition gives $\mathbf{x}^t = (b, a) \approx (0.83, 0.11)$.) At the prices $\mathbf{p}^{t'} = (1,1)$ with wealth equal to 64, the agent chooses $x_1 \in [0, 64]$ to maximize $g(x_1) = (x_1 - 4)^3 + (60 - x_1)^3$. It is straightforward to check that g is strictly convex on $[0, 64]$ and it is thus maximized at the two end points $(0, 64)$ and $(64, 0)$.

Now consider a data set with two observations: the bundle $\mathbf{x}^t = (b, a)$ chosen at $\mathbf{p}^t = (1, 3/2)$ and $\mathbf{x}^{t'} = (64, 0)$ chosen at $\mathbf{p}^{t'} = (1, 1)$. We know that this data set is EU-rationalizable by the Bernoulli function $\hat{u}(y) = (y-4)^3$. We also know from Afriat's Theorem (since these two observations satisfy the generalized axiom of revealed preference) that there is an increasing, continuous, *and concave* utility function defined on $\mathbb{R}^2_+$ that rationalizes these observations. However, such a function cannot be of the EU form because this data set is *not* EU-rationalizable with a concave Bernoulli function.

For readers familiar with classical consumer theory, this is clear because an additive and concave utility function (on the consumption space) must generate a normal demand function, while the demand behavior displayed in these observations requires consumption in state 2 to violate normality.[11] Indeed, in Figure A.1a we have performed a Slutsky

[11] By definition, a good is normal if a parallel outward shift in the budget line leads to an increase in the



(a) Data set with two observations

(b) Data set with $\mathcal{N}$ inserted

Figure A.1: EU-rationalizable but not concave EU-rationalizable data set

decomposition by adding an additional budget line through $\mathbf{x}^{t'}$ that is parallel to the one through $\mathbf{x}^t$; demand on this line has to be between $\mathbf{c}$ and $\mathbf{d}$ for normality to hold, which in turn means (given that utility is strictly increasing) that there is a bundle between $\mathbf{c}'$ and $\mathbf{d}'$ that is strictly preferred to $\mathbf{x}^{t'}$, contradicting the optimality of the latter.

More directly, one could check that these two observations will fail the concave EU test stated in Proposition 1. In this case, the two observations are (b, a) and $(64, 0)$, and 64 is also the largest achievable consumption at these two observations, so we can choose $\mathcal{X} = \{0, a, b, 64\}$. The set $\mathcal{N}$ is depicted by the dotted lines in Figure A.1b. The point $(1, 0)$ can be written as $(b + \alpha(64 - b), 0)$ for some $\alpha \in (0, 1)$ and the point $(64 - a, a)$ can be written as $(64 - \beta(64 - b), a)$ for some $\beta \in (0, 1)$, where $\alpha > \beta$ (because the budget line at t' is steeper than the one at t). Suppose to the contrary that there is a strictly increasing $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ with a linear extension $\bar{u}_\ell$ that passes the test. This means in particular, by the optimality condition on (b, a), that

$$\begin{aligned}\bar{u}(b) + \bar{u}(a) = \bar{u}_\ell(b) + \bar{u}_\ell(a) &\geqslant \bar{u}_\ell(1) + \bar{u}_\ell(0) \\ &= \bar{u}(b) + \alpha\left[\bar{u}(64) - \bar{u}(b)\right] + \bar{u}(0),\end{aligned}$$

demand for that good.

which implies that

$$\bar{u}(a) - \bar{u}(0) \geqslant \alpha(\bar{u}(64) - \bar{u}(b)). \tag{a.22}$$

Similarly the optimality of $(64, 0)$ implies that

$$\begin{aligned}\bar{u}(64) + \bar{u}(0) = \bar{u}_\ell(64) + \bar{u}_\ell(0) \geqslant \bar{u}_\ell(64-a) + \bar{u}_\ell(a) \\ = \bar{u}(64) - \beta\left[\bar{u}(64) - \bar{u}(b)\right] + \bar{u}_\ell(a).\end{aligned}$$

Re-arranging this inequality gives

$$\beta(\bar{u}_\ell(64) - \bar{u}_\ell(b)) \geqslant \bar{u}_\ell(a) - \bar{u}_\ell(0) \tag{a.23}$$

The inequalities (a.22) and (a.23) are incompatible since $\alpha > \beta$ and $\bar{u}(64) > \bar{u}(0)$.

A4.2 Concave RDU-rationalizability with two states

In Section III, we explain how we can test for concave EU-rationalizability of a modified data set, i.e., the data set $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t))\}_{t=1}^T$, with

$$\begin{aligned}B^t(e^t) &= \mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t) \cup \{\mathbf{x}^t\} \\ &= \{\mathbf{x} \in \mathbb{R}_+^{\bar{s}} : \mathbf{p}^t \cdot \mathbf{x} \leqslant e^t\mathbf{p}^t \cdot \mathbf{x}^t\} \cup \{\mathbf{x}^t\}.\end{aligned} \tag{a.24}$$

(see (14) in the main text), in the case where there are just two states. We now show how that test can be modified to accommodate concave RDU-rationalizability (in other words, RDU-rationalizability with a concave Bernoulli function), again in the context of two states. The two-state results (for both the EU and RDU models) are just special cases of the results with multiple states which are covered in the next subsection. Nonetheless, we present the two-state case separately because the results are easier to describe and to understand, and also because this is the relevant case for the empirical tests implemented in Section IV.

We assume that the objective probabilities for both states are strictly positive and, with no loss of generality, assume that $\pi_1 \geqslant \pi_2$. As we explain in Section I.D, in this model, the subject chooses a bundle $\mathbf{x} = (x_1, x_2)$ in the budget set to maximize the utility function $V : \mathbb{R}_+^2 \to \mathbb{R}_+$, with ϕ given by (a.7), so

$$V(x_1, x_2) = \phi(u(x_1), u(x_2)) = \begin{cases} \rho_1 u(x_1) + (1-\rho_1)u(x_2) & \text{if } x_1 \leqslant x_2 \\ (1-\rho_2)u(x_1) + \rho_2 u(x_2) & \text{if } x_1 > x_2 \end{cases} \tag{a.25}$$

The constants ρ_1 and ρ_2 obey (11).

Recall the definitions of $\mathcal{X}^{**}$, $\mathcal{X}$, $\bar{u}$, and $\bar{u}$'s linear extension $\bar{u}_\ell$ in Section III. Suppose that $\mathcal{O}(\mathbf{e})$ is RDU-rationalizable by the concave Bernoulli function u, which means that $\mathbf{x}^t = (x_1^t, x_2^t)$ is optimal in the set $B^t(e^t)$. By construction, $\bar{u}_\ell$ is a Bernoulli function satisfying $\bar{u}_\ell(r) = u(r)$ for all $r \in \mathcal{X}$ and it is linear between adjacent values of $\mathcal{X}$; furthermore, the concavity of u guarantees that $\bar{u}_\ell$ is also concave, with $u(r) \geqslant \bar{u}_\ell(r)$ for all $r \in [0, \bar{r}]$. These properties guarantee that $\phi(u(x_1), u(x_2)) \geqslant \phi(\bar{u}_\ell(x_1), \bar{u}_\ell(x_2))$ for any bundle (x_1, x_2) and $\phi(u(x_1^t), u(x_2^t)) \geqslant \phi(\bar{u}_\ell(x_1^t), \bar{u}_\ell(x_2^t))$ at every choice bundle (x_1^t, x_2^t). It follows immediately from this observation that $\mathcal{O}(\mathbf{e})$ is also RDU-rationalized by $\bar{u}_\ell$. In particular, the following must hold:

$$\phi(\bar{u}_\ell(x_1^t), \bar{u}_\ell(x_2^t)) \geqslant \phi(\bar{u}_\ell(m), \bar{u}_\ell(m)) \text{ for } (m, m) \in \partial\mathscr{B}(\mathbf{p}^t, \mathbf{p}^t \cdot \mathbf{x}^t) \tag{a.26}$$

and

$$\phi(\bar{u}_\ell(x_1^t), \bar{u}_\ell(x_2^t)) \geqslant \phi(\bar{u}_\ell(x_1), \bar{u}_\ell(x_2)) \text{ for all } (x_1, x_2) \in \mathcal{N} \cap \partial\mathscr{B}(\mathbf{p}^t, \mathbf{p}^t \cdot \mathbf{x}^t). \tag{a.27}$$

The first condition states that the observed choice must be preferred to the bundle on the 45 degree line intersecting with the budget line $\partial\mathscr{B}(\mathbf{p}^t, \mathbf{p}^t \cdot \mathbf{x}^t)$. The second states that the observed choice is superior to points on the budget line that intersect with the net $\mathcal{N}$. These conditions follow immediately from the assumed optimality of $\mathbf{x}^t$ in the set $B^t(e^t)$. The next results states that these conditions are also sufficient for concave RDU-rationalizability.

Proposition A.4. *Consider a two-state experiment in which $\pi_1 \geqslant \pi_2$ and suppose a modified data set from this experiment, $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t))\}_{t \in T}$, is RDU-rationalizable with (ρ_1, ρ_2) satisfying (11) (in the main text) and a concave Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$. Let $\mathcal{X} \subset \mathbb{R}_+$ contain $\mathcal{X}^{**}$ (as defined by (15) in the main text). Then u's restriction to $\mathcal{X}$, $\bar{u} : \mathcal{X} \to \mathbb{R}_+$, has the following properties:*

(i) *$\bar{u}(r) < \bar{u}(r')$ for all $r < r'$;*

(ii) *for any three adjacent points $r < r' < r''$ in $\mathcal{X}$,*

$$\frac{\bar{u}(r') - \bar{u}(r)}{r' - r} \geqslant \frac{\bar{u}(r'') - \bar{u}(r')}{r'' - r'}$$

(iii) $\bar{u}_\ell : \mathbb{R}_+ \to \mathbb{R}_+$, *the linear extension of* $\bar{u}$, *satisfies (a.26) and (a.27) at all* t, *with* ϕ *given by (a.25).*

Conversely, if for some modified data set $\mathcal{O}(\mathbf{e})$ *there is* (ρ_1, ρ_2) *satisfying (11) and* $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ *such that properties* (i), (ii), *and* (iii) *are satisfied, then* $\bar{u}$*'s linear extension* $\bar{u}_\ell$ *is a concave Bernoulli function that RDU-rationalizes* $\mathcal{O}(\mathbf{e})$.

Proof. If $\mathcal{O}(\mathbf{e})$ is a data set from this experiment that is RDU-rationalizable with a concave Bernoulli function u, then its restriction $\bar{u}$ must obviously satisfy (i) and (ii). And we have already established that the conditions (a.26) and (a.27) hold at all t. We turn now to the converse.

Suppose there is $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ satisfying conditions (i) to (iii) and let $\bar{u}_\ell$ be its linear extension. Then (i) guarantees that $\bar{u}_\ell$ is strictly increasing on $\mathbb{R}_+$ and (ii) guarantees that $\bar{u}_\ell$ is concave. We claim that $V(\mathbf{x}^t) \geqslant V(\mathbf{x})$ for any $\mathbf{x} \in B^t(e^t)$, where $V(x_1, x_2) = \phi(\bar{u}_\ell(x_1), \bar{u}_\ell(x_2))$. Indeed, $\bar{u}_\ell$ is linear between adjacent values of $\mathcal{X}$ and the probability weights in the definition of ϕ (see (a.25) is constant in the regions of $\mathbb{R}^2_+$ above and below the 45 degree line. It follows that the utility function V (with $u = \bar{u}_\ell$) is linear in every set $W \subset \mathbb{R}^2_+$, where W equals

$$([r, r'] \times [z, z']) \cap \{(x_1, x_2) \in \mathbb{R}^2_+ : x_2 \leqslant x_1\}$$

or

$$([r, r'] \times [z, z']) \cap \{(x_1, x_2) \in \mathbb{R}^2_+ : x_2 \geqslant x_1\},$$

with r and r' being adjacent points in $\mathcal{X}$ (and similarly z and z'). It follows that in any set of the form $\widehat{W} \cap \mathscr{B}(\mathbf{p}^t, e^t \mathbf{p}^t \cdot \mathbf{x}^t)$, for some $\widehat{W} = [\hat{r}, \hat{r}'] \times [\hat{z}, \hat{z}']$, the function V will be maximized at some (x_1^*, x_2^*) that lies on the budget line (i.e., $(x_1^*, x_2^*) \in \partial\mathscr{B}(\mathbf{p}^t, e^t \mathbf{p}^t \cdot \mathbf{x}^t)$, since V is strictly increasing) and either $x_1^* \in \{\hat{r}, \hat{r}'\}$, $x_2^* \in \{\hat{z}, \hat{z}'\}$, or $x_1^* = x_2^*$. More generally, there is (x_1^{**}, x_2^{**}) maximizing $V(x_1, x_2) = \phi(\bar{u}_\ell(x_1), \bar{u}_\ell(x_2))$ in $\mathscr{B}(\mathbf{p}^t, e^t \mathbf{p}^t \cdot \mathbf{x}^t)$, such that either $x_1^{**} = x_2^{**}$, with $(x_1^{**}, x_2^{**}) \in \partial\mathscr{B}(\mathbf{p}^t, e^t \mathbf{p}^t \cdot \mathbf{x}^t)$, or $(x_1^{**}, x_2^{**}) \in \mathcal{N} \cap \partial\mathscr{B}(\mathbf{p}^t, e^t \mathbf{p}^t \cdot \mathbf{x}^t)$. Therefore, (a.26) and (a.27) are sufficient to guarantee the optimality of $\mathbf{x}^t$ in $B^t(e^t)$. **QED**

Given the functional form of ϕ (see (a.25)), conditions (i) to (iii) translate into a finite set of inequalities that are bilinear in the unknowns $\{\rho_1, \rho_2\}$ and $\{\bar{u}(r)\}_{r \in \mathcal{X}}$. In the implementation of this test in Section IV, we let ρ_1 and ρ_2 take different values on a very fine

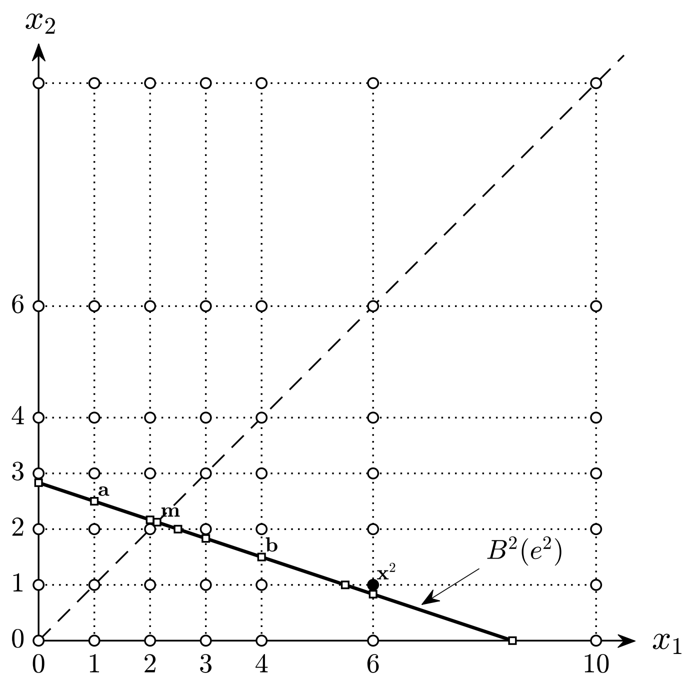


Figure A.2: Test for concave RDU-rationalizability

grid in $[0,1]^2$, subject to (11), and (for each case) perform the corresponding linear test to search for a solution in $\{\bar{u}(r)\}_{r\in\mathcal{X}}$; $\mathcal{O}(\mathbf{e})$ is RDU-rationalizable if such a solution exists for some value of (ρ_1, ρ_2).

As an illustration of how this test works, we consider the data set $\mathcal{O}(\mathbf{e})$ depicted in Figure 4a. (Recall that we also looked at this example when discussing concave EU-rationalizability in Section III.) Given that the three observed choices are $(2,4)$, $(4,3)$, and $(6,1)$ and choosing $\bar{r} = 10$, we obtain $\mathcal{X}^{**} = \{0,1,2,3,4,6,10\}$. Letting $\mathcal{X} = \mathcal{X}^{**}$, the test involves setting up a collection of inequalities that are bilinear in $\{\rho_1, \rho_2\}$ and $\{\bar{u}(r)\}_{r\in\mathcal{X}}$ (corresponding to conditions (i) to (iii)) and checking if it has a solution. Conditions (i) and (ii) are clear enough, so let us explain condition (iii), which guarantees, for each t, the optimality of the observed choice $\mathbf{x}^t$ over a finite set of alternatives in the constraint set $B^t(e^t)$. To be specific, we consider its restrictions on the second observation.

