University of
St Andrews

Revealed preferences over risk and uncertainty

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

Revealed preferences over risk and uncertainty

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

September 13, 2017

Abstract: We develop a nonparametric procedure, called the *lattice method*, for testing the consistency of contingent consumption data with a broad class of models of choice under risk and under uncertainty. Our method allows for risk loving and elation seeking behavior and can be used to calculate, via Afriat's efficiency index, the magnitude of violations from a particular model of choice. We evaluate the performance of different models (including expected utility, disappointment aversion, rank dependent utility, and stochastically monotone utility) in the data collected by Choi, Fisman, Gale, and Kariv (2007), in terms of pass rates, power, and predictive success.

Keywords: expected utility, rank dependent utility, disappointment aversion, generalized axiom of revealed preference, first order stochastic dominance, Afriat efficiency, Bronars power, Selten predictive success

JEL classification numbers: C14, C60, D11, D12, D81

Slides: http://www.mattpolisson.com/uploads/7/6/2/3/76230009/lattice.pdf

1

1. Introduction

This paper is a methodological contribution to the empirical investigation of decision making under risk and under uncertainty. At least since Allais (1953), there has been a large literature developing models of choice under risk or under uncertainty that seek to give a better account of observed behavior than the expected utility (EU) model. An empirical literature that tests the EU and other models on experimental data has also emerged alongside these theoretical developments. These experiments often employ elicitation procedures in which subjects are in effect making repeated choices between two risky or uncertain outcomes; the data obtained in this way consist of a finite number of *binary choices*, which can then be used to partially recover a subject's preference. A more recent strand of experiments employs a different elicitation procedure, which we shall call the *budgetary choice* procedure. In this case subjects are asked to choose a preferred option from a potentially infinite set of alternatives. For example, a subject could be presented with a portfolio problem where she has to allocate her budget between two assets with state-contingent payoffs. An early and influential experiment of this type, the data from which we analyze in this paper, is Choi, Fisman, Gale, and Kariv (2007).[1] Other examples include Loomes (1991), Gneezy and Potters (1997), Bayer *et al.* (2013), Choi *et al.* (2014), Ahn *et al.* (2014), Hey and Pace (2014), Cappelen *et al.* (2015), and Halevy, Persitz, and Zrill (2016).

The main contribution of this paper is to develop an *empirical test* that could be used to analyze data collected from portfolio decisions; it is applicable to experimental data where a budgetary choice elicitation procedure is employed and also to suitable non-experimental or field data. Our test allows us to determine whether a data set is consistent with the EU model or some of its generalizations, without making parametric assumptions on the Bernoulli index or other features of the model. It is worth noting that models of decision making over time (such as the discounted utility model), as well as models of decision making involving both time and risk, are formally very similar to the EU model and its generalizations. Budgetary choice procedures are increasingly used in experiments for studying these types of models,[2] and budgetary choices involving time preferences may occur naturally in the field as well; it

---

[1] See this paper also for an account of the advantages of such an approach.

[2] See Andreoni and Sprenger (2012) and, for a comprehensive list of papers employing such procedures to study time preferences, Imai and Camerer (2016).

is clear that the tests developed in this paper could be used with little or no modification in these contexts.

1.1 Testing EU and other models on a finite lattice

A feature of the budgetary choice procedure, and a reason why it is sometimes favored over binary choice procedures, is that instead of requiring the subject to pick one alternative or another, it allows her to calibrate her response and pick something 'in between.' But this feature is also the crucial reason why a new empirical tool is needed to analyze nonparametrically data collected from a budgetary procedure, whereas no such method is necessary for binary choices. Indeed, suppose we make a finite number of observations, where at observation t an agent chooses a lottery that gives a monetary payoff x_s^t in state s over one that gives y_s^t in state s (for $s = 1, 2, \ldots, \bar{s}$), where the probability of state s is known to be $\pi_s > 0$. Imagine that we would like to test if this data set is consistent with the EU model. Ignoring the issue of errors for the time being, checking for consistency with the EU model simply involves finding a strictly increasing Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}$ such that $\sum_{s=1}^{\bar{s}} \pi_s u(x_s^t) \geqslant \sum_{s=1}^{\bar{s}} \pi_s u(y_s^t)$ holds at every observation t. This amounts to solving a finite set of linear inequalities, and it is computationally straightforward to ascertain if a solution exists. However, it is clear that this method is no longer applicable when choices are instead made from a classical budget sets at every observation t, since even a single observed choice from a budget set reveals an infinite set of binary preferences.

The empirical test that we develop for solving this problem is very simple, and it is worth giving a short explanation here. Consider a data set with three observations and two states, as depicted in Figure 1. The subject chooses the contingent consumption bundle $(2, 5)$ from budget set B^1, $(6, 1)$ from B^2, and $(4, 3)$ from B^3. Assuming that the probability of state s is commonly known to be π_s, consistency with the EU model would require the existence of a strictly increasing Bernoulli function u such that $\pi_1 u(2) + \pi_2 u(5) \geqslant \pi_1 u(x) + \pi_2 u(y)$ for all (x, y) in B^1, and similarly at the other two observations.

In our main methodological result (Theorem 1), we show that this data set can be rationalized by the EU model if it can be rationalized on an appropriately modified consumption set. Specifically, let $\mathcal{X}$ be the set of consumption levels that are observed to have been chosen at some observation and in some state, plus zero; in this example $\mathcal{X} = \{0, 1, 2, 3, 4, 5, 6\}$.
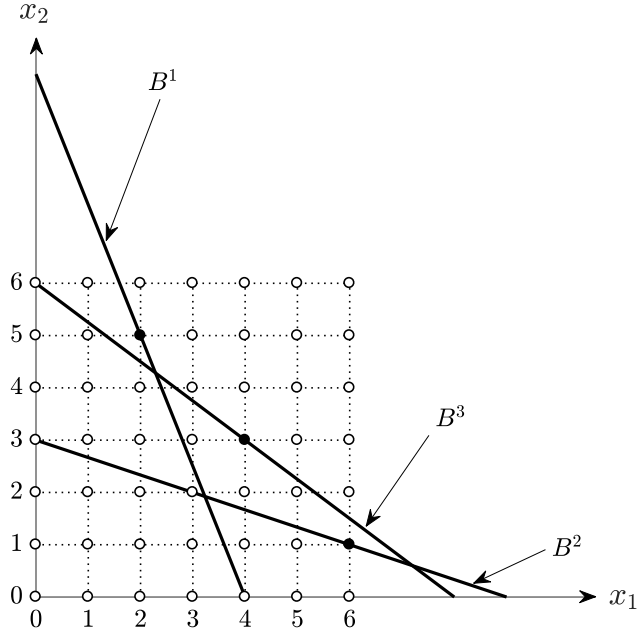
Figure 1: Constructing the finite lattice

Then for the data set to be EU-rationalizable, it is *sufficient* (and obviously necessary) for it to be EU-rationalizable on the reduced consumption set $\mathcal{X}^2$, i.e., there is an increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}$ such that the expected utility of $(2,5)$ is greater than any other bundle in $B^1 \cap \mathcal{X}^2$, and so forth. The set $\mathcal{X}^2$ is a finite lattice, depicted by the open circles in Figure 1. Therefore, checking for EU-rationalizability involves checking if there is a solution to a *finite* set of linear inequalities, a problem which is computationally feasible.[3]

This lattice method turns out to be very flexible: it can be used not just to check for EU-rationalizability, but also for consistency with other models of choice under risk (such as the rank dependent utility (RDU) model (Quiggin, 1982)) and under uncertainty (such as the maxmin expected utility model (Gilboa and Schmeidler, 1989)). The basic idea is always to convert an infinite collection of revealed preference pairs into a finite number involving only bundles on a finite lattice. Note also that the method *does not require linear budget sets*: it works for any type of constraint set, so long as it is compact.

1.2 Empirical implementation and findings

We implement our empirical method on a data set obtained from the well known portfolio

---

[3] This example is EU-rationalizable on $\mathcal{X}^2$ and thus EU-rationalizable. One solution is $\bar{u}(0) = 0$, $\bar{u}(1) = 1$, $\bar{u}(2) = 4$, $\bar{u}(3) = 6$, $\bar{u}(4) = 7$, $\bar{u}(5) = 8$, and $\bar{u}(6) = 9$.

choice experiment in Choi, Fisman, Gale, and Kariv (2007), in order to demonstrate that the lattice method works at a practical level, and also for its own sake. In this experiment, each subject was asked to purchase Arrow-Debreu securities under different budget constraints. There were two states of the world, and it was commonly known that states occurred either symmetrically (each with probability 1/2) or asymmetrically (one with probability 1/3 and the other with probability 2/3). In their analysis, Choi *et al.* (2007) first checked whether a subject's observations were consistent with the maximization of a locally nonsatiated utility function by performing the familiar GARP test (prescribed by Afriat's (1967) Theorem). Those subjects who passed or came sufficiently close to passing GARP were then fitted to a parametric version of the disappointment aversion (DA) model (Gul, 1991), which is a special case of the RDU model when there are two states of the world.