Figure A.2 depicts $B^2(e^2)$, along with $\mathcal{N}$ and the 45 degree line, which are indicated by dashed lines. These lines divide $\mathbb{R}^2_+$ into sets (corresponding to W in the proof of Proposition A.4) that are either boxes or right-angled triangles. There are nine bundles in $\mathcal{N} \cap \partial\mathscr{B}^2(\mathbf{p}^t, e^2\mathbf{p}^t \cdot \mathbf{x}^t)$; these bundles along with the bundle $\mathbf{m} = (m,m)$ on the 45 degree line are indicated by the little squares on the budget line; the coordinates of these bundles could be computed from $\mathcal{O}(\mathbf{e})$. Condition (iii) requires that V (computed with ϕ given by (a.25) and $u = \bar{u}_\ell$) satisfies the following: $V(\mathbf{x}^2) \geqslant V(\mathbf{m})$ and $V(\mathbf{x}^2) \geqslant V(\mathbf{x})$ for $\mathbf{x} \in \mathcal{N} \cap \partial\mathscr{B}^2(\mathbf{p}^t, e^2\mathbf{p}^t \cdot \mathbf{x}^t)$. This translates into ten bilinear inequalities. Since $\mathbf{x}^2$ is below the 45 degree line, $V(\mathbf{x}^2) = (1-\rho_2)\bar{u}(6) + \rho_2\bar{u}(1)$. As an illustration, we write out condition (iii) for bundle $\mathbf{m}$, $\mathbf{a}$, and $\mathbf{b}$. The condition $V(\mathbf{x}^2) \geqslant V(\mathbf{m})$ translates into

$$(1-\rho_2)\bar{u}(6) + \rho_2\bar{u}(1) \geqslant \bar{u}(m).$$

The bundle $\mathbf{a} = (1, 2\lambda + 3(1-\lambda))$ for some $\lambda \in (0,1)$ (which can be calculated). Since $\mathbf{a}$ is above the 45 degree line and $\bar{u}_\ell$ is piecewise linear,

$$V(\mathbf{a}) = \rho_1\bar{u}_\ell(1) + (1-\rho_1)\bar{u}_\ell(2\lambda + (1-\lambda)3) = \rho_1\bar{u}(1) + (1-\rho_1)[\lambda\bar{u}(2) + (1-\lambda)\bar{u}(3)].$$

Thus condition (iii) requires

$$(1-\rho_2)\bar{u}(6) + \rho_2\bar{u}(1) \geqslant \rho_1\bar{u}(1) + (1-\rho_2)\lambda\bar{u}(2) + (1-\rho_1)(1-\lambda)\bar{u}(3).$$

In the case of the bundle $\mathbf{b} = (4, \lambda' + (1-\lambda')2)$ for some $\lambda' \in (0,1)$,

$$V(\mathbf{b}) = (1-\rho_2)\bar{u}(4) + \rho_2\lambda'\bar{u}(1) + \rho_2(1-\lambda')\bar{u}(2)$$

since $\mathbf{b}$ is below the 45 degree line and condition (iii) requires

$$(1-\rho_2)\bar{u}(6) + \rho_2\bar{u}(1) \geqslant (1-\rho_2)\bar{u}(4) + \rho_2\lambda'\bar{u}(1) + \rho_2(1-\lambda')\bar{u}(2).$$

A4.3 Concave EU- and RDU-rationalizability with multiple states

We are interested in testing the EU- or RDU-rationalizability of the modified data set $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t)\}_{t\in T}$ (with $B^t(e^t)$ defined by (14)). Recall the definitions of $\mathcal{X}^{**}$, $\mathcal{X}$, $\bar{u}$, and $\bar{u}$'s linear extension $\bar{u}_\ell$ in Section III. Given any Bernoulli function u, $\bar{u}_\ell$ is also a Bernoulli function that is linear between adjacent values of $\mathcal{X}$ and satisfies $\bar{u}_\ell(r) = u(r)$ for all $r \in \mathcal{X}$; furthermore, if u is concave then $\bar{u}_\ell$ is concave, with $u(r) \geqslant \bar{u}_\ell(r)$ for all $r \in [0, \bar{r}]$. The last property guarantees that

$$\phi(u(x_1), u(x_2), \ldots, u(x_{\bar{s}})) \geqslant \phi(\bar{u}_\ell(x_1), \bar{u}_\ell(x_2), \ldots, \bar{u}_\ell(x_{\bar{s}}))$$

for any bundle $\mathbf{x} = (x_1, x_2, ..., x_{\bar{s}})$, so long as ϕ is increasing. This holds, in particular, for the EU and RDU models, where ϕ is strictly increasing, provided $\boldsymbol{\pi} \gg \mathbf{0}$ (see (3) and (a.7)). Furthermore, at any observed choice $\mathbf{x}^t$,

$$\phi(u(x_1^t), u(x_2^t), \dots, u(x_{\bar{s}}^t)) = \phi(\bar{u}_\ell(x_1^t), \bar{u}_\ell(x_2^t), \dots, \bar{u}_\ell(x_{\bar{s}}^t)).$$

Thus if one replaces u with $\bar{u}_\ell$, the utility of every chosen bundle $\mathbf{x}^t$ stays the same, while the utility of any bundle $\mathbf{x}$ is weakly lower. It follows that if $\mathcal{O}(\mathbf{e})$ is EU-rationalizable with the concave Bernoulli function u, then it is also EU-rationalizable with the concave Bernoulli function $\bar{u}_\ell$. Similarly if $\mathcal{O}(\mathbf{e})$ is RDU-rationalizable with the transformation function ρ and the concave Bernoulli function u then it is RDU-rationalizable with the transformation function ρ and the concave Bernoulli function $\bar{u}_\ell$. Thus to determine whether a data set is EU- or RDU-rationalizable, we can confine our search to piecewise linear extensions of strictly increasing functions $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ (where $\mathcal{X}$ is a finite subset of $\mathbb{R}_+$ containing $\mathcal{X}^{**}$ (as defined by (15)).

We now describe in turn the tests for concave EU-rationalizability and for concave RDU-rationalizability. Note that both tests use the GRID method: a data set is rationalizable by a given model if and only if it is possible, within that model, to guaranteeing the superiority of the chosen bundle against a carefully selected, finite set of alternatives in the budget set (at every observation).

A4.3.1 Concave EU-rationalizability

Since $\bar{u}_\ell$ is linear between adjacent values of $\mathcal{X}$, the utility function

$$V(\mathbf{x}) = \sum_{s=1}^{\bar{s}} \pi_s \bar{u}_\ell(x_s) \tag{a.28}$$

is linear for $\mathbf{x} \in W$, where

$$W = [r_1, r_1'] \times [r_2, r_2'] \times ....[r_{\bar{s}}, r_{\bar{s}}'], \tag{a.29}$$

with r_s and r_s' being adjacent values of $\mathcal{X}$ (for all s). Since $W \cap \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ is a convex set and V is linear within this set, within this set V must be maximized at an extreme point. Thus for any piecewise linear Bernoulli function $\bar{u}_\ell$, the corresponding utility function

V (defined by (a.28)) has the following property: there is $\mathbf{x}^* \in \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ such that $V(\mathbf{x}^*) \geqslant V(\mathbf{x})$ for all $\mathbf{x} \in \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ and $\mathbf{x}^* \in \mathcal{S}^t$, where $\mathcal{S}^t$ is defined as follows:

$$\mathbf{x} \in \mathcal{S}^t \text{ if there is } W \text{ such that } \mathbf{x} \text{ is an extreme point of } W \cap \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t).$$

Since $W \cap \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ is a convex polytope, it has a finite number of extreme points; with only finitely many distinct sets W, $\mathcal{S}^t$ is also a finite set. It follows that to check if $\bar{u}_\ell$ generates V such that $V(\mathbf{x}^t) \geqslant V(\mathbf{x})$ for all $\mathbf{x} \in B^t(e^t)$ it is necessary and sufficient to check if this inequality holds for all $\mathbf{x} \in \mathcal{S}^t$. Thus we have shown the following result.

Proposition A.5. *Suppose the data set $\mathcal{O}(\mathbf{e})$ is EU-rationalizable with probability $\boldsymbol{\pi} \gg \mathbf{0}$ by a concave Bernoulli function u. Let $\mathcal{X}$ be a finite set in $\mathbb{R}_+$ containing $\mathcal{X}^{**}$ (as defined by (15)). Then u's restriction to $\mathcal{X}$, $\bar{u} : \mathcal{X} \to \mathbb{R}_+$, has the following properties:*

(i) *$\bar{u}(r) < \bar{u}(r')$ for all $r < r'$;*

(ii) *for any three adjacent points $r < r' < r''$ in $\mathcal{X}$,*

$$\frac{\bar{u}(r') - \bar{u}(r)}{r' - r} \geqslant \frac{\bar{u}(r'') - \bar{u}(r')}{r'' - r'}$$

(iii) *$\bar{u}_\ell : \mathbb{R}_+ \to \mathbb{R}_+$, the linear extension of $\bar{u}$, satisfies*

$$\sum_{s=1}^{\bar{s}} \pi_s \bar{u}_\ell(x_s^t) \geqslant \sum_{s=1}^{\bar{s}} \pi_s \bar{u}_\ell(x_s) \quad \text{for all } \mathbf{x} \in \mathcal{S}^t, \text{ and for all } t. \tag{a.30}$$

Conversely, if there is $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ with properties (i), (ii), *and* (iii), *then its linear extension $\bar{u}_\ell$ is a concave Bernoulli function that EU-rationalizes $\mathcal{O}(\mathbf{e})$.*

Proof. Clearly conditions (i) and (ii) are necessary because u is a concave Bernoulli function, while (iii) holds because $\bar{u}_\ell$ also EU-rationalizes the data set. Conversely, if (i) and (ii) holds, then its linear extension $\bar{u}_\ell$ is a concave Bernoulli function, and we have already explained why condition (iii) is sufficient to guarantee that the data set is EU-rationalized with this Bernoulli function. **QED**

Note that we have consciously stated this result in a way that it is analogous to Proposition 1. Conditions (i) and (ii) are the same in both propositions, while condition (iii) in this result generalizes condition (iii) in Proposition 1; indeed, when there are there are two states, $\mathcal{S}^t = \mathcal{N} \cap \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$.

A4.3.2 Concave RDU-rationalizability

In this case, the utility function $V : \mathbb{R}^{\bar{s}}_{+} \to \mathbb{R}_{+}$ has the form $V(\mathbf{x}) = \sum_{s=1}^{\bar{s}} \delta(\mathbf{x})u(x_s)$, where δ is defined by (a.6). We claim that we can partition the consumption space $\mathbb{R}^{\bar{s}}_{+}$ such that within each region, the probability weights are the same, i.e., for $\mathbf{x}'$ and $\mathbf{x}''$ in the region, $\delta(\mathbf{x}', s) = \delta(\mathbf{x}'', s)$ for all s. Indeed, let $\theta : S \to S$ be a permutation of the set of states S and $X_\theta = \{\mathbf{u} \in \mathbb{R}^{\bar{s}}_{+} : r(\mathbf{u}, s) = \theta(s) \text{ for all } s \in S\}$, where r is the ranking function defined in Section A2. For example, suppose there are three states and $\hat{\theta}(1) = 2$, $\hat{\theta}(2) = 3$ and $\hat{\theta}(3) = 1$. Then $(5, 4, 8)$ is in $X_{\hat{\theta}}$ because consumption is highest in the third state, second highest in the first state, and lowest in the second state. Similarly, one could check that the element $(1, 1, 5) \in X_{\hat{\theta}}$ but $(5, 1, 5) \notin X_{\hat{\theta}}$. Obviously, the regions X_θ associated with all possible permutations θ partition the space $\mathbb{R}^{\bar{s}}_{+}$. For $\mathbf{x}'$ and $\mathbf{x}''$ in X_θ, we have $r(\mathbf{x}', s) = r(\mathbf{x}'', s)$ for all s, which implies that $\delta(\mathbf{x}', s)) = \delta(\mathbf{x}'', s)$ for all s.

Notice also that the closure of X_θ, which we shall denote by $\bar{X}_\theta$ is an easily characterized, finitely-generated convex cone. For example, $\bar{X}_{\hat{\theta}}$ is the smallest cone containing $(0, 0, 1)$, $(1, 0, 1)$ and $(1, 1, 1)$; as another example, if the permutation is the identity map ι, then $\bar{X}_\iota$ is the smallest cone containing $(1, 0, 0)$, $(1, 1, 0)$, and $(1, 1, 1)$.

In the case where the Bernoulli function is $\bar{u}_\ell$, the linear extension of some strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_{+}$, the corresponding utility function $V : \mathbb{R}_{+} \to \mathbb{R}_{+}$ given by $V(\mathbf{x}) = \sum_{s=1}^{\bar{s}} \delta(\mathbf{x})\bar{u}_\ell(x_s)$ is linear in $W \cap X_\theta$, where W is given by (a.29) and, by the continuity of V, linear in $W \cap \bar{X}_\theta$. This is because $\bar{u}_\ell$ is linear in the set W and $\delta(\cdot, s)$ is constant for any two elements in X_θ. Given that

$$W \cap \bar{X}_\theta \cap \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t),$$

is a convex set, the maximum of V in this set is achieved at one of its extreme points. More generally, there is $\mathbf{x}^* \in \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ such that $V(\mathbf{x}^*) \geqslant V(\mathbf{x})$ for all $\mathbf{x} \in \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ and $\mathbf{x}^* \in \mathcal{C}^t$, where $\mathcal{C}^t$ is defined as follows:

$$\mathbf{x} \in \mathcal{C}^t \text{ if } \mathbf{x} \text{ is an extreme point of } W \cap \bar{X}_\theta \cap \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t) \text{ for some } W \text{ and } \theta.$$

Since $W \cap \bar{X}_\theta \cap \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ is a convex polytope, it has a finite number of extreme points. Furthermore, there are only finitely many distinct sets W and finitely many permutations

θ, which means that $\mathcal{C}^t$ is finite set. To check if $\bar{u}_\ell$ has the property that $V(\mathbf{x}^t) \geqslant V(\mathbf{x})$ for all $\mathbf{x} \in B^t(e^t)$ it is necessary and sufficient to check if this inequality holds for the finite set of $\mathbf{x} \in \mathcal{C}^t$.[12] The following result summarizes our observations.

Proposition A.6. *Suppose the data set $\mathcal{O}(\mathbf{e})$ is RDU-rationalizable with the transformation function ρ and the concave Bernoulli function u. Let $\mathcal{X}$ be a finite set in $\mathbb{R}_+$ that contains $\mathcal{X}^{**}$, where the latter is defined by (15) (in the main paper) and let*

$$\Gamma = \left\{ r \in \mathbb{R} : r = \sum_{s \in S'} \pi_s \text{ for some } S' \subseteq S = \{1, 2, \ldots, \bar{s}\} \right\}.$$

Then the restriction of ρ to Γ, $\bar{\rho} : \Gamma \to \mathbb{R}$, and the restriction of u to $\mathcal{X}$, $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ have the following properties:

(i) $\bar{u}(r) < \bar{u}(r')$ *for all* $r < r'$;

(ii) *for any three adjacent points* $r < r' < r''$ *in* $\mathcal{X}$,

$$\frac{\bar{u}(r') - \bar{u}(r)}{r' - r} \geqslant \frac{\bar{u}(r'') - \bar{u}(r')}{r'' - r'}$$

(iii) $\bar{u}_\ell : \mathbb{R}_+ \to \mathbb{R}_+$, *the linear extension of* $\bar{u}$, *satisfies*

$$\sum_{s=1}^{\bar{s}} \delta(\mathbf{x}^t, s) \bar{u}_\ell(x_s^t) \geqslant \sum_{s=1}^{\bar{s}} \delta(\mathbf{x}, s) \bar{u}_\ell(x_s) \text{ for all } \mathbf{x} \in \mathcal{C}^t, \text{ and for all } t \tag{a.31}$$

(with δ defined by (a.6)).

Conversely, if there are strictly increasing functions $\bar{\rho} : \Gamma \to \mathbb{R}_+$ (with $\bar{\rho}(0) = 0$ and $\bar{\rho}(1) = 1$) and $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ with properties (i), (ii), *and* (iii), *then its linear extension $\bar{u}_\ell$ is a concave Bernoulli function that RDU-rationalizes $\mathcal{O}(\mathbf{e})$ along with any transformation function ρ extending $\bar{\rho}$.*

Proof. Conditions (i) and (ii) are clearly necessary because u is strictly increasing (by the definition of a Bernoulli function) ad concave (by assumption), while (iii) is also necessary because if u RDU-rationalizes $\mathcal{O}(\mathbf{e})$ then so does $\bar{u}_\ell$ (with the same transformation function). Conversely, if (i) and (ii) holds, then its linear extension $\bar{u}_\ell$ is a concave Bernoulli function,

[12] There are efficient algorithms to find the extreme points of a convex polytope (see Matheiss and Rubin (1980)).

and we have already explained why condition (iii) is sufficient to guarantee that the data set is RDU-rationalized with the Bernoulli function $\bar{u}_\ell$ and any transformation function that extends $\bar{\rho}$. **QED**

This result generalizes Proposition A.4 to the case of multiple states. Conditions (i) and (ii) are the same in both propositions, while condition (iii) in this result generalizes condition (iii) in Proposition A.4; indeed, when there are there are two states, $\mathcal{C}^t$ consists precisely of $\mathcal{N} \cap \partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ and the point where the budget line $\partial\mathscr{B}(\mathbf{p}^t, e^t\mathbf{p}^t \cdot \mathbf{x}^t)$ meets the 45 degree line.

Note that the inequality (a.31) in condition (iii) can be written as a bilinear inequality in the unknowns $\{\bar{\rho}(\gamma)\}_{\gamma\in\Gamma}$ and $\{\bar{u}(r)\}_{r\in\mathcal{X}}$ (and the number of unknowns is finite since both Γ and $\mathcal{X}$ are finite sets). Thus the test provided by Proposition A.6 involves solving a finite set of bilinear inequalities.

A5. Description of the GARP and F-GARP tests

In addition to the expected utility, disappointment aversion, and rank dependent utility models, our empirical implementation in Section IV of the main text also implements the (already known) rationalizability tests of two more basic utility models: locally nonsatiated utility and stochastically monotone utility. In this section, we describe these models and explain how they could be tested on classical data sets which have been modified by some vector $\mathbf{e} \in [0,1]^T$, in the sense given by (14) in Section II.