The lattice method developed in this paper makes it possible for us to evaluate other models of choice under risk (beyond basic utility maximization) using a completely nonparametric approach. We test whether a subject's choices are consistent with the EU, DA, and RDU models by applying the lattice method. We also test for consistency with the maximization of a utility function that is *stochastically monotone*, in the sense that if a bundle dominates another with respect to first order stochastic dominance, then it must have higher utility. The EU, DA, and RDU models are all special cases of the SMU model, which is in turn more stringent than basic utility maximization.

With 50 observations collected on each subject, it is unsurprising that hardly any subject would be exactly rationalizable by even the the most permissive model of utility maximization. It is possible to quantify a data set's departure from a particular notion of rationality using the *critical cost efficiency index* (Afriat (1972, 1973)); this index is widely used in the empirical revealed preference literature, including Choi *et al.* (2007).[4] This index runs from 1 to 0, with the index equal to 1 if the data set passes the test exactly. We adopt the same measure of rationality in this paper. Given the data obtained from a *particular subject*, and for each of the models that we consider, it is possible to calculate this index; in the case of the EU, DA, and RDU models, this calculation again relies upon the lattice method.

---

[4] For a recent paper that implements a variant of this index and discusses its merits, see Halevy, Persitz, and Zrill (2016).

In comparing the performance of different models, it is necessary to go beyond comparing (approximate) pass rates since a very permissive model will have a very high pass rate but also little restrictive and predictive power. A standard way of measuring power in the empirical revealed preference literature is to estimate the probability of a randomly generated data set failing the test for a given model (Bronars, 1987); a model has high power if this probability is high. When one is investigating nested models, it is also natural to examine *relative power*: for example, if we randomly select a data set which passes GARP, what is the probability that it is also consistent with the EU model? Even though the power of different models of choice under risk has been investigated in other contexts, we provide the first systematic investigation of this issue in the context of budgetary choice data.

So models can be compared in at least two dimensions: their pass rates and their power. One way of combining these into a single index is to evaluate the difference between the pass rate and the model's *precision*, which is the probability of a random data set passing the test for that model (in other words, 1 minus the power). Selten (1991) provides an axiomatization of this index and calls it the *index of predictive success*. We evaluate the performance of different models according to this index.

The following is a brief summary of our empirical findings:[5]

- At a cost efficiency threshold of 0.9, more than 80% of subjects pass GARP and are therefore consistent with the maximization of a locally nonsatiated utility function.

- Among this group of subjects more than half are rationalizable by the EU model.

- The SMU and RDU models explain a sizable proportion of subjects whose behavior is not captured by the EU model. This is not true of the DA model, even though it is in principle a more permissive model than EU.

- If we randomly generate a data set which passes GARP, the probability of it being consistent with the RDU model, and hence the more stringent DA and EU models, is effectively zero. In other words, the power of these models is close to perfect, even among subjects who are consistent with utility maximization.

- After conditioning on passing GARP, the SMU and RDU models have the highest

---

[5] These findings are broadly consistent with results from earlier studies (see Section 4 for details).

indices of predictive success. These models perform well because they capture significantly more of the population than the EU model, without sacrificing power.

1.3 Relationship with the revealed preference literature

Our paper is related to the revealed preference literature originating from Afriat's (1967) Theorem, which characterizes consumer demand observations that are consistent with the maximization of a locally nonsatiated utility function (see also Diewert (1973) and Varian (1982)). Afriat's Theorem has *theoretical* significance in the sense that its intuitive behavioral characterization of basic rationalizability (through GARP) provides a justification for utility maximization in the consumer demand context, but it also provides a viable *empirical method* for testing rationalizability, which is why it has given rise to a large empirical literature.

A natural follow up to Afriat's contribution is to characterize those data sets which are rationalizable by more specialized utility functions. Among these papers are those which characterize observations of contingent consumption demand that are consistent with the EU model and (in more recent papers) some of its generalizations; these include Varian (1983a, 1983b, 1988), Green and Srivastava (1986), Diewert (2012), Bayer *et al.* (2013), Echenique and Saito (2015), Chambers, Liu, and Martinez (2016), and Chambers, Echenique, and Saito (2016).[6] The principal difference between our results and this literature is that we do not rely on the methods of convex optimization; this means, in particular, that we do not require (or guarantee) the concavity of the Bernoulli function, and our results are applicable to data sets with general constraint sets rather than just linear budget sets. For reasons which we make clear later in the paper, the fact that we allow for nonlinear constraint sets means that our method can also be used to calculate Afriat's efficiency index.

It is worth mentioning that not all revealed preference results (involving budgetary observations) that flow from Afriat's Theorem have the feature of providing both theoretical insight and an empirical method. There are papers where the emphasis is on providing a characterization that offers theoretical insight; in other cases the emphasis is on providing an empirically viable method of model testing.[7] Our main methodological result (Theorem

---

[6] There is also a closely related literature on recovering expected utility from asset or contingent consumption demand *functions*, where, in effect the data set is assumed to be infinite (see, for example, Dybvig and Polemarchakis (1981) and Kubler, Selden, and Wei (2014)).

[7] For example, Chambers, Liu, and Martinez (2016) provides a characterization of the former type, and

1) says that, for a broad class of models, to check for rationalizability it suffices to check for rationalizabilty *as if* the subject's consumption space is some finite lattice (constructed from the data). By itself, the result does not furnish us with any theoretical motivation for one model or another; its principal value is in providing us with an empirical tool.

1.4 Organization of the paper

Section 2 provides a description of the lattice method and explains how it can be used to test the EU, DA, and RDU models in a budgetary choice environment. Further applications of the lattice method, including to models of decision making under uncertainty, can be found in the Online Appendix. In Section 3 we explain Afriat's efficiency index and how the lattice method can be used to calculate this index. The empirical application to the portfolio choice data collected by Choi, Fisman, Gale, and Kariv (2007) can be found in Section 4.

2. THE LATTICE METHOD

We assume that there is a finite set of states, denoted by $S = \{1, 2, \ldots, \bar{s}\}$. The contingent consumption space is $\mathbb{R}_+^{\bar{s}}$; for a typical consumption bundle $\mathbf{x} \in \mathbb{R}_+^{\bar{s}}$, the sth entry, x_s, specifies the consumption level in state s.[8] There are T observations in the data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$; by this we mean that the agent is observed choosing the bundle $\mathbf{x}^t$ from $B^t \subset \mathbb{R}_+^{\bar{s}}$. We assume that B^t is compact and that $\mathbf{x}^t \in \partial B^t$, where ∂B^t denotes the *upper boundary* of B^t.[9] The most important example of B^t is the classical linear budget set under complete markets, i.e.,

$$B^t = \{\mathbf{x} \in \mathbb{R}_+^{\bar{s}} : \mathbf{p}^t \cdot \mathbf{x} \leqslant \mathbf{p}^t \cdot \mathbf{x}^t\}, \tag{1}$$

with $\mathbf{p}^t \gg 0$ denoting the vector of state prices. In this case, we may also write the data set as $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^T$. The experiment conducted by Choi *et al.* (2007), the data from which we analyze in Section 4, involves subjects choosing from linear budget sets.

---

they point out this distinction clearly in the introduction to their paper.

[8] Our results do depend on the realization in each state being one-dimensional (which can be interpreted as a monetary payoff, but not a bundle of goods). This case is the one most often considered in applications and experiments and is also the assumption in a number of recent papers, including Kubler, Selden, and Wei (2014), Echenique and Saito (2015), and Chambers, Echenique, and Saito (2016). The papers by Varian (1983a, 1983b), Green and Srivastava (1986), Bayer *et al.* (2013), and Chambers, Liu, and Martinez (2016) allow for multi-dimensional realizations but (like the three aforementioned papers) they also require the convexity of the agent's preference over contingent consumption and linear budget sets.

[9] An element $\mathbf{y} \in B^t$ is in ∂B^t if there is no $\mathbf{x} \in B^t$ such that $\mathbf{x} > \mathbf{y}$. (For the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{\bar{s}}$, we write $\mathbf{x} \geqslant \mathbf{y}$ if $x_s \geqslant y_s$ for all s, and $\mathbf{x} > \mathbf{y}$ if $\mathbf{x} \geqslant \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$. If $x_s > y_s$ for all s, we write $\mathbf{x} \gg \mathbf{y}$.) For example, if $B^t = \{(x, y) \in \mathbb{R}_+^2 : (x, y) \leqslant (1, 1)\}$, then $(1, 1) \in \partial B^t$ but $(1, 1/2) \notin \partial B^t$.

Bear in mind, however, that our formulation only requires B^t to be compact and, in particular, it does not have to be a linear budget set. A crucial application requiring B^t to be nonlinear is found in Section 3, where we define approximate rationalizability. Another natural example of a nonlinear budget set is when a subject chooses contingent consumption through a portfolio of securities in an incomplete market; in this case, the budget set will be compact so long as the security prices do not admit arbitrage.[10]

Let $\{\phi(\cdot, t)\}_{t=1}^T$ be a collection of functions, where $\phi(\cdot, t) : \mathbb{R}_+^{\bar{s}} \to \mathbb{R}$ is continuous and strictly increasing.[11] The data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is said to be *rationalizable by* $\{\phi(\cdot, t)\}_{t=1}^T$ if there exists a continuous and strictly increasing function $u : \mathbb{R}_+ \to \mathbb{R}_+$, which we shall refer to as the *Bernoulli function*, such that

$$\phi(\mathbf{u}(\mathbf{x}^t), t) \geqslant \phi(\mathbf{u}(\mathbf{x}), t) \ \text{ for all } \mathbf{x} \in B^t, \tag{2}$$

where $\mathbf{u}(\mathbf{x}) = (u(x_1), u(x_2), \ldots, u(x_{\bar{s}}))$. In other words, $\mathbf{x}^t$ maximizes $\phi(\mathbf{u}(\mathbf{x}), t)$ in B^t. It is natural to require u to be strictly increasing since we typically interpret its argument to be money. The requirements on u guarantee that $\phi(\mathbf{u}(\cdot), t)$ is continuous and strictly increasing in $\mathbf{x}$. Note that continuity is an important property because it guarantees that the agent's utility maximization problem always has a solution on a compact constraint set.[12]

Expected utility. This model clearly falls within the framework we have set up. Indeed, suppose that both the observer and the agent know that the probability of state s at observation t is $\pi_s^t > 0$. If the agent is maximizing expected utility (EU),

$$\phi(u_1, u_2, \ldots, u_{\bar{s}}, t) = \sum_{s=1}^{\bar{s}} \pi_s^t u_s, \tag{3}$$

and (2) requires that

$$\sum_{s=1}^{\bar{s}} \pi_s^t u(x_s^t) \geqslant \sum_{s=1}^{\bar{s}} \pi_s^t u(x_s) \ \text{ for all } \mathbf{x} \in B^t, \tag{4}$$

---

[10] Indeed, there is $\mathbf{p}^t \gg 0$ such that $B^t = \{\mathbf{x} \in \mathbb{R}_+^{\bar{s}} : \mathbf{p}^t \cdot \mathbf{x} \leqslant \mathbf{p}^t \cdot \mathbf{x}^t\} \cap \{Z + \boldsymbol{\omega}\}$, where Z is the span of assets available to the agent and $\boldsymbol{\omega}$ is the agent's endowment of contingent consumption. Both B^t and $\mathbf{x}^t$ will be known to the observer, if he knows the asset prices, the agent's holding of securities, the asset payoffs in every state, and the agent's endowment of contingent consumption $\boldsymbol{\omega}$.

[11] By strictly increasing, we mean that $\phi(\mathbf{z}, t) > \phi(\mathbf{z}', t)$ if $\mathbf{z} > \mathbf{z}'$.

[12] The existence of a solution is obviously important if we are to make out-of-sample predictions. More fundamentally, a hypothesis that an agent is choosing a utility-maximizing bundle implicitly assumes that the utility function is such that an optimum exists for a reasonably broad class of constraint sets.

i.e., the expected utility of $\mathbf{x}^t$ is greater than that of any other bundle in B^t. When there exists a Bernoulli function u such that (4) holds, we say that the data set is *EU-rationalizable with the probability weights* $\{\boldsymbol{\pi}^t\}_{t=1}^T$, where $\boldsymbol{\pi}^t = (\pi_1^t, \pi_2^t, \ldots, \pi_{\bar{s}}^t)$.

If $\mathcal{O}$ is rationalizable by $\{\phi(\cdot, t)\}_{t=1}^T$, then since the objective function $\phi(\mathbf{u}(\cdot), t)$ is strictly increasing in $\mathbf{x}$, the rationalizability condition (2) could be strengthened to

$$\phi(\mathbf{u}(\mathbf{x}^t), t) \geqslant \phi(\mathbf{u}(\mathbf{x}), t) \ \text{ for all } \mathbf{x} \in \underline{B}^t, \tag{5}$$

where $\underline{B}^t$ is the *downward extension of* B^t, i.e.,

$$\underline{B}^t = \{\mathbf{y} \in \mathbb{R}_+^{\bar{s}} : \mathbf{y} \leqslant \mathbf{x} \text{ for some } \mathbf{x} \in B^t\}.$$

Furthermore, the inequality in (5) is strict whenever $\mathbf{x} \in \underline{B}^t \backslash \partial \underline{B}^t$ (where $\partial \underline{B}^t$ refers to the upper boundary of $\underline{B}^t$). We define $\mathcal{X} = \{x' \in \mathbb{R}_+ : x' = x_s^t \text{ for some } t, s\} \cup \{0\}$; besides zero, $\mathcal{X}$ contains those levels of consumption that are chosen at some observation and in some state. Since the data set is finite, so is $\mathcal{X}$. Given $\mathcal{X}$, we may construct $\mathcal{L} = \mathcal{X}^{\bar{s}}$, which consists of a finite grid of points in $\mathbb{R}_+^{\bar{s}}$; in formal terms, $\mathcal{L}$ is a finite lattice. Let $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ be the restriction of the Bernoulli function u to $\mathcal{X}$. Given $\mathcal{O}$, the following must hold:

$$\phi(\bar{\mathbf{u}}(\mathbf{x}^t), t) \geqslant \phi(\bar{\mathbf{u}}(\mathbf{x}), t) \ \text{ for all } \mathbf{x} \in \underline{B}^t \cap \mathcal{L} \text{ and} \tag{6}$$

$$\phi(\bar{\mathbf{u}}(\mathbf{x}^t), t) > \phi(\bar{\mathbf{u}}(\mathbf{x}), t) \ \text{ for all } \mathbf{x} \in \left(\underline{B}^t \backslash \partial \underline{B}^t\right) \cap \mathcal{L}, \tag{7}$$

where $\bar{\mathbf{u}}(\mathbf{x}) = (\bar{u}(x_1), \bar{u}(x_2), \ldots, \bar{u}(x_{\bar{s}}))$. Our main theorem says the converse is also true.[13]

THEOREM 1. *Suppose that for some data set* $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ *and collection of continuous and strictly increasing functions* $\{\phi(\cdot, t)\}_{t=1}^T$, *there is a strictly increasing function* $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ *that satisfies conditions (6) and (7). Then there is a Bernoulli function* $u : \mathbb{R}_+ \to \mathbb{R}_+$ *that extends* $\bar{u}$ *and guarantees the rationalizability of* $\mathcal{O}$ *by* $\{\phi(\cdot, t)\}_{t=1}^T$.[14]

---

[13] Note that $\underline{B}^t$ cannot be replaced with B^t in (6) and (7). For example, suppose there are two observations, where $\mathbf{x}^1 = (1, 0)$ is chosen from $B^1 = \{(x_1, x_2) \in \mathbb{R}_+^2 : 2x_1 + x_2 = 2\}$ and $\mathbf{x}^2 = (0, 1)$ is chosen from $B^2 = \{(x_1, x_2) \in \mathbb{R}_+^2 : x_1 + 2x_2 = 2\}$. This pair of observations cannot be rationalized by any increasing utility function (even though the 'budget sets' are just lines) and, in particular, cannot be rationalized in the sense of Theorem 1 (with ϕ constant across t). However, since $\mathcal{L} = \{(0,0), (0,1), (1,0), (1,1)\}$, $B^1 \cap \mathcal{L} = \{(1,0)\}$ and $B^2 \cap \mathcal{L} = \{(0,1)\}$, so conditions (6) and (7) are vacuous. On the other hand $\left(\underline{B}^1 \backslash \partial \underline{B}^1\right) \cap \mathcal{L}$ contains $(0,1)$ and $\left(\underline{B}^2 \backslash \partial \underline{B}^2\right) \cap \mathcal{L}$ contains $(1,0)$, so (7) requires $\phi(\bar{\mathbf{u}}(\mathbf{x}^1)) > \phi(\bar{\mathbf{u}}(\mathbf{x}^2))$ and $\phi(\bar{\mathbf{u}}(\mathbf{x}^1)) < \phi(\bar{\mathbf{u}}(\mathbf{x}^2))$, which plainly cannot happen. This allows us to conclude, correctly, that the data set is not rationalizable.