A5.1 Locally nonsatiated utility (LNU)

In this section, we continue with the discussion of Section II in the main text. Our discussion here partially overlaps with Section A1.1 of the Online Appendix but it could be read independently of that section.

The locally nonsatiated utility (LNU) model is the most permissive of the models that we consider in the empirical implementation in Section IV of the main text since all the other models are special cases of this model. A utility function $U : \mathbb{R}^{\bar{s}}_+ \to \mathbb{R}$ is *locally nonsatiated* if at every open neighborhood N of $\mathbf{x} \in \mathbb{R}^{\bar{s}}_+$, there is $\mathbf{y} \in N$ such that $U(\mathbf{y}) > U(\mathbf{x})$. Afriat's (1967) Theorem tells us that a classical data set $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^T$ is LNU-rationalizable if and only if it obeys a consistency condition known as the Generalized Axiom of Revealed Preference (GARP). There is natural generalization of GARP (which we shall for convenience also simply refer to as GARP) which characterizes the rationalizability of any modified data set $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t)\}_{t=1}^T$, where

$$\mathbf{e} = (e^1, e^2, ..., r^T) \in [0,1]^T$$

and $B^t(e^t)$ is given by (14).

Let $\mathcal{D} = \{\mathbf{x}^t : t = 1, 2, \ldots, T\}$; in other words, $\mathcal{D}$ consists of those bundles that were chosen by the subject at some observation in the data set. For bundles $\mathbf{x}^t$ and $\mathbf{x}^{t'}$ in $\mathcal{D}$, $\mathbf{x}^t$ is said to be *revealed preferred to* $\mathbf{x}^{t'}$ at the efficiency vector $\mathbf{e}$ (we denote this by $\mathbf{x}^t \succcurlyeq^*_{\mathbf{e}} \mathbf{x}^{t'}$) if $\mathbf{x}^{t'} \in B^t(e^t)$;[13] $\mathbf{x}^t$ is said to be *strictly revealed preferred to* $\mathbf{x}^{t'}$ (and we denote this by $\mathbf{x}^t \succ^*_{\mathbf{e}} \mathbf{x}^{t'}$) if $\mathbf{x}^{t'} \in B^t(e^t)$ and $\mathbf{p}^t \cdot \mathbf{x}^{t'} < e^t\, \mathbf{p}^t \cdot \mathbf{x}^t$. Then the following extension of Afriat's Theorem holds: $\mathcal{O}(\mathbf{e})$ is rationalizable by a locally nonsatiated utility function if and only if, whenever there are observations $(\mathbf{p}^{t_i}, \mathbf{x}^{t_i})$ (for $i = 1, 2, \ldots, n$) in $\mathcal{O}$ satisfying

$$\mathbf{x}^{t_1} \succcurlyeq^*_{\mathbf{e}} \mathbf{x}^{t_2},\ \mathbf{x}^{t_2} \succcurlyeq^*_{\mathbf{e}} \mathbf{x}^{t_3},\ \ldots\ ,\ \mathbf{x}^{t_{n-1}} \succcurlyeq^*_{\mathbf{e}} \mathbf{x}^{t_n}, \text{ and } \mathbf{x}^{t_n} \succcurlyeq^*_{\mathbf{e}} \mathbf{x}^{t_1}, \tag{a.32}$$

then we cannot replace $\succcurlyeq^*_{\mathbf{e}}$ with $\succ^*_{\mathbf{e}}$ anywhere in this chain (see Halevy, Persitz, and Zrill (2018)). The latter property states that while there can be revealed preference cycles in $\mathcal{O}$, they cannot contain a strict revealed preference.[14] This property is a generalization of GARP, which is the special case where $\mathbf{e} = (1, 1, \ldots, 1)$. We shall also refer to this generalization as GARP, bearing in mind that it is always conditional on some $\mathbf{e}$.[15]

From an empirical perspective, it is important to note that checking whether or not GARP holds is computationally undemanding: the (strict) revealed preference relations on

[13] Our terminology differs a little from the standard, which refers to $\succcurlyeq^*_{\mathbf{e}}$ as the *direct revealed preference* relation and uses *revealed preference* to refer to the transitive closure of this relation. Since our exposition avoids any discussion of the transitive closure, we have adopted the simpler terminology here.

[14] To be specific: a modified data set $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t))\}_{t=1}^T$ will obey this property whenever it is rationalizable by an LNU function U, in the sense that $U(\mathbf{x}^t) \geqslant U(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^{\bar{s}}_+$ such that $\mathbf{p}^t \cdot \mathbf{x} \leqslant e^t \mathbf{p}^t \cdot \mathbf{x}^t$ (whether or not U is continuous); conversely, whenever a modified data set $\mathcal{O}(\mathbf{e})$ obeys this property, then $\mathcal{O}(\mathbf{e})$ is rationalizable by a continuous, strictly increasing, and concave utility function U.

[15] We gave another generalization of GARP in Section A1.1 of this Appendix for general downward comprehensive constraint sets. The two definitions are closely related. Note that $B^t(e^t) = \{\mathbf{x}^t\} \cup \mathscr{B}(\mathbf{p}^t, \mathbf{p}^t \cdot \mathbf{x}^t)$ is not downward comprehensive for any $e^t < 1$; however its downward closure $\underline{B}^t(e^t) = \{\mathbf{x} \in \mathbb{R}^{\bar{s}}_+ : \mathbf{x} \leqslant \mathbf{x}^t\} \cup \mathscr{B}(\mathbf{p}^t, \mathbf{p}^t \cdot \mathbf{x}^t)$ is downward comprehensive. One could check that the data set $\{(\mathbf{x}^t, \underline{B}^t(e^t))\}_{t=1}^T$ obeys GARP in the sense defined in Section A1.1, if and only if the modified data set $\{(\mathbf{x}^t, B^t(e^t))\}_{t=1}^T$ obeys GARP in the sense defined here.

$\mathcal{D}$ can be easily constructed (for a fixed $\mathbf{e}$); once this has been established, we can apply Warshall's algorithm to compute the transitive closure of the revealed preference relations and then check for the absence of cycles containing strict revealed preferences. This is the test of GARP we use in our empirical implementation.

A5.2 Stochastically monotone utility (SMU)

We discuss in greater detail the stochastically monotone utility (SMU) function model introduced in Section IV and explain how rationalizability by this model of a modified data set can tested.

For $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^{\bar{s}}_{+}$, we write $\mathbf{x} \geqslant_{FSD} \mathbf{y}$ if $\mathbf{x}$ first order stochastically dominates $\mathbf{y}$ (given the payoffs and the objectively known probabilities) and write $\mathbf{x} >_{FSD} \mathbf{y}$ if $\mathbf{x} \geqslant_{FSD} \mathbf{y}$ and the two distributions are distinct. One way of sharpening the locally nonsatiated utility model is to require that the utility function $U : \mathbb{R}^{\bar{s}}_{+} \to \mathbb{R}$ be *stochastically monotone*. By this we mean that $U(\mathbf{x}) > (\geqslant)\, U(\mathbf{y})$ whenever $\mathbf{x} >_{FSD} \mathbf{y}$ ($\mathbf{x} \geqslant_{FSD} \mathbf{y}$). Note that the rank dependent utility, disappointment aversion, and expected utility models all obey this property. In the Choi *et al.* (2007) experiment, there are two states; it is straightforward to check that when $\pi_1 = \pi_2 = 1/2$, a utility function is stochastically monotone if and only if it is strictly increasing and symmetric; when $\pi_1 > \pi_2$, a utility function U is stochastically monotone if and only if it is strictly increasing and $U(a,b) > U(b,a)$ whenever $a > b$.

Since a utility function U that is stochastically monotone is also strictly increasing, any data set that is rationalizable by a stochastically monotone utility function is also rationalizable by a locally nonsatiated utility function. However, the converse is not true; indeed, the single observation given in Example 1 in the main text passes GARP trivially, but it cannot be rationalized by any symmetric and strictly increasing utility function.

Nishimura, Ok, and Quah (2017) have developed a test for rationalizability by stochastically monotone utility functions. The test can be thought of as a version of GARP, but with suitably modified revealed preference relations. Let $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t)\}_{t=1}^{T}$ be a modification by the vector $\mathbf{e}$ of a classical data set $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^{T}$. For bundles $\mathbf{x}^t$ and $\mathbf{x}^{t'}$ in $\mathcal{D} = \{\mathbf{x}^t : t = 1, 2, \ldots, T\}$, we say that $\mathbf{x}^t$ *is F-revealed preferred to* $\mathbf{x}^{t'}$ (at vector $\mathbf{e}$) if there is a bundle $\mathbf{y}$ such that $\mathbf{y} \in B^t(e^t)$ and $\mathbf{y} \geqslant_{FSD} \mathbf{x}^{t'}$; this revealed preference is *strict* if $\mathbf{y}$ can be chosen to satisfy either $\mathbf{p}^t \cdot \mathbf{y} < e^t\,\mathbf{p}^t \cdot \mathbf{x}^t$ or $\mathbf{y} >_{FSD} \mathbf{x}^t$. Nishimura, Ok, and Quah (2017) show that $\mathcal{O}(\mathbf{e})$ is rationalizable by a stochastically monotone utility function if and only if it does not admit F-revealed preference cycles (such as (a.32)) containing strict F-revealed preferences; we call the latter property F-GARP (at the efficiency threshold $\mathbf{e}$), where 'F' stands for first order stochastic dominance.[16] Clearly this result is analogous to the characterization of rationalizability by locally nonsatiated utility functions, except that the revealed preferences are defined differently.

It is not difficult to see that in checking whether $\mathbf{x}^t$ is F-revealed preferred or F-strictly revealed preferred to $\mathbf{x}^{t'}$ we can always choose the intermediate vector $\mathbf{y}$, if it exists, to be a permutation of $\mathbf{x}^{t'}$. In particular, suppose that there are just two equiprobable states. Then $\mathbf{x}^t$ is F-revealed preferred to $\mathbf{x}^{t'} = (a, b)$ if one of two conditions hold: either (i) $\mathbf{p}^t \cdot (a, b) \leqslant e^t\mathbf{p}^t \cdot \mathbf{x}^t$ or (ii) $\mathbf{p}^t \cdot (b, a) \leqslant e^t\mathbf{p}^t \cdot \mathbf{x}^t$ (since (a, b) and (b, a) are stochastically equivalent). In the case where there are two states and $\pi_1 > \pi_2$, $\mathbf{x}^t$ is revealed preferred to $\mathbf{x}^{t'} = (a, b)$ if either one of two conditions hold: (i) $\mathbf{p}^t \cdot (a, b) \leqslant e^t\mathbf{p}^t \cdot \mathbf{x}^t$ or (ii) $a \leqslant b$ and $\mathbf{p}^t \cdot (b, a) \leqslant e^t\mathbf{p}^t \cdot \mathbf{x}^t$ (since (b, a) first order stochastically dominates (a, b) when $a \leqslant b$ and $\pi_1 > \pi_2$). All the experimental data analyzed in this paper involves just two states, so their F-revealed preference relations are easily obtained in this manner. Once the relations are in place, checking for the absence of cycles via Warshall's algorithm is also straightforward.

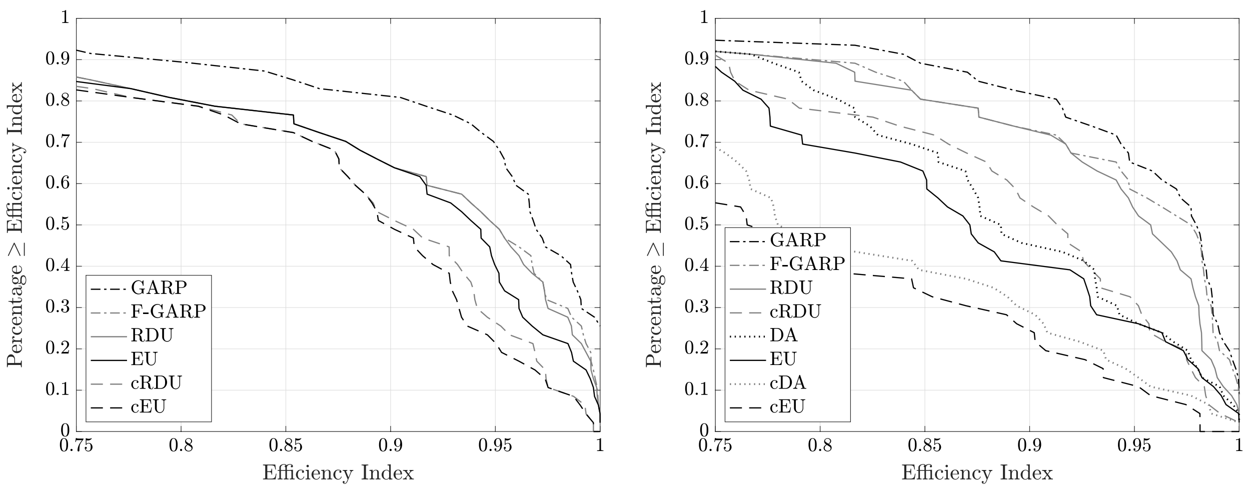
A6. Empirical Application: More on Choi *et al.* (2007)

A6.1 Generating random data sets obeying GARP or F-GARP

The process of generating a random data set obeying GARP (F-GARP) *at a given efficiency threshold* is as follows. First, we generate 50 budget sets as in Choi *et al.* (2007). Next, we select a budget line and randomly (uniformly) choose a bundle on that line. Then we select a second budget line and randomly choose a bundle from that part of the line which guarantees that this observation, along with the first, obeys GARP (F-GARP) at the given efficiency threshold. A third budget line is then selected and a bundle randomly chosen

[16] The term F-GARP is ours and not found in Nishimura, Ok, and Quah (2017). Their result is applicable to general data sets with compact constraint sets, and not just to modified data sets. Their result states that if F-GARP holds then the data set can be rationalized by a stochastically monotone *and continuous* utility function.

from that part of the line so that all three observations together obey GARP (F-GARP). Note that such a bundle *must exist*; indeed, the demand (on the third budget line) arising from any locally nonsatiated (stochastically monotone) utility function rationalizing the first two observations will have this property. We then choose a fourth budget line and a bundle on that line randomly so that the first four observations obey GARP (F-GARP), and so on. We generate 30,000 data sets (with 50 observations each) which pass (GARP) F-GARP at each of the two efficiency thresholds (0.9 and 0.95) in this manner. (So there are four distinct collections of data sets, with each collection containing 30,000 data sets.) After that we subject each data set to a test for a more specialized model (whether it is EU, DA, or RDU). By the Azuma-Hoeffding inequality, in order to be $100(1-\delta)$ percent confident that the sample pass rate resulting from a simulation is within ϵ of the true probability of passing the test, we require at least $N = (1/2\epsilon^2)\log(2/\delta)$ samples; with 30,000 samples, we can be 99.5 percent sure that the estimates in Table 5 are within 0.01 of the true value.



(a) $\pi_1 = 1/2$ (b) $\pi_1 \neq 1/2$

Figure A.3: Distributions of efficiency indices (Choi *et al.*, 2007)

A6.2 Probability distortions in the RDU model

The RDU model generalizes the EU model by permitting a distortion of the objective probabilities. With two states, the probability of the less favorable state is distorted to be $g(\pi)$ when π is the true probability. In the asymmetric treatment of Choi *et al.* (2007), π is either 1/3 or 2/3. It turns out that, at the 0.9 threshold, all of the 15 subjects who fail EU but pass RDU (see Table 6 in the main text) continue to do so if we restrict $g(2/3) \in [0.55, 0.75]$ and $g(1/3) \in [0.25, 0.45]$. (Note that g may differ across subjects.) At the 0.95 threshold, the same restrictions on the distorted probabilities capture 11 of the 12 subjects who pass RDU and fail EU. So it seems that those who pass RDU do so with fairly modest distortions of the true probabilities.

Furthermore, there is some evidence that subjects deflate the probability of the less favorable state when it is objectively 2/3 and inflate the probability when it is 1/3, so that the cumulative probability weighting function has the shape favored by cumulative prospect theory. Indeed, if we restrict ourselves to choosing $g(2/3) \in [0.55, 2/3]$ and $g(1/3) \in [1/3, 0.45]$, we still manage to capture every subject who passes the RDU test at the 0.9 threshold and all but two who pass at the 0.95 threshold. On the other hand, the mirror restriction performs very badly: if we insist on choosing $g(2/3) \in [2/3, 0.75]$ and $g(1/3) \in [0.25, 1/3]$, the RDU model captures *no subject* at either efficiency threshold other than those who are already EU-rationalizable. (Note that for any subject who passes RDU, there will typically be more than one set of distorted probabilities at which the subject is rationalizable.)

We know from Table 4 in the main text that, for the symmetric treatment the pass rates of the EU and RDU/DA models differ only at the efficiency threshold 0.95, where 5 subjects pass RDU/DA but fail EU. All 5 subjects pass the RDU test for some $g(1/2) < 0.5$, which is consistent with disappointment aversion, and 4 of them pass with values of $g(1/2)$ chosen from the interval $[0.45, 0.5)$.

A6.3 Distributions of efficiency indices

Figure A.3 depicts the distributions of efficiency indices, including the cRDU and cDA models, thereby augmenting Figure 6 in the main text.

A7. Empirical Application: More on Choi *et al.* (2014)

As described in the main text, the experiment in Choi *et al.* (2014) was conducted on 1,182 CentERpanel adult members, which is a representative sample of the Dutch-speaking population of the Netherlands. Each subject was asked to make allocation decisions on

25 linear budget sets; price vectors were drawn randomly as in Choi *et al.* (2007) and varied across rounds and subjects; income was normalized to one; and state probabilities were symmetric ($\pi_1 = \pi_2 = 1/2$) and commonly known. The distributions of critical cost efficiency indices corresponding to utility maximization (GARP), stochastically monotone utility maximization (F-GARP), expected utility (EU) maximization, and concave expected utility (cEU) maximization are all depicted in Figure 7 in the main text.

	$e = 0.90$	$e = 0.95$	$e = 1.00$
GARP	683/1,182 (58%)	535/1,182 (45%)	231/1,182 (20%)
F-GARP	396/1,182 (34%)	273/1,182 (23%)	14/1,182 (1%)
EU	384/1,182 (32%)	253/1,182 (21%)	12/1,182 (1%)
cEU	330/1,182 (28%)	205/1,182 (17%)	0/1,182 (0%)

Table A.1: Pass rates by efficiency threshold (Choi *et al.*, 2014)

	$e = 0.9$		$e = 0.95$	
	Sample Prop.	Conf. Interval	Sample Prop.	Conf. Interval
cEU	330/1,182 (0.279)	[0.254, 0.306]	205/1,182 (0.173)	[0.152, 0.196]
EU	384/1,182 (0.325)	[0.298, 0.352]	253/1,182 (0.214)	[0.191, 0.239]
EU\cEU	54/1,182 (0.046)	[0.035, 0.059]	48/1,182 (0.041)	[0.030, 0.053]

Table A.2: Confidence intervals on preference types (Choi *et al.*, 2014)

In Tables A.1 and A.2, we display pass rates by efficiency threshold and confidence intervals on preference types for the Choi *et al.* (2014) data, as we did for the Choi *et al.* (2007) data in the main text. It is worth reiterating some observations already made in the main text: (1) the EU model performs well, at least in the sense that around half of GARP-consistent subjects (at some given threshold) are also consistent with the EU model; (2) because the pass rates of the EU model are so close to that for F-GARP, the contribution of the RDU model to account for the behavior of subjects who are not EU-rationalizable must be limited; and (3) unlike the other two experimental studies we analyze, imposing concavity on the Bernoulli function does not significantly lower pass rates for the expected utility model in this case.

Choi *et al.* (2014) related the efficiency index for GARP to various covariates of interest, including sex, age, education, income, work status, occupation, and wealth. We could extend this analysis by relating these covariates with efficiency indices for the other models. Figure

A.4 is an *exact* replication of Figure 3 in Choi *et al.* (2014). It displays sample means and 95% confidence intervals for the efficiency index in different subpopulations. Figures A.5, A.6, and A.7 are new and display sample means and 95% confidence intervals in the same subpopulations, but for the F-GARP, EU, and cEU efficiency indices. While the confidence intervals corresponding to F-GARP are necessarily lower, the patterns just described are nonetheless the same; notice, however, that the 95% confidence intervals are wider, indicating that the efficiency indices corresponding to F-GARP are estimated with less precision. Lastly, the indices corresponding to F-GARP and EU are virtually identical, and for cEU, the indices are very slightly lower than for EU.

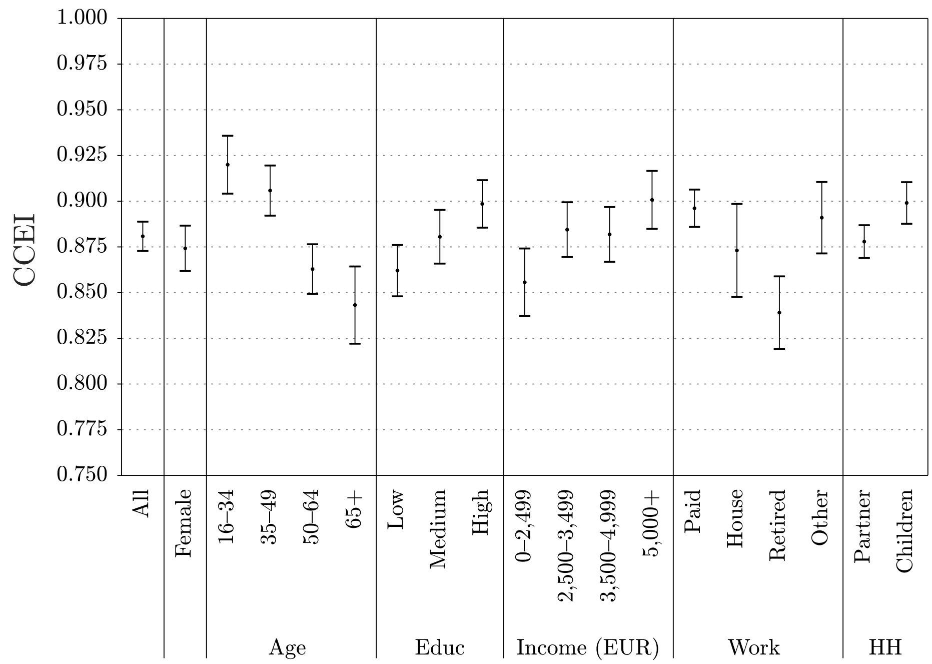


Figure A.4: Mean CCEI Scores (GARP)

In Table A.3, we replicate and extend Table 2 in Choi *et al.* (2014), which presents the results of linear regressions between the efficiency indices and a number of covariates of interest. Columns (1) and (2) correspond to GARP and F-GARP, i.e., with the efficiency indices for GARP and F-GARP as the *dependent* variables, and are simple replications of Choi *et al.* (2014); columns (3) and (4) apply the same linear regression analysis, but to the EU and cEU indices that we have calculated. There are no major changes to the story presented in Choi *et al.* (2014), which is that all else equal, there are (statistically significant)

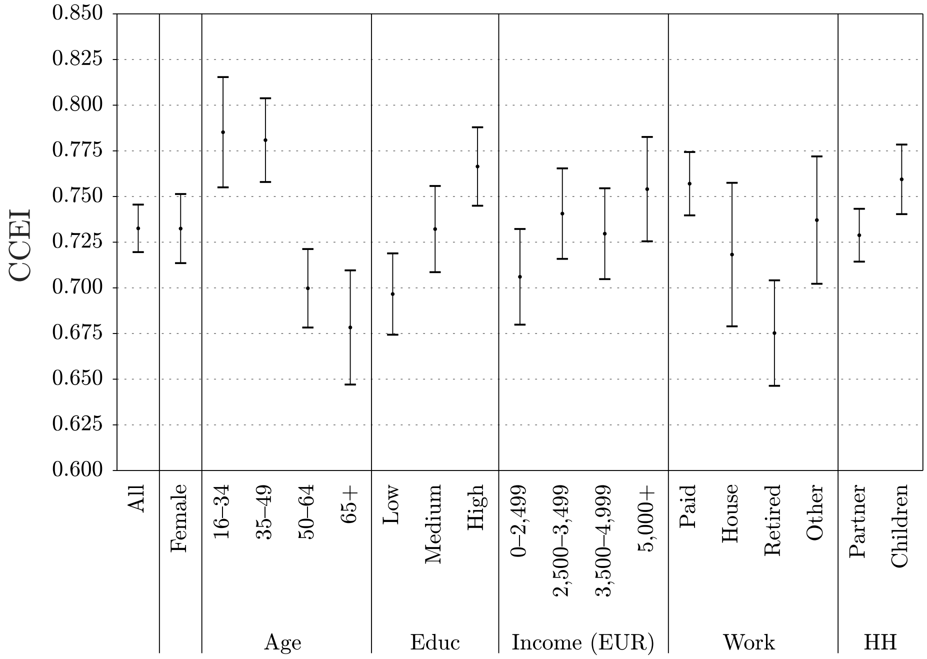


Figure A.5: Mean CCEI Scores (F-GARP)

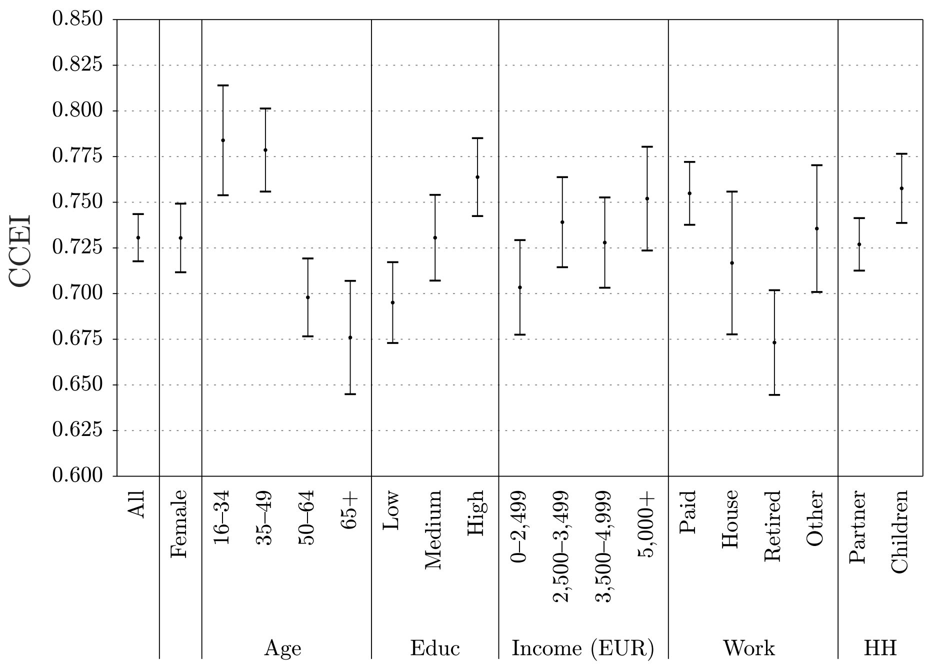


Figure A.6: Mean CCEI Scores (EU)

negative age effects and positive education effects on the various rationality indices.

Lastly, in Tables A.4–A.6, we replicate and extend Table 3 in Choi *et al.* (2014), which

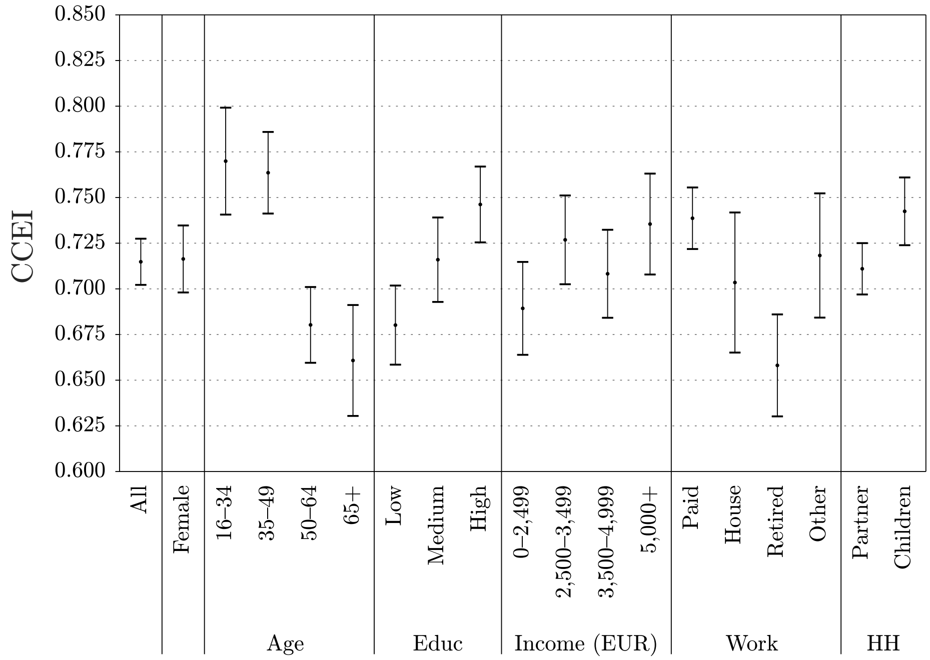


Figure A.7: Mean CCEI Scores (cEU)

contains the results from a set of linear regressions of wealth measures (assets net of liabilities) on efficiency indices plus a set of covariates. They consider three cases: log wealth with sample restricted to households where the respondent is 35 years or older, log wealth with no sample restrictions, and wealth with no sample restrictions (corresponding to columns 1, 2, and 3 respectively in their Table 3). In all three cases, Choi *et al.* (2014) find that an increase in the efficiency index corresponding to GARP is associated with an increase in wealth, after controlling for other factors such as income. Tables A.4–A.6 indicate that this positive association remains true if one uses the efficiency index corresponding to F-GARP, EU, and cEU, though the effect is statistically weaker than if one uses the GARP efficiency index (compare column (1) with columns (2), (3), and (4) in Tables A.4–A.6).

	(1)	(2)	(3)	(4)
Constant	0.887	0.735	0.735	0.725
	(0.022)	(0.037)	(0.037)	(0.036)
Female	−0.024	−0.012	−0.012	−0.009
	(0.009)	(0.015)	(0.015)	(0.014)
Age				
35–49	−0.016	−0.007	−0.008	−0.009
	(0.011)	(0.020)	(0.019)	(0.019)
50–64	−0.052	−0.077	−0.078	−0.082
	(0.011)	(0.020)	(0.020)	(0.019)
65+	−0.051	−0.081	−0.082	−0.087
	(0.020)	(0.032)	(0.032)	(0.031)
Education				
Medium	0.009	0.021	0.021	0.021
	(0.011)	(0.017)	(0.017)	(0.016)
High	0.026	0.059	0.058	0.056
	(0.011)	(0.018)	(0.018)	(0.017)
Income				
€2,500–3,499	0.026	0.026	0.027	0.028
	(0.012)	(0.019)	(0.019)	(0.018)
€3,500–4,999	0.020	0.005	0.006	0.001
	(0.013)	(0.020)	(0.019)	(0.019)
€5,000+	0.033	0.016	0.017	0.015
	(0.014)	(0.022)	(0.022)	(0.022)
Occupation				
Paid Work	0.028	0.030	0.029	0.026
	(0.018)	(0.026)	(0.026)	(0.026)
House Work	0.046	0.039	0.039	0.037
	(0.021)	(0.030)	(0.030)	(0.030)
Other	0.037	0.034	0.034	0.028
	(0.019)	(0.030)	(0.030)	(0.029)
Household Composition				
Partner	−0.026	−0.022	−0.022	−0.022
	(0.011)	(0.018)	(0.018)	(0.018)
Number of Children	0.001	0.001	0.001	0.001
	(0.004)	(0.007)	(0.007)	(0.007)
R^2	0.068	0.057	0.058	0.061
Observations	1,182	1,182	1,182	1,182

Table A.3: CCEI Correlations (OLS)

	(1)	(2)	(3)	(4)
GARP	1.343			
	(0.567)			
F-GARP		0.558		
		(0.303)		
EU			0.580	
			(0.307)	
cEU				0.571
				(0.314)
Log Income	0.584	0.595	0.594	0.594
	(0.132)	(0.131)	(0.131)	(0.131)
Female	−0.313	−0.327	−0.327	−0.328
	(0.177)	(0.178)	(0.178)	(0.178)
Household Composition				
Partner	0.652	0.653	0.654	0.654
	(0.181)	(0.181)	(0.181)	(0.181)
Number of Children	0.090	0.087	0.086	0.086
	(0.093)	(0.093)	(0.093)	(0.093)
Age	−0.303	−0.299	−0.299	−0.300
	(0.347)	(0.345)	(0.345)	(0.346)
Age^2	0.007	0.007	0.007	0.007
	(0.006)	(0.006)	(0.006)	(0.006)
Age^3	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Education				
Pre-Vocational	0.269	0.246	0.246	0.248
	(0.464)	(0.464)	(0.464)	(0.464)
Pre-University	0.634	0.609	0.609	0.616
	(0.478)	(0.478)	(0.478)	(0.478)
Senior Vocational	0.416	0.395	0.395	0.399
	(0.474)	(0.476)	(0.476)	(0.476)
Vocational College	0.490	0.471	0.472	0.477
	(0.451)	(0.452)	(0.452)	(0.452)
University	0.725	0.700	0.700	0.706
	(0.473)	(0.474)	(0.474)	(0.474)
Occupation				
Paid Work	0.207	0.234	0.236	0.232
	(0.322)	(0.329)	(0.329)	(0.328)
House Work	0.552	0.606	0.606	0.602
	(0.406)	(0.415)	(0.415)	(0.414)
Retired	0.131	0.125	0.126	0.116
	(0.318)	(0.324)	(0.324)	(0.323)
Constant	6.308	6.950	6.920	6.972
	(6.419)	(6.432)	(6.430)	(6.437)
R^2	0.204	0.198	0.199	0.198
Observations	517	517	517	517

Table A.4: CCEI Scores and Log Wealth (age 35 and above)