[14] The increasing assumptions on ϕ and $\bar{u}$ ensure that we may confine ourselves to checking (6) and (7) for undominated elements of $\underline{B}^t \cap \mathcal{L}$, i.e., $\mathbf{x} \in \underline{B}^t \cap \mathcal{L}$ such that there does not exist $\mathbf{x}' \in \underline{B}^t \cap \mathcal{L}$ with $\mathbf{x} < \mathbf{x}'$.

What Theorem 1 achieves is *domain reduction*: checking the rationalizability of $\mathcal{O}$ is equivalent to checking rationalizability in the case where the agent's consumption space is considered to be $\mathcal{L}$ rather than $\mathbb{R}_+^{\bar{s}}$, which (crucially) reduces the rationality requirements to a *finite* number of inequalities, each involving the observed choice and an alternative (see (6) and (7)), and with the Bernoulli function defined on $\mathcal{X}$ rather than $\mathbb{R}_+$.

The intuition for Theorem 1 ought to be strong. Given $\bar{u}$ satisfying (6) and (7), we can define the step function $\hat{u} : \mathbb{R}_+ \to \mathbb{R}_+$ where $\hat{u}(r) = \bar{u}([r])$, with $[r]$ being the largest element of $\mathcal{X}$ weakly lower than r, i.e., $[r] = \max\{r' \in \mathcal{X} : r' \leqslant r\}$. Notice that $\phi(\hat{\mathbf{u}}(\mathbf{x}^t), t) = \phi(\bar{\mathbf{u}}(\mathbf{x}^t), t)$ and, for any $\mathbf{x} \in \underline{B}^t$, $\phi(\hat{\mathbf{u}}(\mathbf{x}), t) = \phi(\bar{\mathbf{u}}([\mathbf{x}]), t)$, where $[\mathbf{x}] = ([x_1], [x_2], \ldots, [x_{\bar{s}}])$ in $\underline{B}^t \cap \mathcal{L}$. Clearly, if $\bar{u}$ obeys (6) and (7) then $\mathcal{O}$ is rationalized by $\{\phi(\cdot, t)\}_{t=1}^T$ and $\hat{u}$ (in the sense that (2) holds). This falls short of the claim in the theorem only because $\hat{u}$ is neither continuous nor strictly increasing; the proof in the Appendix shows how one could in fact construct a Bernoulli function with these additional properties.

2.1 Testing the expected utility model

Theorem 1 provides us with a very convenient way of testing EU-rationalizability. The theorem tell us that $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is EU-rationalizable with the probability weights $\{\boldsymbol{\pi}^t\}_{t=1}^T$ if and only if there is a collection of real numbers $\{\bar{u}(r)\}_{r \in \mathcal{X}}$ such that

$$0 \leqslant \bar{u}(r') < \bar{u}(r) \text{ whenever } r' < r, \tag{8}$$

and the inequalities (6) and (7) hold, where $\phi(\cdot, t)$ is defined by (3). This is a linear program and it is both formally *solvable* (in the sense that there is an algorithm that can decide within a known number of steps whether or not there is a solution to this set of linear inequalities) and also computationally feasible.

At this point it is worth emphasizing that requiring a data set to be EU-rationalizable is certainly more stringent than simply requiring it to be rationalizable by a locally nonsatiated utility function. Indeed, while a data set with a single observation $(\mathbf{x}^t, \mathbf{p}^t)$ must be consistent with the maximization of a strictly increasing (and hence locally nonsatiated) utility function, even a single observation can be incompatible with the EU model.

EXAMPLE 1. Suppose that there are two equiprobable states of the world, and at the price vector $\mathbf{p}^t = (p_1^t, p_2^t)$ such that $p_1^t > p_2^t$, the agent purchases a bundle $\mathbf{x}^t$ such that

$x_1^t > x_2^t$. We claim that this is not EU-rationalizable, or in other words, the agent cannot buy strictly more of the more expensive good. Indeed, such an observation is not compatible with the maximization of any symmetric and strictly increasing utility function on $\mathbb{R}_+^2$: with symmetry, the bundle (y_1, y_2), where $y_1 = x_2^t$ and $y_2 = x_1^t$, is strictly cheaper than $\mathbf{x}^t$ but gives the same utility, so $\mathbf{x}^t$ is not optimal. Such an observation will also fail the lattice test, since (y_1, y_2) is in $\mathcal{L} \cap (\underline{B}^t \backslash \partial \underline{B}^t)$ but the condition (7) is not satisfied. $\qquad \square$

On the other hand, our test is strictly less stringent than a test of EU-rationalizability that also requires the Bernoulli function to be concave (such as Green and Srivastava (1986)); imposing concavity on the Bernoulli function has observable implications over and above those which flow simply from the EU model, as the following example demonstrates.

EXAMPLE 2. Suppose an agent maximizes expected utility and has the Bernoulli function $u(y) = (y - 4)^3$, which is strictly concave for $y < 4$ and strictly convex otherwise.[15] There are two states of the world, which occur with equal probability. At $\mathbf{p}^t = (1, 3/2)$ and with wealth 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 $\mathbf{x}^t$ (see Figure 2).[16] 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 consisting of two observations: the bundle $\mathbf{x}^t$ chosen at $\mathbf{p}^t = (1, 3/2)$ and $\mathbf{x}^{t'} = (64, 0)$ chosen at $\mathbf{p}^{t'} = (1, 1)$. This data set is EU-rationalizable and it will pass the lattice test, but it cannot be rationalized by a concave Bernoulli function. Indeed,

$$u(x_1^t) + u(x_2^t) \geqslant u(1) + u(0) \tag{9}$$

since $(1, 0)$ is affordable to the agent when $\mathbf{x}^t$ is chosen. If we further assume that u is concave, $u(1) - u(x_1^t) \geqslant u(64) - u(x_1^t + 63)$; substituting this into (9), we obtain $u(x_1^t + 63) + u(x_2^t) \geqslant u(64) + u(0)$. The bundle $((x_1^t + 63), x_2^t)$ is strictly cheaper than $\mathbf{x}^{t'} = (64, 0)$ at $\mathbf{p}^{t'}$ (see Figure 2), so $\mathbf{x}^{t'}$ cannot be optimal. To conclude, while $\mathcal{O} = \{(\mathbf{p}^t, \mathbf{x}^t), (\mathbf{p}^{t'}, \mathbf{x}^{t'})\}$ is indeed

---

[15] A Bernoulli function with a concave region followed by a convex region is used by Friedman and Savage (1948, Figure 2) to explain why an agent can simultaneously buy insurance and accept risky gambles.

[16] Solving the (quadratic) first order condition gives $\mathbf{x}^t \approx (0.83, 0.11)$.
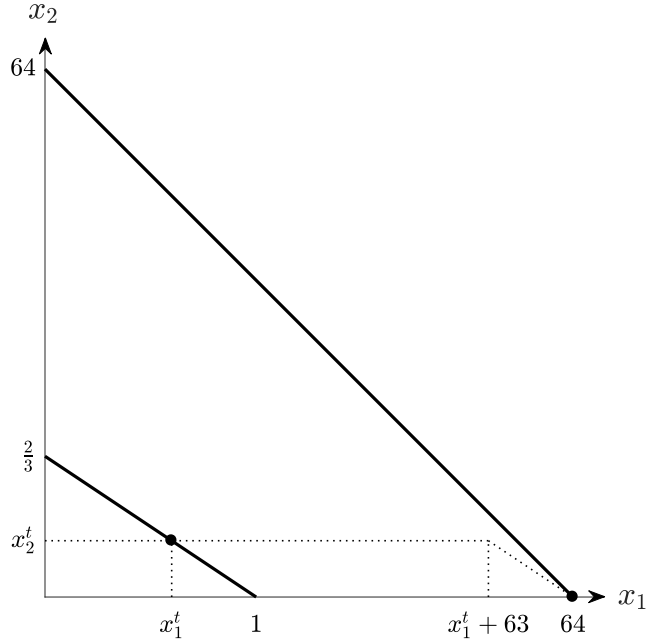
Figure 2: EU-rationalizable but not concave EU-rationalizable data set

EU-rationalizable, it is not EU-rationalizable with a concave Bernoulli function. □

We have shown that a data set is EU-rationalizable if and only if it is EU-rationalizable on $\mathcal{L}$ and the latter is in turn equivalent to the existence of a function $\bar{u}$ obeying conditions (6), (7), and (8) (with $\phi(\mathbf{u}, t) = \sum_{s=1}^{\bar{s}} \pi_s^t u_s$). Conditions (6) and (7) generate a finite list of preference pairs between some chosen bundle $\mathbf{x}^t$ and another bundle $\mathbf{x}$ in $\underline{B}^t \cap \mathcal{L}$ or $(\underline{B}^t \backslash \partial \underline{B}^t) \cap \mathcal{L}$. Condition (8) 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 $N < M$), the bundle $\mathbf{a}^j$ is weakly preferred to $\mathbf{b}^j$ (so the pairs are drawn from (6)) and for $j > N$, $\mathbf{a}^j$ is strictly preferred to $\mathbf{b}^j$ (so the pairs are drawn from (7) and (8)). Each bundle $\mathbf{a}^j$ can be written in its lottery form $\hat{\mathbf{a}}^j$, where $\hat{\mathbf{a}}^j$ is the 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 $\hat{\mathbf{b}}^j$. For example, in the example given in the introduction, $\mathcal{X} = \{0, 1, 2, 3, 4, 5, 6\}$ and the two states are equiprobable, so the bundle $(2, 5)$ chosen from B^1 has the lottery form $(0, 0, 1/2, 0, 0, 1/2, 0)$.