	(1)	(2)	(3)	(4)
GARP	1.103			
	(0.535)			
F-GARP		0.508		
		(0.285)		
EU			0.533	
			(0.287)	
cEU				0.532
				(0.294)
Log Income	0.606	0.616	0.615	0.615
	(0.127)	(0.126)	(0.126)	(0.126)
Female	−0.356	−0.366	−0.365	−0.367
	(0.164)	(0.165)	(0.165)	(0.165)
Household Composition				
Partner	0.595	0.598	0.599	0.600
	(0.171)	(0.171)	(0.171)	(0.172)
Number of Children	0.109	0.105	0.105	0.104
	(0.086)	(0.085)	(0.085)	(0.085)
Age	−0.008	−0.003	−0.003	−0.004
	(0.208)	(0.206)	(0.205)	(0.205)
Age^2	0.002	0.002	0.002	0.002
	(0.004)	(0.004)	(0.004)	(0.004)
Age^3	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Education				
Pre-Vocational	0.246	0.225	0.225	0.226
	(0.462)	(0.462)	(0.462)	(0.462)
Pre-University	0.562	0.538	0.538	0.544
	(0.476)	(0.477)	(0.477)	(0.477)
Senior Vocational	0.421	0.402	0.402	0.406
	(0.468)	(0.470)	(0.470)	(0.470)
Vocational College	0.527	0.505	0.505	0.509
	(0.449)	(0.449)	(0.449)	(0.449)
University	0.685	0.662	0.661	0.666
	(0.465)	(0.466)	(0.466)	(0.466)
Occupation				
Paid Work	0.227	0.252	0.254	0.250
	(0.321)	(0.327)	(0.327)	(0.326)
House Work	0.604	0.640	0.640	0.635
	(0.413)	(0.419)	(0.419)	(0.419)
Retired	0.191	0.187	0.188	0.179
	(0.318)	(0.322)	(0.322)	(0.321)
Constant	0.476	0.907	0.888	0.921
	(3.598)	(3.565)	(3.564)	(3.554)
R^2	0.239	0.236	0.237	0.237
Observations	566	566	566	566

Table A.5: CCEI Scores and Log Wealth

	(1)	(2)	(3)	(4)
GARP	101,018.2			
	(52,669.0)			
F-GARP		47,934.5		
		(34,707.4)		
EU			50,981.1	
			(35,188.0)	
cEU				51,377.3
				(35,928.7)
Log Income	1.775	1.792	1.791	1.789
	(0.353)	(0.354)	(0.354)	(0.354)
Female	−32,484.4	−34,186.3	−34,151.2	−34,273.8
	(17,523.6)	(17,635.0)	(17,629.3)	(17,647.3)
Household Composition				
Partner	46,193.4	45,674.5	45,776.0	45,901.9
	(17,172.5)	(17,171.2)	(17,162.9)	(17,161.5)
Number of Children	14,094.0	13,804.6	13,749.5	13,682.4
	(8,353.9)	(8,416.9)	(8,416.8)	(8,421.1)
Age	−19,199.0	−18,243.8	−18,094.1	−18,159.3
	(30,167.8)	(30,091.1)	(30,050.2)	(30,084.5)
Age^2	469.5	450.9	449.1	449.8
	(523.7)	(521.6)	(520.7)	(521.6)
Age^3	−2.9	−2.8	−2.8	−2.8
	(2.9)	(2.9)	(2.9)	(2.9)
Education				
Pre-Vocational	14,151.0	10,648.6	10,483.9	10,806.9
	(43,449.1)	(43,652.8)	(43,689.0)	(43,652.2)
Pre-University	59,056.5	55,270.8	55,068.3	55,574.2
	(44,747.8)	(44,981.1)	(45,012.0)	(44,976.8)
Senior Vocational	28,328.7	25,421.0	25,316.7	25,599.8
	(42,419.1)	(42,953.3)	(42,994.4)	(42,937.2)
Vocational College	31,402.7	28,377.5	28,228.3	28,725.9
	(42,048.1)	(42,454.9)	(42,477.9)	(42,396.6)
University	77,652.0	74,224.2	73,980.6	74,548.6
	(47,707.2)	(47,801.5)	(47,831.9)	(47,814.5)
Occupation				
Paid Work	−12,600.7	−10,779.3	−10,670.8	−10,862.2
	(26,603.1)	(26,982.1)	(26,981.7)	(26,931.9)
House Work	16,923.7	21,259.1	21,220.4	21,008.1
	(31,113.5)	(31,765.7)	(31,760.8)	(31,722.8)
Retired	16,729.7	16,143.8	16,261.0	15,597.9
	(35,168.0)	(35,546.6)	(35,533.4)	(35,445.1)
Constant	77,961.5	118,492.1	112,639.1	114,680.700
	(559,749.6)	(556,893.4)	(556,328.3)	(556,381.8)
R^2	0.211	0.209	0.210	0.210
Observations	568	568	568	568

Table A.6: CCEI Scores and Wealth

A8. Empirical Application: More on Halevy, Persitz, and Zrill (2018)

A8.1 Afriat efficiency indices

In Tables A.7 and A.8 we display pass rates by efficiency threshold and confidence intervals on preference types for the Halevy, Persitz, and Zrill (2018) data, as we did for the Choi *et al.* (2007) data in the main text. These tables confirm the observations we have already made in the main text: (1) more than half of subjects who pass GARP are also compatible with the EU model; (2) whether or not concavity is imposed on the Bernoulli function, the rank-dependent model explains only a modest fraction of the population not consistent with the EU model; and (3) the pass rates of the parametric versions of the rank-dependent and expected utility models have distinctly lower pass rates, suggesting a high level of parametric mis-specification.

	$e = 0.90$	$e = 0.95$	$e = 1.00$
GARP	194/207 (94%)	178/207 (86%)	92/207 (44%)
F-GARP	170/207 (82%)	155/207 (75%)	63/207 (30%)
RDU	170/207 (82%)	155/207 (75%)	62/207 (30%)
EU	170/207 (82%)	153/207 (74%)	59/207 (29%)
cRDU	152/207 (73%)	114/207 (55%)	26/207 (13%)
cEU	151/207 (73%)	103/207 (50%)	18/207 (9%)
RDU-CRRA	120/207 (58%)	78/207 (38%)	7/207 (3%)
EU-CRRA	94/207 (45%)	45/207 (22%)	7/207 (3%)

Table A.7: Pass rates by efficiency threshold (Halevy, Persitz, and Zrill, 2018)

	$e = 0.9$		$e = 0.95$	
	Sample Prop.	Conf. Interval	Sample Prop.	Conf. Interval
cEU	151/207 (0.729)	[0.664, 0.789]	103/207 (0.498)	[0.428, 0.568]
EU	170/207 (0.821)	[0.762, 0.871]	153/207 (0.739)	[0.674, 0.798]
EU\cEU	19/207 (0.092)	[0.056, 0.140]	50/207 (0.242)	[0.185, 0.306]
RDU\EU	0/207 (0.000)	[0.000, 0.018]	2/207 (0.010)	[0.001, 0.034]
cRDU\cEU	1/207 (0.005)	[0.000, 0.027]	11/207 (0.053)	[0.027, 0.093]

Table A.8: Confidence intervals on preference types (Halevy, Persitz, and Zrill, 2018)

A8.2 Varian efficiency indices

In the analysis of the data from their experiments, Halevy, Persitz, and Zrill (2018) measured departures from exact rationalizability, not with Afriat's efficiency index, but with Varian's inconsistency index. As we have pointed out in Section II of the main text, calculating Varian's index is computationally a lot more demanding than Afriat's index, because one would have to search across all efficiency vectors $\mathbf{e}$. Since Halevy, Persitz, and Zrill (2018) only calculated Varian indices for GARP and the parametric models they consider, they did manage to calculate the indices for the vast majority of their subjects. (For a handful of subjects, Halevy, Persitz, and Zrill (2018) did not manage to calculate the index for GARP but only obtained an approximation.)

We have avoided the use of Varian's index, partly because it is still less commonly used than Afriat's index, but also because calculating this index for *all* the nonparametric models we consider is not computationally feasible. However, we did calculate Varian-type indices for GARP, F-GARP, and for the EU model, using a new algorithm described in Section A9. Our algorithm works very well for GARP and F-GARP, giving exact answers for all subjects in the Halevy, Persitz, and Zrill (2018) experiment. In the case of the EU model, it gives a good approximation in the form of upper and lower bounds on the index.

Recall that Varian's efficiency index is given by

$$\sup \left\{ 1 - \sqrt{\frac{\sum_{t=1}^{T}(1-e^t)^2}{T}} \ : \ \mathcal{O}(\mathbf{e}) \text{ is rationalizable by } \mathcal{U} \right\}.$$

It is clear from the formula that this index varies between 0 and 1, with 1 being the value when the subject is exactly $\mathcal{U}$-rationalizable. For a given data set and model $\mathcal{U}$, Afriat's index will always be lower than Varian's, since it is the solution to the maximization problem with the additional constraint that $e^1 = e^2 = \ldots = e^T$.

The distributions of the Varian indices for different models are depicted in Figure A.8. The distributions for GARP, F-GARP, and EU are obtained from our calculations. The distributions for the parametric models – RDU-CRRA and EU-CRRA – are based on the numbers reported in Halevy, Persitz, and Zrill (2018). Note that in the case of the EU model, we assume that each subject's Varian index is the midpoint between the upper and lower

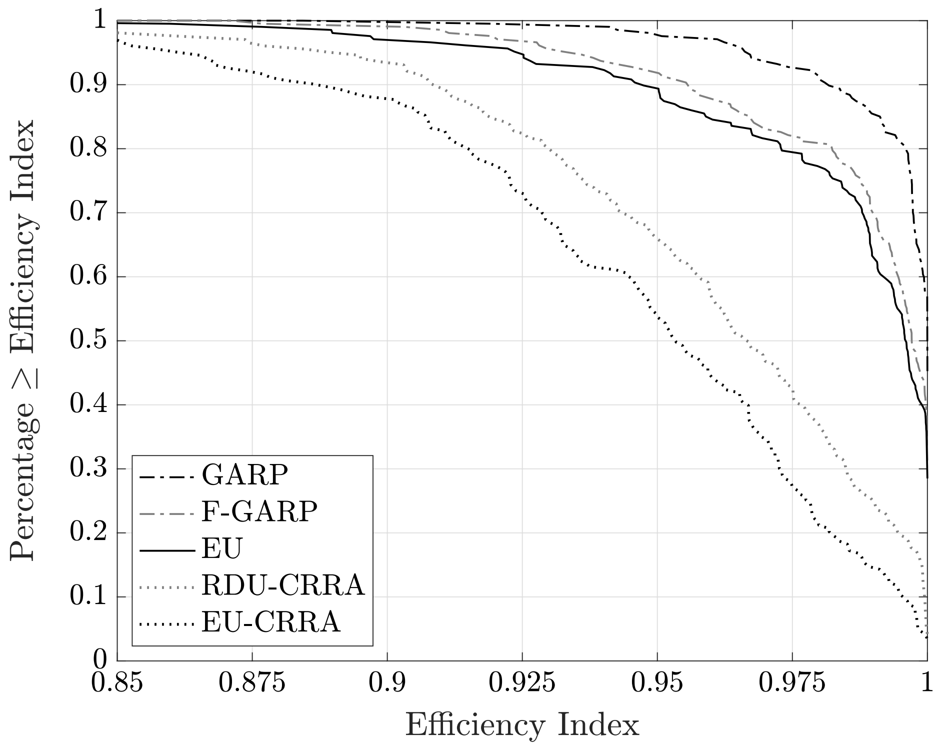


Figure A.8: Varian efficiency distributions (Halevy *et al.*, 2018)

bounds of the index. For the vast majority of subjects, the EU estimate we obtain is a good one: the difference between upper and lower bounds has a mean of 0.007 and a median lower than 0.001, with 90% of subjects having a difference smaller than 0.0253.

Comparing Figure A.8 with Figure 8 in the main text, one would notice that the Varian indices are significantly higher than the Afriat indices for the same model. However, certain qualitative features of the data are notably similar. In particular, (1) the nonparametric EU model performs well, with well over half of subjects who pass GARP at some threshold (see

	$e = 0.95$	$e = 0.99$	$e = 1.00$
GARP	202/207 (98%)	176/207 (85%)	92/207 (44%)
F-GARP	190/207 (92%)	144/207 (70%)	63/207 (30%)
EU (AVG)	185/207 (89%)	130/207 (63%)	59/207 (29%)
EU (LB)	178/207 (86%)	127/207 (61%)	59/207 (29%)
RDU-CRRA	133/203 (66%)	50/203 (25%)	7/203 (3%)
EU-CRRA	109/203 (54%)	29/203 (14%)	7/203 (3%)

Table A.9: Pass rates by Varian efficiency threshold (Halevy *et al.*, 2018)

	Corr.	Rank Corr.
GARP	0.989	0.999
F-GARP	0.975	0.998
EU(LB)	0.890	0.986
RDU-CRRA	0.966	0.982
EU-CRRA	0.961	0.974

Table A.10: Afriat/Varian correlations (Halevy, Persitz, and Zrill, 2018)

Table A.9[17]) being EU-rationalizable as well. (2) Indeed the distribution for the EU model is close to that for F-GARP which it cannot, by definition, surpass. Since the RDU model respects first order stochastic dominance, its distribution (which we did not calculate) must lie between the F-GARP and EU curves. In other words, echoing the analysis in the main text using Afriat's index, it appears that the RDU model does not explain the behavior of a significant number of subjects not explained by the EU model. (3) The imposition of a parametric form lowers pass rates quite sharply, suggesting that parametric misspecification is significant. This observation echoes our analysis of the same data using Afriat's index in Section IV.A of the main text.

Lastly, we calculate the correlation and rank correlation between the Afriat and Varian indices for the different models under consideration. There is a very high correlation between the two ways of measuring inconsistency (see Table A.10) suggesting that, at least for certain types of analysis, conclusions may not be sensitive to the index used.

A9. A new algorithm for calculating Varian's efficiency index

As we explained in Section II of the main paper, calculating Varian's index is computationally challenging and, in the case of the locally nonsatiated utility (LNU) model, it is known to be an NP hard problem (Smeulders *et al.*, 2014). In this section, we describe a new algorithm for calculating Varian's efficiency index for the LNU (GARP), SMU (F-GARP) and EU models. As far as we know there are no available algorithms for calculating Varian's index in the case of the SMU and EU models. For the LNU model, some other methods for

[17] There are only 203 subjects for RDU-CRRA and EU-CRRA (as opposed to 207 for the other tests) because results for those two models are taken from Halevy, Persitz, and Zrill (2018) where four of the subjects were dropped from the study. Note that EU (LB) denotes the lower bound on Varian's index and EU (AVG) the midpoint of the upper and lower bounds.

calculating Varian's index are available; we discuss their relationship with ours in Section A9.8. The Varian indices reported in Section A8 are calculated using the algorithm described here.

A9.1 The relaxed Varian efficiency problem

Let $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^T$ be a classical data set. As usual, we denote its modification by the efficiency vector $\mathbf{e} \in [0,1]^T$ by $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t)\}_{t=1}^T$, where $B^t(e^t)$ is given by (14). Given a collection $\mathcal{U}$ of utility functions defined on $\mathbb{R}_+^{\bar{s}}$, we define

$$\mathbf{E} = \{\mathbf{e} \in [0,1]^T : \mathcal{O}(\mathbf{e}) \text{ is } \mathcal{U}\text{-rationalizable}\}.$$

This set is nonempty since, for any family $\mathcal{U}$, it follows immediately from the definition of rationalizability that $(0, 0, \ldots, 0) \in \mathbf{E}$. In the case where a data set $\mathcal{O}$ is (exactly) $\mathcal{U}$-rationalizable, $(1, 1, \ldots, 1) \in \mathbf{E}$. Finally, note that this set is *downward comprehensive* in the sense that if $\mathbf{e} \in \mathbf{E}$, then $\mathbf{e}' < \mathbf{e}$ is also in $\mathbf{E}$.

Varian proposes to measure the degree of a data set $\mathcal{O}$'s inconsistency with a particular model, as captured by some collection of utility functions $\mathcal{U}$, by looking at the Euclidean distance between $(1, 1, \ldots, 1)$ and $\mathbf{E}$. We generalize the measure proposed by Varian slightly. We refer to a function $f : [0,1]^T \to \mathbb{R}$ such that $f(1,1,...,1) = 1$, with f continuous and increasing, as an *efficiency index*. Our objective is to find a way of solving the following problem, which we shall refer to as the *Varian efficiency problem*:

$$\sup f(\mathbf{e}) \text{ subject to } \mathbf{e} \in \mathbf{E}. \tag{a.33}$$

Note that this problem includes as a special case the measure proposed by Varian, which could be thought of as the case where

$$f(\mathbf{e}) = 1 - \sqrt{\frac{\sum_{t=1}^T (1 - e^t)^2}{T}}.$$

Afriat's critical cost efficiency index is also covered; in that case

$$f(\mathbf{e}) = \max\{e^1, e^2, \ldots, e^T\}.$$

Instead of solving the Varian efficiency problem, it is convenient to examine instead a related problem where the constraint on $\mathbf{e}$ is relaxed. We say that $\mathcal{O}(\mathbf{e})$ is *almost $\mathcal{U}$-rationalizable* if there is a sequence of vectors $\mathbf{e}_n$ with the following properties:

(i) $\mathbf{e}_n$ converges to $\mathbf{e}$;

(ii) $\mathbf{e}_n$ is *strictly below* $\mathbf{e}$ in the sense that $e_n^t \leqslant e^t$ and $e_n^t < e^t$ whenever $e^t > 0$; and

(iii) $\mathcal{O}(\mathbf{e}_n)$ is $\mathcal{U}$-rationalizable.

We denote by $\bar{\mathbf{E}}$ the set of efficiency vectors where $\mathcal{O}(\mathbf{e})$ is almost $\mathcal{U}$-rationalizable. Obviously this set is weakly larger than $\mathbf{E}$. By the *relaxed Varian efficiency problem*, we refer to the following:

$$\sup f(\mathbf{e}) \text{ subject to } \mathbf{e} \in \bar{\mathbf{E}}. \tag{a.34}$$

The advantage of this formulation is that the set $\bar{\mathbf{E}}$ is somewhat better behaved than $\mathbf{E}$. Firstly, it is also downward comprehensive. Indeed suppose $\mathbf{e} \in \bar{\mathbf{E}}$ and let $\mathbf{e}' < \mathbf{e}$. Let $\mathbf{d}_n$ be any sequence of vectors strictly below $\mathbf{e}'$ and converging towards $\mathbf{e}'$. Since $\mathbf{e} \in \bar{\mathbf{E}}$ there is another sequence of vectors $\mathbf{e}_n \in \mathbf{E}$, strictly below $\mathbf{e}$ and converging to $\mathbf{E}$ such that $\mathcal{O}(\mathbf{e}_n)$ is $\mathcal{U}$-rationalizable. Then the sequence $\mathbf{c}_n$, where $c_n^t = \min\{e_n^t, d_n^t\}$ is strictly below $\mathbf{e}'$ and converging towards $\mathbf{e}'$; furthermore, because $\mathbf{E}$ is downward comprehensive, $\mathcal{O}(\mathbf{c}_n)$ is $\mathcal{U}$-rationalizable. So we have shown that $\mathbf{e}' \in \bar{\mathbf{E}}$. Secondly, $\bar{\mathbf{E}}$ is a closed (and therefore, compact) set. Suppose the sequence $\mathbf{e}_n$ in $\bar{\mathbf{E}}$ converges to $\mathbf{e}$. Then the sequence $\mathbf{d}_n$, where $d_n^t = \min\{e_n^t, e^t\}$ also converges to $\mathbf{e}$, with $\mathbf{d}_n \leqslant \mathbf{e}$; furthermore, since $\bar{\mathbf{E}}$ is downward comprehensive, $\mathbf{d}_n \in \bar{\mathbf{E}}$ for all n. Obviously, we can then choose another sequence $\mathbf{g}_n$ in $\mathbf{E}$, such that $\mathbf{g}_n$ is strictly below $\mathbf{d}_n$ for each n (hence strictly below $\mathbf{e}$) and converging towards $\mathbf{e}$, which shows that $\mathbf{e} \in \bar{\mathbf{E}}$.

The following result is an immediate consequence of these observations; note that it holds whatever the definition of $\mathcal{U}$.

Lemma A.1. *There is an efficiency vector $\mathbf{e}^* \in [0,1]^T$ that solves the relaxed Varian efficiency problem (a.34), and its value coincides with that of the Varian efficiency problem (a.33).*

Proof. The existence of $\mathbf{e}^*$ is guaranteed since f is continuous and $\bar{\mathbf{E}}$ is a compact set. Then there is a sequence $\mathbf{e}_n \in \mathbf{E}$ that converges to $\mathbf{e}^*$. By the continuity of f, $f(\mathbf{e}_n)$ tends to $f(\mathbf{e}^*)$. So the problems (a.33) and (a.34) have the same value. QED

We now narrow our attention to the case where $\mathcal{U}$ is one of three cases: where it represents the family of (i) locally nonsatiated utility (LNU) functions, (ii) stochastically monotone utility (SMU) functions, and (iii) expected utility (EU) functions. We shall establish that in each of these cases, we can find the solution to the relaxed Varian efficiency problem, and thus Varian efficiency problem, by confining our search to a *finite* subset of $\bar{\mathbf{E}}$. In other words, one could (in principle) always solve the Varian efficiency problem by working out the values of $f(\mathbf{e})$ for different $\mathbf{e}$ in that finite set. As a practical matter, finding that subset in $\bar{\mathbf{E}}$ also requires us to check whether $\mathcal{O}(\mathbf{e})$ is *almost* $\mathcal{U}$*-rationalizable*, so we shall also address that issue, in the case where $\mathcal{U}$ belongs to any one of the three cases we listed.

We know (see Section A5) that $\mathcal{O}(\mathbf{e})$ is LNU-rationalizable if and only if it obeys GARP. There is a closely related necessary and sufficient condition for $\mathcal{O}(\mathbf{e})$ to be almost LNU-rationalizable.

A test for almost LNU-rationalizability Recall that for $\mathbf{x}^t$ and $\mathbf{x}^{t'}$ in $\mathcal{D} = \{\mathbf{x}^t : t = 1, 2, \ldots, T\}$, $\mathbf{x}^t$ is strictly revealed preferred to $\mathbf{x}^{t'}$ (denoted by $\mathbf{x}^t \succ^*_{\mathbf{e}} \mathbf{x}^{t'}$) if $\mathbf{p}^t \cdot \mathbf{x}^{t'} < e^t\, \mathbf{p}^t \cdot \mathbf{x}^t$. We claim that $\mathcal{O}(\mathbf{e})$ is almost rationalizable by a locally nonsatiated utility function if and only if $\succ^*_{\mathbf{e}}$ has no cycles, i.e., there does not exist observations $(\mathbf{p}^{t_i}, \mathbf{x}^{t_i})$ (for $i = 1, 2, \ldots, n$) in $\mathcal{O}$ such that

$$\mathbf{x}^{t_1} \succ^*_{\mathbf{e}} \mathbf{x}^{t_2},\ \mathbf{x}^{t_2} \succ^*_{\mathbf{e}} \mathbf{x}^{t_3},\ \ldots\ ,\ \mathbf{x}^{t_{n-1}} \succ^*_{\mathbf{e}} \mathbf{x}^{t_n}, \text{ and } \mathbf{x}^{t_n} \succ^*_{\mathbf{e}} \mathbf{x}^{t_1} \tag{a.35}$$

It is clear that this condition is necessary. To see that it is sufficient, notice that (crucially) $\mathcal{D}$ is a finite set. Thus, we can always choose a sequence $\mathbf{d}_n$ strictly below $\mathbf{e}$ and tending towards $\mathbf{e}$ such that the only revealed preference relations in $\mathcal{O}(\mathbf{d}_n)$ are strict revealed preference relations. (Note that, for any vector $\mathbf{c}$, $\mathcal{O}(\mathbf{c})$ could have non-strict revealed preference relations only if, for some $\mathbf{x}^t$ and $\mathbf{x}^{t'}$, we have $\mathbf{p}^t \cdot \mathbf{x}^{t'} = c^t \mathbf{p}^t \cdot \mathbf{x}^t$; we could always choose c^t to avoid this since $\mathcal{D}$ is a finite set.) When $\mathcal{O}(\mathbf{d}_n)$ has only strict revealed preference relations, GARP would coincide with the absence of cycles involving these strict relations, and so we know that $\mathcal{O}(\mathbf{d}_n)$ is LNU-rationalizable. We conclude that $\mathcal{O}(\mathbf{e})$ is almost LNU-rationalizable.

Let A^t be a set in $[0, 1]$ defined in the following way:

$$a \in A^t \text{ if } \mathbf{p}^t \cdot \mathbf{x} = a\mathbf{p}^t \cdot \mathbf{x}^t \text{ for some } \mathbf{x} \in \mathcal{D}. \tag{a.36}$$

Note by definition $1 \in A^t$ and that A^t is finite since $\mathcal{D}$ is finite. The next result says that the solution to (a.34) can be obtained by searching through the set

$$\mathbf{A} = A^1 \times A^2 \times \ldots \times A^T.$$

PROPOSITION A.7. *Suppose $\mathcal{U}$ is the collection of LNU functions. Then $\mathbf{e}^*$ is a solution to the relaxed Varian efficiency problem (a.34) if $f(\mathbf{e}^*) \geqslant f(\mathbf{e})$ for all $\mathbf{e} \in \mathbf{A} \cap \bar{\mathbf{E}}$, where A^t is given by (a.36) and $\bar{\mathbf{E}}$ is the set of efficiency vectors $\mathbf{e}$ that render $\mathcal{O}(\mathbf{e})$ almost LNU-rationalizable.*

Proof. Suppose that relaxed Varian efficiency problem (a.34) is solved at $\hat{\mathbf{e}} = (\hat{e}^1, \hat{e}^2, \ldots, \hat{e}^T)$ and for some $t = t'$, we have $\hat{e}^{t'} \notin A^{t'}$. Then there must be $\bar{a}$ in $A^{t'}$ such that $\hat{e}^{t'} < \bar{a}$. Observe that $\tilde{\mathbf{e}} \in \bar{\mathbf{E}}$, where $\tilde{e}^{t'} = \bar{a}$ and $\tilde{e}^t = \hat{e}^t$ for all $t \neq t'$. This is because, the increase in the vector from $\hat{\mathbf{e}}$ to $\tilde{\mathbf{e}}$ does not alter the set of revealed preference relations generated at observation t and, more generally, $\mathcal{O}(\tilde{\mathbf{e}})$ and $\mathcal{O}(\hat{\mathbf{e}})$ induce the same strict revealed preference relations. Thus $\mathcal{O}(\tilde{\mathbf{e}})$, like $\mathcal{O}(\hat{\mathbf{e}})$, is almost LNU-rationalizable. Since f is increasing, $\tilde{\mathbf{e}}$ must also solve (a.34). Repeating this procedure if necessary, we eventually end up with an efficiency vector in $\mathbf{A}$ that solves (a.34). QED

This proposition gives a potentially practical way of solving the relaxed Varian efficiency problem (and thus the Varian efficiency problem). We go through the elements of $\mathbf{A}$ and, for each element, check if it is in $\bar{\mathbf{E}}$ by implementing the relaxed version of the GARP test. If it is work out $f(\mathbf{e})$; the highest value obtained in this manner will solve the Varian efficiency problem.

Suppose that $A^t = \{a_1, a_2, a_3, \ldots, a_{m(t)}\}$, with

$$1 = a_1 > a_2 > a_3 > \ldots > a_{m(t)} \geqslant 0.$$

Each value in A^t is associated with a constraint set $B^t(a_j)$ and hence a certain number of strict revealed preference relations between $\mathbf{x}^t$ and elements of $\mathcal{D}$ in $B^t(a_j)$. The values of A^t are precisely those values at which the number of such relations drop. Thus the choice of $a_j \in A^t$ could be thought of as a choice over the number of strict revealed preference relations to ignore. For instance, at $a_1 = 1$, we are retaining all the strict revealed preference relations in the sense that $\mathbf{x}^t$ is strictly revealed preferred to every element $\mathbf{x} \in \mathcal{D}$ for which

$\mathbf{p}^t \cdot \mathbf{x} < \mathbf{p}^t \cdot \mathbf{x}^t$; on the other hand, at $a = a_2$, $\mathbf{x}^t$ is strictly revealed preferred to every element $\mathbf{x} \in \mathcal{D}$ such that $\mathbf{p}^t \cdot \mathbf{x} < a_2 \mathbf{p}^t \cdot \mathbf{x}^t$, which means that we are removing those elements $\mathbf{x} \in \mathcal{D}$ (at least one and possibly more) for which $\mathbf{p}^t \cdot \mathbf{x} = a_2 \mathbf{p}^t \cdot \mathbf{x}^t$. There is a one-to-one map between the elements of A^t and the number of elements in $\mathcal{D}$ which are removed as strict revealed preference relations. We define $C^t = \{c_1, c_2, ..., c_{m(t)}\}$, such that

$$0 = c_1 < c_2 < \ldots < c_{m(t)},$$

the following manner: for $j \geqslant 1$, c_j is the difference in the number of strict revealed preference relations generated at observation t between $a = a_1 = 1$ and $a = a_j$. We denote this one-to-one and onto map from C^t and A^t by ϕ. For example, suppose $A^t = \{1, 0.9, 0.6\}$ and $C^t = \{0, 2, 3\}$. This means that there are exactly three elements $\mathbf{x} \in \mathcal{D}$ such that $\mathbf{p}^t \cdot \mathbf{x} < \mathbf{p}^t \cdot \mathbf{x}^t$, with two elements satisfying $\mathbf{p}^t \cdot \mathbf{x} = 0.9\mathbf{p}^t \cdot \mathbf{x}^t$ and one element satisfying $\mathbf{p}^t \cdot \mathbf{x} = 0.6\mathbf{p}^t \cdot \mathbf{x}^t$. So as we 'go' from $a = 1$ to $a = 0.9$ to $a = 0.6$, the budget shrinks at each step, and number of revealed preference relations removed are 0, 2, and finally 3.

Defining $\mathbf{C} = C^1 \times C^2 \times \ldots C^T$, we can associate to each element $\mathbf{c} \in \mathbf{C}$, the element

$$\Phi(\mathbf{c}) = (\phi^1(c^1), \phi^2(c^2), \ldots, \phi^T(c^T)) \in \mathbf{A}.$$

Abusing notation, we shall write $f(\mathbf{c})$ to mean $f(\Phi(\mathbf{c}))$. Note that $f(\mathbf{c})$ is *decreasing in* $\mathbf{c}$ because f is increasing in the efficiency vector. We shall refer to $\mathbf{c}$ as being almost LNU-rationalizable when we mean that $\mathcal{O}(\Phi(\mathbf{c}))$ is almost LNU-rtionalizable. We shall also refer to $\mathbf{c}$ as being in $\bar{\mathbf{E}}$ when we mean that $\Phi(\mathbf{c})$ is in $\bar{\mathbf{E}}$. The problem of solving the Varian efficiency problem can be understood as follows:

$$\max f(\mathbf{c}) \text{ subject to } \mathbf{c} \in \mathbf{C} \text{ being almost LNU-rationalizable.}$$

A9.3 The Varian efficiency problem for SMU functions

In the previous section, we explained that SMU-rationalizability can be characterized by F-GARP. A closely related test could be used to check for almost SMU-rationalizability of some modified data set $\mathcal{O}(\mathbf{e})$.

A test for almost SMU-rationalizability Recall that for $\mathbf{x}^t$ and $\mathbf{x}^{t'}$ in $\mathcal{D}$, we had defined $\mathbf{x}^t$ as being strictly revealed preferred to $\mathbf{x}^{t'}$ if there is $\mathbf{y} \in B^t(e^t)$ such that either

$\mathbf{p}^t \cdot \mathbf{y} < e^t \mathbf{p}^t \cdot \mathbf{x}^t$ or $\mathbf{p}^t \cdot \mathbf{y} = e^t \mathbf{p}^t \cdot \mathbf{x}^t$ and $\mathbf{y} >_{FSD} \mathbf{x}^t$, with *either* inequality strict. For testing almost LNU-rationalizability, only one of these two types of strict revealed preference relations are relevant. We define $\mathbf{x}^t$ as being *type 1 strictly revealed preferred* to $\mathbf{x}^{t'}$ if there is $\mathbf{y}$ such that $\mathbf{p}^t \cdot \mathbf{y} < e^t \mathbf{p}^t \cdot \mathbf{x}^t$ and $\mathbf{y} \geqslant_{FSD} \mathbf{x}^t$. (In fact, if $\mathbf{y}$ exists then we can always choose it to be a permutation of $\mathbf{x}^t$.) We claim that $\mathcal{O}(\mathbf{e})$ is almost LNU-rationalizable if and only if the type 1 strict revealed preference relations admit no cycles; in other words, we cannot find observations $(\mathbf{p}^{t_i}, \mathbf{x}^{t_i})$ (for $i = 1, 2, \ldots, n$) in $\mathcal{O}$ such that we have a cycle like (a.35), interpreting $>^*_{\mathbf{e}}$ to be a type 1 strict revealed preference relation. It is clear that this condition is necessary. Indeed, if a type 1 strict revealed preference cycle exists, then for *any* sequence $\mathbf{d}_n$ strictly below $\mathbf{e}$ and tending towards $\mathbf{e}$, $\mathcal{O}(\mathbf{d}_n)$ will violate F-GARP when n is sufficiently large. Hence $\mathcal{O}(\mathbf{d}_n)$ will not be SMU-rationalizable and we conclude that $\mathcal{O}(\mathbf{e})$ is not almost SMU-rationalizable. To see that this condition is also sufficient, let $\mathcal{G}^t$ be the following set of vectors:

$$\mathbf{y} \in \mathcal{G}^t \text{ if } \mathbf{p}^t \cdot \mathbf{y} \leqslant \mathbf{p}^t \cdot \mathbf{x}^t \text{ and, for some } \mathbf{x}' \in \mathcal{D}, \mathbf{y} \text{ solves } \min \mathbf{p}^t \cdot \mathbf{y}' \text{ subject to } \mathbf{y}' \geqslant_{FSD} \mathbf{x}'.$$

In words, $\mathbf{y} \in \mathcal{G}^t$ if it is in the budget at observation t, it dominates some element in $\mathbf{x}' \in \mathcal{D}$, and (evaluated by $\mathbf{p}^t$) it is the cheapest bundle to dominate $\mathbf{x}'$. Hence we can always choose a sequence $\mathbf{d}_n$ strictly below $\mathbf{e}$ and tending towards $\mathbf{e}$ such that d^t_n does *not* satisfy $\mathbf{p}^t \cdot \mathbf{y} = d^t_n \mathbf{p}^t \cdot \mathbf{x}^t$ for some $\mathbf{y} \in \mathcal{G}^t$ (since $\mathcal{D}$ is a finite set). In this way, the only revealed preference relations in $\mathcal{O}(\mathbf{d}_n)$ are type 1 strict revealed preference relations. Since there are no cycles with these relations, $\mathcal{O}(\mathbf{d}_n)$ obeys F-GARP and is thus SMU-rationalizable. We conclude that $\mathcal{O}(\mathbf{e})$ is almost SMU-rationalizable.

Let A^t be a set in $[0, 1]$ defined in the following way:

$$a \in A^t \text{ if } a = 1 \text{ or } \mathbf{p}^t \cdot \mathbf{y} = a\mathbf{p}^t \cdot \mathbf{x}^t \text{ for some } \mathbf{y} \in \mathcal{G}^t. \tag{a.37}$$

A^t is finite since $\mathcal{D}$ is finite. The next result says that the solution to (a.34) can be obtained by searching through the set $\mathbf{A} = A^1 \times A^2 \times \ldots \times A^T$. The proof is closely analogous to that for Proposition A.7.