We know from Fishburn (1975) that the list $\{(\mathbf{a}^j, \mathbf{b}^j)\}_{j=1}^M$ is rationalizable by EU (i.e., there is $\bar{u}$ that solves (6), (7), and (8) with $\phi(\mathbf{u}, t) = \sum_{s=1}^{\bar{s}} \pi_s^t u_s$) if and only if 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

$$\sum_{j=1}^{M} \lambda^j \hat{\mathbf{a}}^j = \sum_{j=1}^{M} \lambda^j \hat{\mathbf{b}}^j. \tag{10}$$

This condition is very intuitive: assuming that the agent has a preference over lotteries, the independence axiom says that the lottery $\sum_{j=1}^{M} \lambda^j \hat{\mathbf{a}}^j$ must be strictly preferred to $\sum_{j=1}^{M} \lambda^j \hat{\mathbf{b}}^j$, and therefore (10) is excluded.[17] Put another way, a violation of Fishburn's condition must imply a violation of the independence axiom.

To summarize, 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 it is EU-rationalizable with probability weights* $\{\boldsymbol{\pi}^t\}_{t=1}^{T}$ *on the domain* $\mathcal{L}$ *and this in turn holds if and only if the preference pairs on* $\mathcal{L}$ *(as revealed by the data) do not contain a contradiction of the independence axiom of the form (10).* Example 3, to be explained later in the paper, gives an example of a data set which violates Fishburn's condition on $\mathcal{L}$ and is therefore not EU-rationalizable.

2.2 Other applications of the lattice method

So far, we have considered tests of EU-rationalizability in the case where the probability of each state is known to both the agent and the observer. The testing procedure extends to the case where no objective probabilities can be attached to each state. A data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^{T}$ is *rationalizable by subjective expected utility* (SEU) if there exist probability weights $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_{\bar{s}}) \gg 0$, with $\sum_{s=1}^{\bar{s}} \pi_s = 1$, and a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that, for all $t = 1, 2, \ldots, T$,

$$\sum_{s=1}^{\bar{s}} \pi_s u(x_s^t) \geqslant \sum_{s=1}^{\bar{s}} \pi_s u(x_s) \text{ for all } \mathbf{x} \in B^t.$$

In this case, ϕ is independent of t and instead of being fixed, it is required to belong to the family of functions Φ_{SEU} such that $\phi \in \Phi_{SEU}$ if $\phi(\mathbf{u}) = \sum_{s=1}^{\bar{s}} \pi_s u_s$ for some $\boldsymbol{\pi} \gg 0$. By Theorem 1, the data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^{T}$ can be rationalized by some $\phi \in \Phi_{SEU}$ if *there is a strictly increasing* $\bar{u}$ *such that (6) and (7) holds* and it is clear that these conditions are also

---

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

necessary. These conditions form a system of bilinear inequalities with unknowns $\{\pi_s\}_{s=1}^{\bar{s}}$ and $\{\bar{u}(r)\}_{r \in \mathcal{X}}$.

For many of the standard models of decision making under risk or uncertainty, the rationalizability problem has a structure similar to that of SEU in the sense that rationalizability by a particular model involves finding a Bernoulli function u and a function ϕ belonging to some family Φ that together rationalize the data, and this problem can in turn be transformed (via Theorem 1) into a problem of solving a system of bilinear inequalities. In the Online Appendix, we show how tests for various models, including choice acclimating personal equilibrium (Köszegi and Rabin, 2007), maxmin expected utility (Gilboa and Schmeidler, 1989), and variational preferences (Maccheroni, Marinacci, and Rustichini, 2006), can be devised using Theorem 1.

Even though solving a bilinear problem may be computationally intensive, the Tarski-Seidenberg Theorem tells us that this problem is decidable, in the sense that there is a known algorithm that can determine in a finite number of steps whether or not a solution exists. Nonlinear tests are not new to the revealed preference literature; for example, they appear in tests of weak separability (Varian, 1983a), in tests of maxmin expected utility and other models of ambiguity (Bayer *et al.*, 2013), and in tests of Walrasian general equilibrium (Brown and Matzkin, 1996). Solving such problems can be computationally demanding, but some cases can be computationally straightforward because of certain special features and/or when the number of observations is small. In the case of the tests that we develop, they simplify dramatically and are implementable in practice when there are only two states (though they remain nonlinear). The two-state case, while special, is very common in applied theoretical settings and laboratory experiments. For example, to implement the SEU test, simply condition on the probability of state 1 (and hence on the probability of state 2), and then perform a linear test to check whether there is a collection of real numbers $\{\bar{u}(r)\}_{r \in \mathcal{X}}$ solving (6), (7), and (8) (with $\phi(\cdot, t)$ is defined by (3)). If not, choose another probability for state 1, implement, and repeat (if necessary). Even a uniform grid search of up to two decimal places on the probability of state 1 will lead to no more than 99 linear tests, which can be implemented with little difficulty.[18]

---

[18] While we have not found it necessary to use them in our implementation in this paper, there are solvers available for mixed integer nonlinear programs (for example, as surveyed in Bussieck and Vigerske (2010))

Rank dependence and disappointment aversion. The rank dependent utility (RDU) model (Quiggin, 1982) is a prominent model of choice under risk. In Section 4, we report the findings of a test of this model, so we explain it here in greater detail. Let $\pi_s > 0$ be the objective probability of state s.[19] Given a contingent consumption bundle $\mathbf{x}$, we can rank the entries of $\mathbf{x}$ from the smallest to the largest, with ties broken by the rank of the state. We denote by $r(\mathbf{x}, s)$, the rank of x_s in $\mathbf{x}$. 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 $V(\mathbf{x}) = \sum_{s=1}^{\bar{s}} \delta(\mathbf{x}, s) u(x_s)$ where $u : \mathbb{R}_+ \to \mathbb{R}$ is a Bernoulli function,

$$\delta(\mathbf{x}, s) = g\left(\sum_{\{s' : r(\mathbf{x}, s') \leqslant r(\mathbf{x}, s)\}} \pi_{s'}\right) - g\left(\sum_{\{s' : r(\mathbf{x}, s') < r(\mathbf{x}, s)\}} \pi_{s'}\right),$$

and $g : [0, 1] \to \mathbb{R}$ is a continuous and strictly increasing function. (If $\{s' : r(\mathbf{x}, s') < r(\mathbf{x}, s)\}$ is empty, we let $g\left(\sum_{\{s' : r(\mathbf{x}, s') < r(\mathbf{x}, s)\}} \pi_{s'}\right) = g(0)$.) The function g distorts the cumulative distribution of the bundle $\mathbf{x}$, so that an agent maximizing rank dependent utility can behave as though the probability he attaches to a state depends on the relative attractiveness of the outcome in that state. Since u is strictly increasing, $\delta(\mathbf{x}, s) = \delta(\mathbf{u}(\mathbf{x}), s)$ and therefore $V(\mathbf{x}) = \phi(\mathbf{u}(\mathbf{x}))$, where for any vector $\mathbf{u} = (u_1, u_2, \ldots, u_{\bar{s}})$,

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

Note that the function ϕ is continuous and strictly increasing in $\mathbf{u}$. So V has the form assumed in Theorem 1, and we can use that result to devise a test for RDU-rationalizability.

We discuss a multiple-state version of the RDU test in the Appendix; at this point it suffices to explain the two-state case, which is the one relevant to the implementation in Section 4. Let $\rho_s = g(\pi_s)$ be the distorted value of π_s (the true probability of state s, for $s = 1, 2$). Then $\phi(u_1, u_2) = \rho_1 u_1 + (1 - \rho_1) u_2$ if $u_1 \leqslant u_2$ and $\phi(u_1, u_2) = (1 - \rho_2) u_1 + \rho_2 u_2$ if $u_1 > u_2$; by Theorem 1, a sufficient (and obviously necessary) condition for a data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ to be RDU-rationalizable is for there to be a solution to (6) and (7), with this formula for ϕ. This test involves solving a set of inequalities that are bilinear in

---

that are potentially useful for implementing bilinear tests more generally.

[19] 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.

the unknowns $\{\bar{u}(r)\}_{r \in \mathcal{X}}$ and $\{\rho_1, \rho_2\}$. In our implementation, we simply let ρ_1 and ρ_2 take different values on a very fine grid in $[0,1]^2$, subject to $\rho_1 \leqslant \rho_2$ (if and only if $\pi_1 \leqslant \pi_2$) and (for each case) perform the corresponding linear test.

It is worth emphasizing at this point that for certain values of ρ_1 and ρ_2, the function ϕ is clearly not concave or even quasiconcave, and therefore we cannot guarantee the quasiconcavity of the agent's utility over contingent consumption, *even if* we restrict ourselves to concave Bernoulli functions. While the lattice method still works in these cases, it is not possible to formulate a test for rationalizability that allows for non-quasiconcave utility functions using concave optimization methods (such as those cited in Section 1.3) because the first order conditions are no longer *sufficient* for optimality.

We also implement a lattice test of Gul's (1991) model of disappointment aversion (DA). When there are two states, the DA model is a special case of the RDU model with a further restriction on ρ_1 and ρ_2. Specifically, there is some $\beta \in (-1, \infty)$ such that, for $s = 1, 2$,

$$\rho_s = \frac{(1 + \beta)\pi_s}{1 + \pi_s \beta}. \tag{11}$$

Note that this restriction has bite only if $\rho_1 \neq \rho_2$, so in fact the RDU and DA models are identical when $\pi_1 = \pi_2$. If $\beta = 0$, the agent simply maximizes expected utility. If $\beta > 0$, we have $\rho_s > \pi_s$, so the agent attaches a probability on s that is higher than the objective probability when s is the less favorable state; in this case, the agent is said to be *disappointment averse*. If $\beta < 0$, then $\rho_s < \pi_s$, and the agent is said to be *elation seeking*; this is an instance where the function ϕ is not quasiconcave. As in the RDU model, we test the DA model by letting β take on different values and performing the associated linear test.

While it is well known that the RDU and EU models lead to different predictions, it not immediately clear that they are observationally distinct in the context of observations drawn from linear budgets. We end this section with an example of a data set that is RDU-rationalizable but not EU-rationalizable.

EXAMPLE 3. Suppose the data set consists of three observations $(\mathbf{x}^t, \mathbf{p}^t)$, for $t = 1, 2, 3$, where $\mathbf{p}^1 = (1, q)$, $\mathbf{x}^1 = (a, a)$; $\mathbf{p}^2 = (1, 1/q)$, $\mathbf{x}^2 = (b, b)$; and $\mathbf{p}^3 = (1, (1/q^2) + \epsilon)$, $\mathbf{x}^3 = (a + (a - b)/q, b + (b - a)q)$, with $q > 1$ and $a < b$, and $\epsilon > 0$ is a small number. The three observations are depicted Figure 3, where $c = a + (a - b)/q$ and $d = b + (b - a)q$.
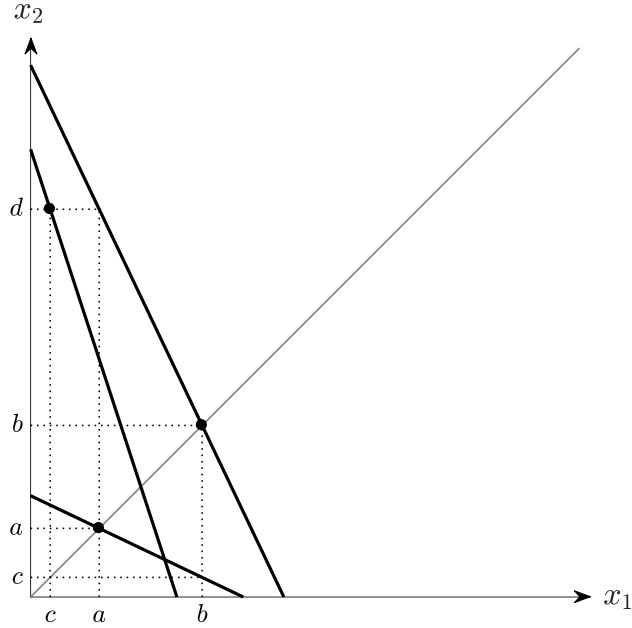
17

Figure 3: RDU-rationalizable but not EU-rationalizable data set

We claim that these observations are not EU-rationalizable if the two states are equiprobable. Suppose that they are, for some Bernoulli function u. Then the first observation tells us that $2u(a) \geqslant u(b) + u(c)$, since (b, c) is available when (a, a) is chosen. Similarly, from the second observation, we know that $2u(b) \geqslant u(a) + u(d)$. Together this gives

$$u(b) - u(d) \geqslant u(a) - u(b) \geqslant u(c) - u(a),$$

from which we obtain $u(a) + u(b) \geqslant u(c) + u(d)$. On the other hand, it is straightforward to check that, with $\epsilon > 0$, the bundle (a, b) is strictly cheaper than (c, d) at $\mathbf{p}^3$, which leads to a contradiction since (c, d) is chosen over (a, b) at the third observation.[20]

We claim that these observations are RDU-rationalizable; in fact, they can be rationalized with a smooth and concave Bernoulli function. Suppose $V(x_1, x_2) = \rho u(x_1) + (1 - \rho)u(x_2)$ when $x_1 \leqslant x_2$ and $V(x_1, x_2) = (1 - \rho)u(x_1) + \rho u(x_2)$ when $x_1 > x_2$, with $\rho = q/(q+1)$. Since $\rho > 1/2$, the agent displays disappointment aversion. So long as u is strictly concave, the agent's utility is maximized at $x_1 = x_2$ whenever $p_1 = 1$ and $p_2 \in [1/q, q]$. So V rationalizes the first two observations. To justify the third, it suffices to find u such that $u' > 0$ and

---

[20] Equivalently, note that there is a violation of Fishburn's condition. The bundle/lottery (a, a) is preferred to (b, c), (b, b) is preferred to (a, d), and (c, d) is *strictly* preferred to (a, b). However, the compound lottery where (a, a), (b, b) and (c, d) each occur with probability $1/3$ is stochastically equivalent to the compound lottery where (b, c), (a, d) and (a, b) each occur with probability $1/3$.

$u'' < 0$ satisfying the first order condition

$$\frac{\rho u'(c)}{(1-\rho)u'(d)} = q\frac{u'(c)}{u'(d)} = \frac{p_1^3}{p_2^3} = \frac{q^2}{1+q^2\epsilon}.$$

If ϵ is sufficiently small, this is possible since the price ratio p_1^3/p_2^3 is greater than q.[21] □

3. Goodness of fit

The revealed preference tests presented in the previous section are 'sharp', in the sense that a data set either passes the test for a given model or it fails. This either/or feature of the tests is not particular to our results but is true of all classical revealed preference tests, including Afriat's. It would, of course, be desirable to develop a way of measuring *the extent* to which a certain class of utility functions succeeds or fails in rationalizing a data set, and the most common approach adopted in the revealed preference literature to address this issue was developed by Afriat (1972, 1973) and Varian (1990).[22, 23] We now give an account of this approach and explain why implementing it in our setting is possible (or at least no more difficult than implementing the exact tests).

Suppose that the observer collects a data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$; following the earlier literature, we focus attention on the case where B^t is a classical linear budget set given by (1). For any number $e^t \in [0,1]$, we define

$$B^t(e^t) = \{\mathbf{x} \in \mathbb{R}_+^{\bar{s}} : \mathbf{p}^t \cdot \mathbf{x} \leqslant e^t \mathbf{p}^t \cdot \mathbf{x}^t\} \cup \{\mathbf{x}^t\}. \tag{12}$$

Clearly $B^t(e^t)$ is smaller than B^t and shrinks with the value of e^t. Let $\mathcal{U}$ be a collection of utility functions defined on $\mathbb{R}_+^{\bar{s}}$ belonging to a given family; for example, $\mathcal{U}$ could be the family of locally nonsatiated utility functions (which was the family considered by Afriat (1972, 1973) and Varian (1990) in their work). We define the set $E(\mathcal{U})$ in the following manner: a vector $\mathbf{e} = (e^1, e^2, \ldots, e^T)$ is in $E(\mathcal{U})$ if there is some function $U \in \mathcal{U}$ that rationalizes the

---

[21] If it were smaller than q, this would not be possible since the concavity of u requires $u'(c) \geqslant u'(d)$.

[22] For examples where Afriat-Varian type indices are used to measure a model's fit, see Mattei (2000), Harbaugh, Krause, and Berry (2001), Andreoni and Miller (2002), Choi *et al.* (2007, 2014), Beatty and Crawford (2011), and Halevy, Persitz, and Zrill (2016), and Pastor-Bernier, Plott, and Schultz (2017). Echenique, Lee, and Shum (2011) develops and applies a related index called the money pump index.