PROPOSITION A.8. *Suppose $\mathcal{U}$ is the collection of SMU functions. Then $\mathbf{e}^*$ is a solution to the relaxed Varian efficiency problem (a.34) if $f(\mathbf{e}^*) \geqslant f(\mathbf{e})$ for all $\mathbf{e} \in \mathbf{A} \cap \bar{\mathbf{E}}$, where A^t*

is given by (a.37) and $\bar{\mathbf{E}}$ *is the set of efficiency vectors* $\mathbf{e}$ *that render* $\mathcal{O}(\mathbf{e})$ *almost SMU-rationalizable.*

Proof. Suppose that the relaxed Varian efficiency problem (a.34) is solved at $\hat{\mathbf{e}} = (\hat{e}^1, \hat{e}^2, \ldots, \hat{e}^T)$ and for some $t = t'$, we have $\hat{e}^{t'} \notin A^{t'}$. Then there must be $\bar{a}$ in $A^{t'}$ such that $\hat{e}^{t'} < \bar{a}$. Observe that $\tilde{\mathbf{e}} \in \bar{\mathbf{E}}$, where $\tilde{e}^{t'} = \bar{a}$ and $\tilde{e}^t = \hat{e}^t$ for all $t \neq t'$. This is because the absence of any $\mathbf{y} \in \mathcal{G}^t$ with $\mathbf{p}^t \cdot \mathbf{y}$ lying in the interval $(\hat{e}^{t'}, \bar{a})$ means that increase in the efficiency vector from $\hat{\mathbf{e}}$ to $\tilde{\mathbf{e}}$ does not alter the set of revealed preference relations generated at observation t (and more generally, at all observations). Thus $\mathcal{O}(\tilde{\mathbf{e}})$, like $\mathcal{O}(\hat{\mathbf{e}})$, is almost SMU-rationalizable. Since f is increasing, $\tilde{\mathbf{e}}$ must also solve (a.34). Repeating this procedure if necessary, we eventually end up with an efficiency vector in $\mathbf{A}$ that solves (a.34). QED

As in the LNU case, it is convenient to think of the process of choosing an element in A^t with choosing the number of type 1 strict revealed preference relations to drop. In this way (analogous to our discussion in the LNU case) there is a one-to-one map between A^t and the set C^t, with the latter keeping count of the number of revealed preference relations dropped as we move towards lower values of A^t. An example should suffice to explain what we mean. Suppose $A^t = \{1, 0.9, 0.6\}$ and $C^t = \{0, 2, 3\}$. This means that there are exactly three elements $\mathbf{x} \in \mathcal{D}$ to which $\mathbf{x}^t$ is type 1 strictly revealed preferred when $e^t = 1$ (that is, before the budget set is shrunk). Suppose the elements are $\mathbf{x}'$, $\mathbf{x}''$, and $\mathbf{x}'''$. When $e^t = 0.9$, two of those elements are removed, in the sense that $\mathbf{x}^t$ is no longer type 1 strictly revealed preferred to two of three elements. Suppose those elements are $\mathbf{x}'$ and $\mathbf{x}''$ (with $\mathbf{x}'''$ remaining). This means there is $\mathbf{y}'$ such that $\mathbf{y}' \cdot \mathbf{p}^t = 0.9\mathbf{x}^t \cdot \mathbf{p}^t$ and $\mathbf{y}'$ minimizes $\mathbf{p}^t \cdot \mathbf{y}$ among those $\mathbf{y}$ satisfying $\mathbf{y} \geqslant_{FSD} \mathbf{x}'$. Similarly, there is $\mathbf{y}''$ such that $\mathbf{y}'' \cdot \mathbf{p}^t = 0.9\mathbf{x}^t \cdot \mathbf{p}^t$ and $\mathbf{y}''$ minimizes $\mathbf{p}^t \cdot \mathbf{y}$ among those $\mathbf{y}$ satisfying $\mathbf{y} \geqslant_{FSD} \mathbf{x}''$. Finally, when $e^t = 0.6$, $\mathbf{x}^t$ is no longer type 1 strictly revealed preferred to any element in $\mathcal{D}$, so there is $\mathbf{y}'''$ such that $\mathbf{y}'' \cdot \mathbf{p}^t = 0.6\mathbf{x}^t \cdot \mathbf{p}^t$ and $\mathbf{y}'''$ minimizes $\mathbf{p}^t \cdot \mathbf{y}$ among those $\mathbf{y}$ satisfying $\mathbf{y} \geqslant_{FSD} \mathbf{x}'''$. There is a one-to-one map between the elements of $\mathbf{C}$ and of $\mathbf{A}$ and thus (analogous to the LNU case) the problem of solving the Varian efficiency problem can be understood as follows:

$$\max f(\mathbf{c}) \text{ subject to } \mathbf{c} \in \mathbf{C} \text{ being almost SMU-rationalizable.}$$

The test we developed for EU-rationalizability can be readily modified to test for almost EU-rationalizability.

A test for almost EU-rationalizability We know that $\mathcal{O}(\mathbf{e})$ is EU-rationalizable if and only if there is a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ for which $\sum_{s=1}^{\bar{s}} \pi_s^t \bar{u}(x_s^t) \geqslant \sum_{s=1}^{\bar{s}} \pi_s^t \bar{u}(x_s)$ for all $\mathbf{x} \in B^t(e^t) \cap \mathcal{G}$, with the inequality being strict if $\mathbf{x} \in \mathcal{G}$ satisfies $\mathbf{p}^t \cdot \mathbf{x} < e^t \mathbf{p}^t \cdot \mathbf{x}^t$. We claim that this implies that $\mathcal{O}(\mathbf{e})$ is almost EU-rationalizable if and only if there is a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ satisfying

$$\sum_{s=1}^{\bar{s}} \pi_s^t \bar{u}(x_s^t) > \sum_{s=1}^{\bar{s}} \pi_s^t \bar{u}(x_s) \text{ for all } \mathbf{x} \in \mathcal{G} \text{ such that } \mathbf{p}^t \cdot \mathbf{x} < e^t \mathbf{p}^t \cdot \mathbf{x}^t. \tag{a.38}$$

In other words, the test is the same as that for EU-rationalizability except that only the strict revealed preference conditions at each observation are retained. This condition is clearly necessary. To see that it is sufficient, we define $\mathcal{H}^t$ and A^t in the following way:

$$\mathbf{y} \in \mathcal{H}^t \text{ if } \mathbf{p}^t \cdot \mathbf{y} \leqslant \mathbf{p}^t \cdot \mathbf{x}^t \text{ and } \mathbf{y} \in \mathcal{G}.$$

(in order words, $\mathcal{H}^t$ are the elements of $\mathcal{G}$ which are affordable at observation t) and

$$a \in A^t \text{ if } \mathbf{p}^t \cdot \mathbf{y} = a \mathbf{p}^t \cdot \mathbf{x}^t \text{ for some } \mathbf{y} \in \mathcal{H}^t. \tag{a.39}$$

Since $\mathbf{x}^t \in \mathcal{H}^t$, $1 \in A^t$. Furthermore, since $\mathcal{G}$ is finite, so is A^t. Thus, we can always choose a sequence $\mathbf{d}_n$ strictly below $\mathbf{e}$ and tending towards $\mathbf{e}$ such that $d_n^t \notin A^t$ for all n. In this way, the only revealed preference conditions required for $\mathcal{O}(\mathbf{d}_n)$ to be EU-rationalizable are the strict conditions, which for n sufficiently large, are identical with having a strictly increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ such that (a.38) holds. Thus $\mathcal{O}(\mathbf{d}_n)$ is EU-rationalizable and we conclude that $\mathcal{O}(\mathbf{e})$ is almost EU-rationalizable.

PROPOSITION A.9. *Suppose $\mathcal{U}$ is the collection of EU functions. Then $\mathbf{e}^*$ is a solution to the relaxed Varian efficiency problem (a.34) if $f(\mathbf{e}^*) \geqslant f(\mathbf{e})$ for all $\mathbf{e} \in \mathbf{A} \cap \bar{\mathbf{E}}$, where A^t is given by (a.39) and $\bar{\mathbf{E}}$ is the set of efficiency vectors $\mathbf{e}$ that render $\mathcal{O}(\mathbf{e})$ almost EU-rationalizable.*

Proof. Suppose that relaxed Varian efficiency problem (a.34) is solved at $\hat{\mathbf{e}} = (\hat{e}^1, \hat{e}^2, \ldots, \hat{e}^T)$ and for some $t = t'$, we have $\hat{e}^{t'} \notin A^{t'}$. Then there must be $\bar{a}$ in $A^{t'}$ such that $\hat{e}^{t'} < \bar{a}$. Observe that $\tilde{\mathbf{e}} \in \bar{\mathbf{E}}$, where $\tilde{e}^{t'} = \bar{a}$ and $\tilde{e}^t = \hat{e}^t$ for all $t \neq t'$. This is because, the increase in the vector from $\hat{\mathbf{e}}$ to $\tilde{\mathbf{e}}$ notwithstanding, $\mathcal{O}(\tilde{\mathbf{e}})$ and $\mathcal{O}(\hat{\mathbf{e}})$ induce the same set strict revealed conditions (a.38) since the set of elements in $\mathbf{x} \in \mathcal{G}$ such that $\mathbf{p}^{t'} \cdot \mathbf{x} < \hat{e}^{t'} \mathbf{p}^{t'} \cdot \mathbf{x}^{t'}$ is the same as the set of $\mathbf{x} \in \mathcal{G}$ such that $\mathbf{p}^{t'} \cdot \mathbf{x} < \bar{a} \mathbf{p}^{t'} \cdot \mathbf{x}^{t'}$. Thus $\mathcal{O}(\tilde{\mathbf{e}})$, like $\mathcal{O}(\hat{\mathbf{e}})$, is almost EU-rationalizable. Since f is increasing, $\tilde{\mathbf{e}}$ must also solve (a.34). Repeating this procedure if necessary, we eventually end up with an efficiency vector in $\mathbf{A}$ that solves (a.34). QED

As in the LNU and SMU cases, it is convenient to think of the process of choosing an element in A^t with choosing the number of strict revealed preference comparisons to drop. There is a one-to-one map between A^t and the set C^t, with the latter keeping count of the number of revealed preference relations removed from consideration as we move towards lower values of A^t. An example should suffice to explain what we mean. Suppose $A^t = \{1, 0.9, 0.6, 0\}$ and $C^t = \{0, 2, 3, 4\}$. This means that there are exactly four elements $\mathbf{x} \in \mathcal{G}$ such that $\mathbf{p}^t \cdot x < \mathbf{p}^t \cdot \mathbf{x}^t$. Suppose the elements are $\mathbf{x}'$, $\mathbf{x}''$, $\mathbf{x}'''$, and $\mathbf{0}$. (Note that, by construction, $\mathbf{0}$ is always in $\mathcal{G}$.) Then two of them – say $\mathbf{x}'$ and $\mathbf{x}''$ – must satisfy $\mathbf{p}^t \cdot \mathbf{x}' = \mathbf{p}^t \cdot \mathbf{x}'' = 0.9\mathbf{p}^t \cdot \mathbf{x}^t$, while $\mathbf{x}'''$ must satisfy $\mathbf{p}^t \cdot \mathbf{x}''' = 0.6\mathbf{p}^t \cdot \mathbf{x}^t$. When $a = 1$, the condition (a.38) gives rise to four revealed preference conditions (in other words, four inequalities) guaranteeing that $\mathbf{x}^t$ is superior to $\mathbf{x}'$, $\mathbf{x}''$, $\mathbf{x}'''$, and $\mathbf{0}$. When $a = 0.9$, two of these conditions are removed and we only require $\mathbf{x}^t$ is superior to be superior to $\mathbf{x}'$ and to $\mathbf{0}$. And so on. Thus we can identify elements of $\mathbf{C}$ with elements of $\mathbf{A}$, and thus the problem of solving the Varian efficiency problem can be understood as follows:

$$\max f(\mathbf{c}) \text{ subject to } \mathbf{c} \in \mathbf{C} \text{ being almost EU-rationalizable.}$$

A.9.5 Three ways of expediting the search through $\mathbf{C}$

To recap, we have shown that in the case of the LNU, SMU, and EU models, solving the Varian efficiency problem could be thought of as a search through some *finite* space $\mathbf{C}$ which specifies the number of revealed preference relations to drop at each observation. This does not in itself lead to a tractable way of solving the problem, since $\mathbf{C}$ can be a very large set, even if it is finite. We now discuss several ways in which we can speed up the search across the elements of $\mathbf{C}$; these shortcuts are applicable to all the three models under consideration.

SHORTCUT 1. *Does a subset fail the test?*

There is a one-to-one map between $e^t \in A^t$ and $c^t \in C^t$ and thus a one-to-one map between $(c^t)_{t\in T'}$ and $\{(\mathbf{x}^t, B^t(e^t)\}_{t\in T'}$ for any $(e^t)_{t\in T'} \in (A^t)_{t\in T'}$, where T' is a subset of observations. We say that $(c^t)_{t\in T'}$ is almost $\mathcal{U}$-rationalizable if its corresponding data set is almost $\mathcal{U}$-rationalizable. Note that a data set $\{(\mathbf{x}^t, B^t(e^t)\}_{t\in T}$ cannot be almost $\mathcal{U}$-rationalizable if a subset of its observations, $\{(\mathbf{x}^t, B^t(e^t)\}_{t\in T'}$ is not almost $\mathcal{U}$-rationalizable. In terms of the elements of $\mathbf{C}$, this means that $(\bar{c}^t)_{t\in T'}$ is not almost $\mathcal{U}$-rationalizable if and only if $\mathbf{c} \in \mathbf{C}$ is not in $\bar{\mathbf{E}}$ whenever $c^t = \bar{c}^t$ for $t \in T'$.

Illustrative data set. In all the examples considered in the rest of this section, we shall refer to a data set with three observations, where $C^1 = \{0, 1, 2\}$, $C^2 = \{0, 1\}$, and $C^3 = \{0, 1\}$.

As an example of Shortcut 1 at work, suppose that the first observation, when considered in isolation, is not almost $\mathcal{U}$-rationalizable. (This is impossible if $\mathcal{U}$ is the family of LNU functions, but it is possible if it is the family of SMU or EU functions (see Example 1 in Section I.B of the main text).) This means that any modified data set that is almost $\mathcal{U}$-rationalizable will involve the first budget set shrinking. Put another way, we know that $\mathbf{c} \notin \bar{\mathbf{E}}$ for all $\mathbf{c}$ such that $c^1 = 0$. Thus we can remove from consideration the following values of $\mathbf{c}$:

$$(0, 0, 0),\ (0, 1, 0),\ (0, 1, 1),\ (0, 0, 1).$$

Shortcut 2. *Can we still improve on the optimal value?*

This is best explained with our illustrative data set. Suppose we find that $f(1, 0, 0) = 0.9$, $f(0, 1, 0) = 0.7$ and that $(1, 0, 0) \in \bar{\mathbf{E}}$. Clearly then $(0, 1, 0)$ cannot be a solution to the relaxed Varian efficiency problem. In addition, because f is decreasing in $\mathbf{c}$, we may also remove the following from consideration:

$$(1, 1, 0),\ (0, 1, 1),\ (1, 1, 1),\ (2, 1, 1),\ (2, 1, 0)$$

since these vectors are greater than $(0, 1, 0)$.

Shortcut 3. *Is an observation involved in a violation?*

Suppose that at some $\mathbf{c}$ we find that $\mathbf{c} \notin \bar{\mathbf{E}}$. We say that observation t' is *not involved in a violation at* $\mathbf{c}$ if the following holds: if $\bar{\mathbf{c}} \in \bar{\mathbf{E}}$, then $\hat{\mathbf{c}} \in \bar{\mathbf{E}}$, where $\hat{c}^t = \bar{c}^t$ for all $t \neq t'$ and $\hat{c}^{t'} = c^{t'}$. In other words, if $\mathbf{c} \notin \bar{\mathbf{E}}$ then its entries will have to be altered for it to be in $\bar{\mathbf{E}}$ but there is no alteration that crucially involves changing (in particular raising) the t'th entry of $\mathbf{c}$. Thus we need not consider any alteration of $\mathbf{c}$ with a higher value for $c^{t'}$ because this is not crucial and could only lower the value of f. At $\mathbf{c} \in \mathbf{C}$, we define $V(\mathbf{c})$ as the subset of T such that $\tilde{t} \in V(\mathbf{c})$ if $\tilde{t}$ *is* involved in a violation at $\mathbf{c}$. Returning to our illustrative data set, if $\mathbf{c} = (1, 0, 0) \notin \bar{\mathbf{E}}$ but the first observation is not involved in the violation (i.e., $1 \notin V(1, 0, 0)$), then we need not consider those values of $\mathbf{c}$ that removes more revealed preference relations from observation 1, i.e., we need not consider

$$\{(2, 0, 0), (2, 1, 0), (2, 0, 1), (2, 1, 1)\}.$$

To implement this shortcut, what is needed is some way of determining whether some observation is not in $V(\mathbf{c})$. In all the three families of utility functions we are considering, there are simple sufficient conditions to guarantee that $t' \notin V(\mathbf{c})$.