[23] For an account of why such measures may be more suitable then other measures of goodness-of-fit, such as the sum of squared errors between observed and predicted demands, see Varian (1990) and Halevy, Persitz, and Zrill (2016).

modified data set $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t))\}_{t=1}^T$, i.e., $U(\mathbf{x}^t) \geqslant U(\mathbf{x})$ for all $\mathbf{x} \in B^t(e^t)$. Clearly, the data set $\mathcal{O}$ is rationalizable by a utility function in $\mathcal{U}$ if and only if the unit vector $(1, 1, \ldots, 1)$ is in $E(\mathcal{U})$. We also know that $E(\mathcal{U})$ must be nonempty since it contains the vector $\mathbf{0}$, and it is clear that if $\mathbf{e} \in E(\mathcal{U})$ then $\mathbf{e}' \in E(\mathcal{U})$, where $\mathbf{e}' < \mathbf{e}$. The closeness of the set $E(\mathcal{U})$ to the unit vector is a measure of how well the utility functions in $\mathcal{U}$ can explain the data. Afriat (1972, 1973) suggests measuring this distance with the supnorm, so the distance between $\mathbf{e}$ and 1 is $D_A(\mathbf{e}) = 1 - \min_{1 \leqslant t \leqslant T} \{e^t\}$, while Varian (1990) suggests that we choose the square of the Euclidean distance, i.e., $D_V(\mathbf{e}) = \sum_{t=1}^T (1 - e^t)^2$.

Measuring distance by the supnorm has the advantage that it is computationally straightforward, and it is also the measure most commonly used in the empirical revealed preference literature, so this is the approach that we adopt in our implementation (see Section 4). Note that $D_A(\mathbf{e}) = D_A(\widetilde{\mathbf{e}})$ where $\widetilde{\mathbf{e}}$ is the vector with identical entries equal to $\min \{e^1, e^2, \ldots, e^T\}$, where $\mathbf{e} = (e^1, e^2, \ldots, e^T)$. Since $\widetilde{\mathbf{e}} \leqslant \mathbf{e}$, we obtain $\widetilde{\mathbf{e}} \in E(\mathcal{U})$ whenever $\mathbf{e} \in E(\mathcal{U})$. Therefore, $\min_{\mathbf{e} \in E(\mathcal{U})} D_A(\mathbf{e}) = \min_{\mathbf{e} \in \widetilde{E}(\mathcal{U})} D_A(\mathbf{e})$, where $\widetilde{E}(\mathcal{U}) = \{\mathbf{e} \in E(\mathcal{U}) : e^t = e^{t'} \text{ for any } t, t'\}$, i.e., in searching for $\mathbf{e} \in E(\mathcal{U})$ that minimizes the supnorm distance from $(1, 1, \ldots, 1)$, we can focus our attention on the set $\widetilde{E}(\mathcal{U})$, which consists of those vectors in $E(\mathcal{U})$ that shrink each observed budget set by the same proportion. Given a data set $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^T$, Afriat refers to $\sup \{e : (e, e, \ldots, e) \in E(\mathcal{U})\}$ as the *critical cost efficiency index*; we say that $\mathcal{O}$ is *rationalizable in $\mathcal{U}$ at the efficiency index/threshold e'* if $(e', e', \ldots, e') \in E(\mathcal{U})$.

Suppose that for a given data set, the critical cost efficiency index is 0.95. In that case, while we cannot guarantee that $\mathbf{x}^t$ is optimal in the true budget set, we know that there is some utility function in $\mathcal{U}$ for which $\mathbf{x}^t$ is optimal in $B^t(0.95)$ at every observation t. With this utility function, there could be bundles in B^t which the subject prefers to $\mathbf{x}^t$, but choosing such a bundle (instead of $\mathbf{x}^t$) will not lead to savings of more than 5%. Furthermore, this number is tight in the following sense: given any $\epsilon > 0$, then for every utility function in $\mathcal{U}$, there is at least one observation t where the subject could indeed have saved $(5 - \epsilon)\%$ of her expenditure.[24] We can interpret this index as a characteristic of the subject and, specifically, a measure of her bounded rationality; the bounded rationality could have arisen because she is simply incapable of better decision making, or it could be that she has consciously or

---

[24]Formally, for every $U \in \mathcal{U}$ there is t such that $\max \{U(\mathbf{x}) : \mathbf{p}^t \cdot \mathbf{x} \leqslant \mathbf{p}^t \cdot \mathbf{x}^t(0.95 + 0.01\epsilon)\} > U(\mathbf{x}^t)$.

otherwise judged that it is not, from a broader perspective, rational for her to expend the mental powers needed for exactly rational portfolio decisions.

Calculating the efficiency index (or, more generally, an index based on the Euclidean metric) will require checking whether a particular vector $\mathbf{e} = (e^1, e^2, \ldots, e^T)$ is in $E(\mathcal{U})$, i.e., whether $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t))\}_{t=1}^T$ is rationalizable by a member of $\mathcal{U}$. When $\mathcal{U}$ is the family of all locally nonsatiated utility functions, Afriat (1972, 1973) provides a necessary and sufficient condition for the rationalizability of $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t))\}_{t=1}^T$ (which we describe in greater detail in the Online Appendix).

More generally, the calculation of the efficiency index will hinge on whether there is a suitable test for the rationalizability of $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t))\}_{t=1}^T$ by members of $\mathcal{U}$. Even if a test of the rationalizability of $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ by members of $\mathcal{U}$ is available, this test may rely on the convexity or linearity of the budget sets B^t; in this case, extending the test so as to check for the rationalizability of $\mathcal{O}(\mathbf{e}) = \{(\mathbf{x}^t, B^t(e^t))\}_{t=1}^T$ is not straightforward since the modified budget sets $B^t(e^t)$ are clearly nonconvex. Crucially, this is *not* the case with the lattice method, which is applicable even for nonconvex constraint sets, so long as they are compact. Thus extending our testing procedure to measure goodness of fit in the form of the efficiency index involves no additional difficulties.

3.1 Approximate smooth rationalizability

While Theorem 1 guarantees that there is a Bernoulli function u that extends $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ and rationalizes the data when the required conditions are satisfied, the Bernoulli function is not necessarily smooth (though it is continuous and strictly increasing by definition). Of course, the smoothness of u is commonly assumed in applications of expected utility and related models and its implications can appear to be stark. For example, suppose that it is commonly known that states 1 and 2 occur with equal probability and we observe the agent choosing $(1, 1)$ at a price vector (p_1, p_2), with $p_1 \neq p_2$. This observation is incompatible with a smooth EU model; indeed, given that the two states are equiprobable, the slope of the indifference curve at $(1, 1)$ must equal -1 and thus it will not be tangential to the budget line and will not be a local optimum. On the other hand, it is trivial to check that this observation is EU-rationalizable in our sense. In fact, one could even find a *concave*

Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ for which $(1, 1)$ maximizes expected utility. (Such a u will, of course, have a kink at 1.)

These two facts can be reconciled by noticing that, even though this observation cannot be exactly rationalized by a smooth Bernoulli function, it is in fact possible to find a smooth function that comes arbitrarily close to rationalizing it. Indeed, given any strictly increasing and continuous function u defined on a compact interval of $\mathbb{R}_+$, there is a strictly increasing and smooth function $\tilde{u}$ that is uniformly and arbitrarily close to u on that interval. As such, if a Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}_+$ rationalizes $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ by $\{\phi(\cdot, t)\}_{t=1}^T$, then for any efficiency threshold $e \in (0, 1)$, there is a smooth Bernoulli function $\tilde{u} : \mathbb{R}_+ \to \mathbb{R}_+$ that rationalizes $\mathcal{O}' = \{(\mathbf{x}^t, B^t(e))\}_{t=1}^T$ by $\{\phi(\cdot, t)\}_{t=1}^T$. In other words, if a data set is rationalizable by some Bernoulli function, then it can also be rationalized by a smooth Bernoulli function, for any efficiency threshold arbitrarily close to 1. In this sense, imposing a smoothness requirement on the Bernoulli function does not radically alter a model's ability to explain a given data set.

4. Implementation

We examine the data collected from the well known portfolio choice experiment in Choi, Fisman, Gale, and Kariv (2007). The experiment was performed on 93 undergraduate subjects at the University of California, Berkeley. Every subject was asked to make consumption choices on 50 decision problems under risk. The subject divided her budget between two Arrow-Debreu securities, with each security paying one token if the corresponding state was realized, and zero otherwise. In a symmetric treatment applied to 47 subjects, each state of the world occurred with probability 1/2, and in a (balanced) asymmetric treatment applied to 46 subjects, the probabilities of the states were 1/3 and 2/3. These probabilities were objectively known. Lastly, income was normalized to one, and the state prices were chosen at random and varied across subjects.