In the case where $\mathcal{U}$ is the collection of LNU or SMU functions, it is clear that $t' \notin V(\mathbf{c})$ if none of the revealed preference relations generated by t' are part of some strict revealed preference cycle in the (unmodified) data set $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}$. This is an easy-to-check condition. In the case where $\mathcal{U}$ is the family of EU functions, there is a similar sufficient condition. Fishburn's result (see Section A1.2) tells us that $\{(\mathbf{x}^t, \mathbf{p}^t)\}$ is *not* almost EU-rationalizable if and only if there are weights $\lambda^j \geqslant 0$ with $\sum_{j=1}^{M} \lambda^j = 1$, such that

$$\sum_{j=1}^{M} \lambda^j g(\mathbf{a}^j) = \sum_{j=1}^{M} \lambda^j g(\mathbf{b}^j), \tag{a.40}$$

where $\{(g(\mathbf{a}^j), g(\mathbf{b}^j))\}_{j=1}^{M}$ is the complete set of strict revealed preference relations. Within this set are those elements generated by the observation t' if its budget is shrunk by the factor $e^{t'}$, i.e., the cases where $\mathbf{a}_j = \mathbf{x}^{t'}$ and $\mathbf{b}_j \in B^{t'}(e^{t'}) \cap \mathcal{G}$; let that subset of conditions be denoted by J'. If $t' \in V(\mathbf{c})$ then, by definition, there are $\hat{e}^t$ (for all $t \neq t'$) such that $\{(\mathbf{x}^t, B^t(\hat{e}^t))\}_{t \in T \setminus t'}$ is almost EU-rationalizable but $\{(\mathbf{x}^t, B^t(\hat{e}^t))\}_{t \in T \setminus t'} \cup \{(\mathbf{x}^{t'}, B^{t'}(e^{t'}))\}$ is *not* almost EU-rationalizable (where $\mathbf{e} = (e^t)_{t \in T}$ is the efficiency vector associated with $\mathbf{c}$). By Fishburn's condition, this implies that there are weights $\lambda^j \geqslant 0$ with $\sum_{j=1}^{M} \lambda^j = 1$, that solve (a.40) with $\sum_{j \in J'} \lambda_j > 0$. In other words, a sufficient condition for $t' \notin V(\mathbf{c})$ is the following: for any $\lambda_j \geqslant 0$ such that $\sum_{j=1}^{M} \lambda^j = 1$ and (a.40) holds, we have $\sum_{j \in J'} \lambda_j = 0$. This is equivalent to the following easy-to-check condition: the solution to $\max \sum_{j \in J'} \lambda_j$ subject to $\lambda^j \geqslant 0$, $\sum_{j=1}^{M} \lambda^j = 1$ and (a.40) is zero.

Note that the condition we have identified in each case of $\mathcal{U}$ is sufficient but not necessary to guarantee that $t' \notin V(\mathbf{c})$. This does not mean that in applying this shortcut we miss out on checking elements of $\mathbf{C}$ which should be checked (and potentially leading to a solution lower than the true solution), but it does mean that the algorithm checks some elements of $\mathbf{C}$ which a sharper criterion when applying Shortcut 2 would eliminate.[18]

A9.6 Example of the algorithm at work

Suppose we iterate over the elements of $\mathbf{C}$ in our illustrative data set in the following order:

$$(0,0,0),\ (1,0,0),\ (2,0,0),\ (2,1,0),\ (2,1,1),\ (2,0,1),\\ (1,1,0),\ (1,1,1),\ (1,0,1),\ (0,1,0),\ (0,1,1),\ (0,0,1) \quad \text{(a.41)}$$

It is unimportant to understand at the moment how this ordering was chosen. The only thing to note is that this list exhausts $\mathbf{C}$. Suppose that the following values of $\mathbf{c}$ are in $\bar{\mathbf{E}}$:

$$(2,1,0),\ (2,1,1),\ (2,0,1),\ (1,1,0),\ (1,1,1),\ (1,0,1). \quad \text{(a.42)}$$

Note that this implies that observation 1 is not involved in a violation at any $\mathbf{c} = (1, v, w)$ for $(v, w) \in C^2 \times C^3$. Suppose also that

$$f(2,1,1) < f(2,0,1) < f(2,1,0) < f(2,0,0) < f(1,1,1)\\ < f(0,1,1) < f(1,0,1) < f(1,1,0) < f(1,0,0) < f(0,1,0) < f(0,0,1) \quad \text{(a.43)}$$

which implies that $(1,1,0)$ maximizes $f(\mathbf{c})$ among elements of $\mathbf{c} \in \bar{\mathbf{E}}$.

The algorithm we propose proceeds as follows. The algorithm starts by creating a variable $s = \min f$. The algorithm then sets b equal to the first element in the list (a.41). That is, the algorithm sets $\mathbf{c} = (0,0,0)$. Next, we test to see if $(0,0,0) \in \bar{\mathbf{E}}$. We have assumed this

[18] For example, suppose we wish to solve the Varian efficiency problem in the LNU case, for a data set with three observations. Suppose also that $\mathbf{x}^1$ strictly revealed preferred to $\mathbf{x}^2$, $\mathbf{x}^2$ strictly revealed preferred to $\mathbf{x}^3$, $\mathbf{x}^3$ strictly revealed preferred to $\mathbf{x}^1$, and $\mathbf{x}^3$ strictly revealed preferred to $\mathbf{x}^2$, so there are two strict revealed preference cycles. In this case, observation 1 is not involved in a violation according to our definition, since the re-inclusion of that observation does not introduce a cycle so long as the revealed preference for $\mathbf{x}^2$ over $\mathbf{x}^3$, or vice versa is removed. However, the revealed preferences generated by observation 1 *are* part of a strict cycle.

is not the case (see (a.42)). Next, the algorithm sets $\mathbf{c} = (1, 0, 0)$ which is the next item on the list (a.41). We test if $(1, 0, 0) \in \bar{\mathbf{E}}$. We have assumed that it is not. Before going to the next element in the list in (a.41) the algorithm checks if observation 1 is involved in the violation at $(1, 0, 0)$. We have assumed that it is not. As this is the case there will be no optimum of the form $(2, v, w)$ for any v and w. The algorithm then skips over $(2, 0, 0)$, $(2, 1, 0)$, $(2, 1, 1)$, and $(2, 0, 1)$. So the algorithm goes on to the next element in the list which is $\mathbf{c} = (1, 1, 0)$. The algorithm will then confirm that $(1, 1, 0) \in \bar{\mathbf{E}}$. The algorithm then writes $s = f(1, 1, 0)$. The algorithm does not consider $(1, 1, 1)$ as $f(1, 1, 1) < f(1, 1, 0)$ because f is strictly decreasing. (This is an implementation of Shortcut 2.) Next, the algorithm sets $\mathbf{c} = (1, 0, 1)$. Before testing if it is in $\bar{\mathbf{E}}$ the algorithm observes that $s > f(1, 0, 1)$ and skips to $\mathbf{c} = (0, 1, 0)$ (which is again Shortcut 2 at work). The algorithm then determines that $c^1 = 0$ is not almost $\mathcal{U}$-rationalizable; applying Shortcut 1, it skips the remaining elements in the list in (a.41).

We next give the pseudocode to express the general idea of how the algorithm functions (see Algorithm 1). This code will terminate having saved the optimal value for the VEP in the variable s. It can be easily checked that this code will reproduce the flow expressed in a.41.

Because the algorithm can run faster or slower depending on the order of the observations we modify the above algorithm slightly. First, we shuffle the dataset so that the observations are in a random order. We then let the algorithm run until line 6 in Algorithm 1 is called 100 times. After this we record the highest solution value which the algorithm has found thus far. We shuffle the dataset again and repeat. We do this 20 times. We then put the dataset in the order in which we found the highest solution value. We then run the above algorithm on this dataset. This helps ensure that the solution space is searched in a more efficient manner.

A9.7 Approximate Algorithm

We are able to use the above algorithm to calculate Varian's efficiency index for the LNU snd SMU models for all subjects in the data collected by Halevy, Persitz, and Zrill (2018). However, the algorithm simply took too long for some subjects when determining the index

Algorithm 1 Calculate Varian Index

```
function NEXTc(remainingobs)
    finishedobs = list of obs not in remainingobs
    while remainingobs not empty do
        t = lowest observation number from remainingobs
        if t ∈ V(c) then                                                    ▷ Shortcut 3
            c^t = c^t + 1
            if f(c) > s then                                                ▷ Shortcut 2
                nextc( remainingobs )
            end if
            c^t = c^t − 1
        end if
        Remove observation t from remainingobs and add it to finishedobs
        if (c^t)_{∪ t∈finishedobs} is not almost 𝒰-rationalizable then   ▷ Shortcut 1
            Terminate Function
        end if
    end while
    if f(c) > s then
        s = f(c)
    end if
end function
function VARIANINDEX
    s = inf f
    remainingobs = {1, . . . , T}
    nextc( remainingobs )
    return best solution found
end function
```

in the case of the EU model. The reason is two fold. First, checking for GARP or F-GARP involves checking the acyclicality of graphs with at most T^2 directed edges. On the other hand, the test for the EU model involves solving a linear programming problem with at most T^3 variables. Solving this linear programming problem takes longer to solve than checking the acyclicality conditions. Second, there are typically more revealed preference conditions in the EU test, leading to a set $\mathbf{C}$ with a larger cardinality and thus more checks.

However, even for the EU model, we are able to calculate bounds for each subject's Varian efficiency index. We now discuss how this was accomplished. First, we modify Algorithm 1 to make the order in which $\mathbf{C}$ is traversed random. To do this we make a shuffled copy of the remainingobs variable just before line 3 is run. Line 3 is then changed to a while statement which loops while this shuffled copy is not empty. This ensures that B is traversed in a random order. This means that if the algorithm were called many times some of these times would traverse more efficient paths while others would be less efficient. We shall refer to this new algorithm as Algorithm 1 Shuffle.

Often the space $\mathbf{C}$ is too large for the algorithm to traverse the solution space in a reasonable time frame. To deal with this we turn the original solution space $[0,1]^T$ in the definition of the Varian efficiency problem into a grid of points where each grid point is 2^{-k} away from its nearest neighbor for some non-negative integer k. We say that we are traversing the k-grid when we search the grid whose points are separated by a distance of 2^{-k}. We refer to this new algorithm as Algorithm 2; note that the three shortcuts used in Algorithm 1 are also applicable, when modified in the obvious way. Further, define Algorithm 2 Shuffle to be algorithm where we traverse the search space in a random order just as Algorithm 1 Shuffle.

To explain further how we proceed we must first introduce some terminology. Every time Algorithm 1 or 2 (or its shuffle versions) runs line 6 we say that the algorithm *steps down*. Every time the algorithm runs line 10 we say the algorithm *steps up*. Define an algorithm which is identical to Algorithm 2 Shuffle except that the algorithm terminates after it steps up for the first time. We refer to this as the Drill Algorithm. Calling the drill algorithm dives deep into the discretized space of efficiency vectors (randomly) until it either finds a single feasible solution or realizes there are no feasible solutions nearby which improve on the current best solution.

Algorithm 2 Calculate Approximate Varian Index

```
function NEXTe(remainingobs)
    finishedobs = list of obs not in remainingobs
    while remainingobs not empty do
        t = lowest observation number from remainingobs
        if t is involved in a violation at e then                  ▻ Shortcut 3
            e^t = e^t − 1/(1 + k)
            if f(e) > s then                                       ▻ Shortcut 2
                nexte( remainingobs )
            end if
            e^t = e^t + 1/(1 + k)
        end if
        Remove observation t from remainingobs and add it to finishedobs
        if (e^t)_{∪ t∈finishedobs} is not almost U-rationalizable then   ▻ Shortcut 1
            Terminate Function
        end if
    end while
    if f(e) > s then
        s = f(e)
    end if
end function
function VARIANINDEX2
    s = inf f
    remainingobs = {1, . . . , T}
    nexte( remainingobs )
    return best solution found
end function
```

We now describe the general approach we take to approximate the Varian index for the expected utility model. First, we run Algorithm 1 Shuffle until either it terminates or the algorithm steps down 100 times; we include this stage to get an exact solution for those cases where they are easy to calculate. If this algorithm has not terminated by then we set a variable k to 5 and run the Drill Algorithm 50 times traversing the k-grid. Each time the drill algorithm is run we wait until it terminates or we terminate it after it has taken 200 steps down. We repeat the process for k equal to 6, 7, 8, and 9. If for any k we do not improve our solution but the algorithm steps up at least once (that is, the algorithm has located a dead end in at least one of the 50 times) we skip over the remaining values of k that we have yet to try. This stage allows us to get a lower bound for the optimal efficiency score. This lower bound is useful in its own right and also improves the performance of the algorithm in the next stage.

In the next stage, we run Algorithm 2 Shuffle traversing the k-grid for k equal to 5, 6, 7, 8, and 9. If the Algorithm fails to terminate after 500 steps down then we terminate the algorithm and skip the remaining values of k. This step allows us to deduce an upper bound for the optimal solution. Suppose Algorithm 2 Shuffle successfully terminates and finds the optimal efficiency vector e^* in the 5-grid. Then, we know that the true solution to the Varian efficiency problem must be no greater than $f(\mathbf{e}^* + (2^{-6}, 2^{-6}, \ldots, 2^{-6}))$, so long as f is a symmetric function (which is true of Varian's index). Lastly, we compare the upper bound obtained in this way with the solution to the Varian efficiency problem for SMU functions (which can be calculated exactly for all subjects in the data of Halevy, Persitz, and Zrill (2018)); this number also constitutes an upper bound since the family of SMU functions contains the family of EU functions. The lower of these two numbers is then reported as the upper bound for the solution to the Varian efficiency problem for EU functions.

A9.8 Other Approaches to Estimating Varian's Efficiency Index for the LNU case

The Varian Efficiency score for the LNU model (but not for the SMU and EU models) is also calculated in Halevy, Persitz, and Zrill (2018). They employ three different methods to calculate this index. The first method is an exact algorithm (delivers the exact value of the index) but is potentially very computationally intense. The latter two methods provide approximations to the Varian efficiency index. Let us refer to these methods as Method 1, Method 2, and Method 3. Method 1 constructs a list of all violations of GARP. This means making a list $\{\ell_1, \ldots, \ell_M\}$ where each element of the list is an ordered sequence of observation numbers where each observation is revealed preferred to its successor in the list and the last element is revealed preferred to the first element. Method 1 then iterates over each element of $\ell_1 \times \ldots \ell_M$ calculating Varian's index corresponding to removing these revealed preference relations. The algorithm outputs the best solution found. The method considers all possible ways of resolving each violation of GARP and so will eventually calculate Varian's efficiency index. However, this method will not be practicable if the number of ways of resolving GARP violations is too large. In that case, Method 2 is considered. Method 2 creates a list of violations of the Weak Axiom of Revealed Preference $\{\ell_1, \ldots, \ell_M\}$ where each ℓ_m is a pair of observations each revealed preferred to the other. The method iterates over each element of $\ell_1 \times \ldots \ell_M$ recording the Varian index associated with removing the corresponding revealed preference relations. This method will calculate the Varian Index corresponding to removing all violations of WARP. As noted by Halevy, Persitz, and Zrill (2018) this will be an *upper* bound on the true Varian Index. If there are too many violations of WARP then Method 3 is used. Method 3 is to run Algorithm 3 in Alcantud *et al.* (2010). This method provides a heuristic for selecting one feasible solution to the Varian efficiency problem and is thus gives a *lower* bound to Varian's efficiency index.

References

AFRIAT, S. N. (1967): "The Construction of Utility Functions from Expenditure Data," *International Economic Review*, 8(1), 67–77.

ALCANTUD, J. C. R., D. L. MATOS, and C. R. PALMERO (2010): "Goodness-of-Fit in Optimizing a Consumer Model," *Mathematical and Computer Modelling*, 52(7–8), 1088–1094.

ANDREONI, J., and C. SPRENGER (2012): "Estimating Time Preferences from Convex Budgets," *American Economic Review*, 102(7), 3333–3356.

BRONARS, S. G. (1987): "The Power of Nonparametric Tests of Preference Maximization," *Econometrica*, 55(3), 693–698.

CHOI, S., R. FISMAN, D. GALE, and S. KARIV (2007): "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review*, 97(5), 1921–1938.

CHOI, S., S. KARIV, W. MÜLLER, and D. SILVERMAN (2014): "Who Is (More) Rational?," *American Economic Review*, 104(6), 1518–1550.

FISHBURN, P. C. (1975): "Separation Theorems and Expected Utilities," *Journal of Economic Theory*, 11(1), 16–34.

FORGES, F., and E. MINELLI (2009): "Afriat's Theorem for General Budget Sets," *Journal of Economic Theory*, 144(1), 135–145.

FRIEDMAN, M., and L. J. SAVAGE (1948): "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, 56(4), 279–304.

GILBOA, I., and D. SCHMEIDLER (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18(2), 141–153.

HALEVY, Y., D. PERSITZ, and L. ZRILL (2018): "Parametric Recoverability of Preferences," *Journal of Political Economy*, 126(4), 1558–1593.

KŐSZEGI, B., and M. RABIN (2007): "Reference-Dependent Risk Attitudes," *American Economic Review*, 97(4), 1047–1073.

MACCHERONI, F., M. MARINACCI, and A. RUSTICHINI (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74(6), 1447–1498.

MASATLIOGLU, Y., and C. RAYMOND (2016): "A Behavioral Analysis of Stochastic Reference Dependence," *American Economic Review*, 106(9), 2760–2782.

NISHIMURA, H., E. A. OK, and J. K.-H. QUAH (2017): "A Comprehensive Approach to Revealed Preference Theory," *American Economic Review*, 107(4), 1239–1263.

QUIGGIN, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3(4), 323–343.

SMEULDERS, B., F. C. R. SPIEKSMA, L. CHERCHYE, and B. DE ROCK (2014): "Goodness-of-Fit Measures for Revealed Preference Tests: Complexity Results and Algorithms," *ACM Transactions on Economics and Computation*, 2(1), 3:1–3:16.

TVERSKY, A., and D. KAHNEMAN (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5(4), 297–323.