In their analysis, Choi *et al.* (2007) first tested whether each subject could have been maximizing a locally nonsatiated utility function by performing a GARP test (or, strictly speaking, an extended version of the GARP test that characterizes rationalizability at a given efficiency threshold). Those subjects who passed GARP at a sufficiently high efficiency

threshold were then fitted individually to a two-parameter version of the disappointment aversion model of Gul (1991). The lattice method developed in this paper makes it possible for us to re-analyze the same data using purely revealed preference techniques, without appealing to any parametric assumptions. In this section we evaluate different models of decision making with these newly developed tests according to three criteria: (1) the ability of the model to explain the observed data; (2) the precision of the model's predictions (in various senses that we shall define); and (3) an index combining (1) and (2).

We consider five nested models in our empirical analysis. The most stringent of these is the expected utility (EU) model, followed by the disappointment aversion (DA) and rank dependent utility (RDU) models. (Recall from Section 2, however, that the RDU and DA models are identical when the two states are equiprobable.) We test these three models using the lattice method developed in Sections 2 and 3. In addition we test for basic rationalizability, i.e., consistency with locally nonsatiated utility maximization. A more stringent criterion is rationalizability by a stochastically monotone utility (SMU) function; this is a utility function that gives strictly higher utility to the bundle $\mathbf{x}$ compared to $\mathbf{y}$ whenever $\mathbf{x}$ first order stochatically dominates $\mathbf{y}$ (with respect to the objective probabilities attached to each state) and gives them the same utility whenever they are stochastically equivalent.

A test of consistency with a stochastically monotone utility function (at a given efficiency threshold) was recently developed by Nishimura, Ok, and Quah (2017); this test has features similar to GARP and we shall refer to it as F-GARP (see the Online Appendix for details). In the Choi *et al.* (2007) experiment, there are just two states. In this case 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, and when $\pi_2 > \pi_1$, a utility function U is stochastically monotone if and only if it is strictly increasing and $U(a, b) > U(b, a)$ whenever $b > a$. Notice that the RDU, DA, and EU models all obey this property; in fact, it is well known that these three models respect first order stochastic dominance, even when there are more than two states of the world.

4.1 Exact pass rates and efficiency indices

We first test all five models on the Choi *et al.* (2007) data, and the results from these exact tests are displayed in Table 1, where each cell contains a pass rate. Across 50 decision

$\pi_1 = 1/2$		$\pi_1 \neq 1/2$		Total	
GARP	12/47 (26%)	GARP	4/46 (9%)	GARP	16/93 (17%)
F-GARP	1/47 (2%)	F-GARP	3/46 (7%)	F-GARP	4/93 (4%)
RDU/DA	1/47 (2%)	RDU	2/46 (4%)	RDU	3/93 (3%)
		DA	1/46 (2%)	DA	2/93 (2%)
EU	1/47 (2%)	EU	1/46 (2%)	EU	2/93 (2%)

Table 1: Pass rates
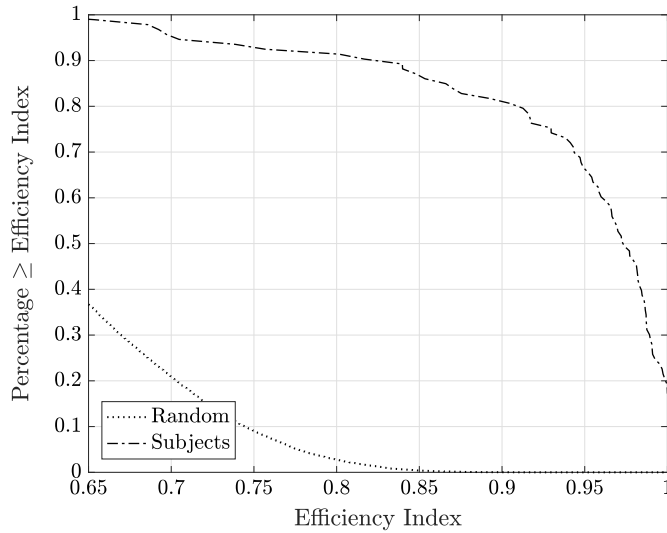
Figure 4: Distribution of efficiency indices for utility maximization

problems, 16 out of 93 subjects obey GARP and are therefore consistent with basic utility maximization; subjects in the symmetric treatment perform distinctly better than those in the asymmetric treatment. Of the 16 subjects who pass GARP, only 4 pass F-GARP; still fewer subjects are rationalizable by the RDU, DA, and EU models.

Given that we observe 50 decisions for every subject, it may not be intuitively surprising that so many subjects should have violated GARP (let alone more stringent conditions). We next investigate the efficiency thresholds at which subjects pass the different tests. First, we calculate the efficiency index at which each of the 93 subjects passes GARP; this empirical distribution is depicted in Figure 4. (Note that this figure is essentially a replication of Figure 4 in Choi *et al.* (2007).) We see that more than 80% of subjects have an efficiency index above 0.9, and more than 90% have an index above 0.8. A first glance at these results suggest that the data are largely compatible with the locally nonsatiated utility model.

To better understand what the observed distribution of efficiency indices says about the

24

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

Figure 5: Distributions of efficiency indices

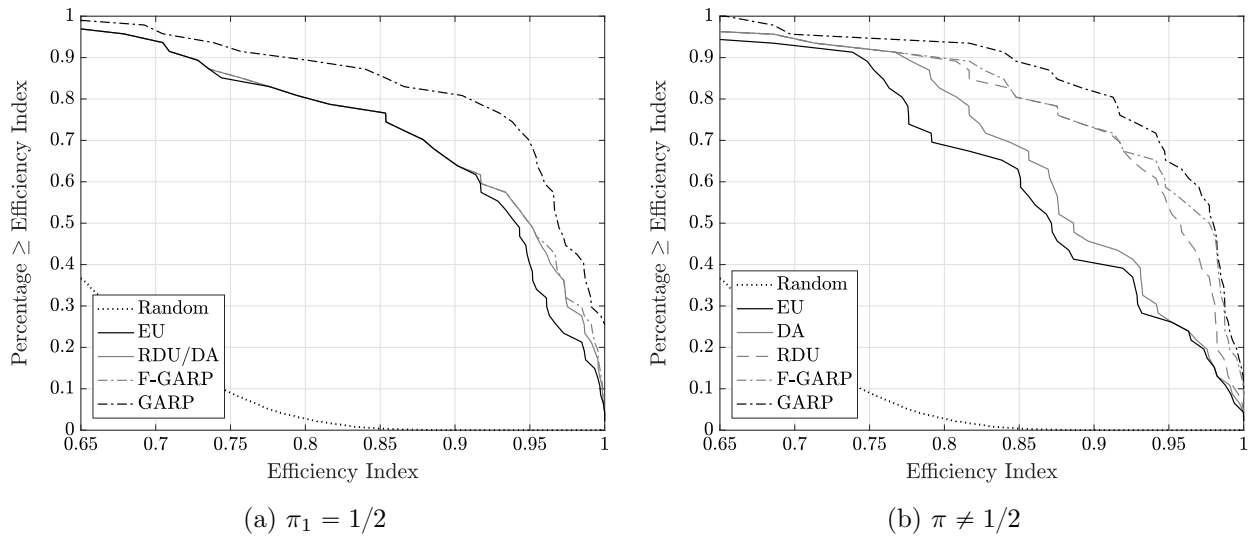success or failure of a particular model to explain the data collected, it is useful to see what distribution of efficiency indices will arise if we postulate an alternative form of behavior. We adopt an approach first suggested by Bronars (1987) that simulates random uniform consumption, i.e., which posits that consumers are choosing randomly uniformly from their budget frontiers. The Bronars (1987) approach has become common practice in the revealed preference literature as a way of assessing the *power* or *precision* of revealed preference tests. We follow exactly the procedure of Choi *et al.* (2007) and generate a random sample of 25,000 simulated subjects. Each simulated subject chooses randomly uniformly from 50 budget lines that are selected in the same random fashion as in the experimental setting. The dotted curve in Figure 4 corresponds to the distribution of efficiency indices for the simulated subjects. The experimental and simulated distributions are starkly different. For example, while 80% of subjects have an efficiency index of 0.9 or higher, the chance of a randomly generated data set passing GARP at an efficiency index of 0.9 is negligible. In other words, even though the locally nonsatiated utility model could accommodate much of the choice behavior observed in the experiment, it is also precise enough to exclude behavior that is simply randomly generated, which lends support to basic utility maximization as an accurate and discriminating model of choice among contingent consumption bundles.

Going *beyond* Choi *et al.* (2007), we then calculate the distributions of efficiency indices associated with the SMU, RDU, DA, and EU models among the 93 subjects. These

distributions are shown in Figures 5a and 5b, which correspond to the symmetric and asymmetric treatments, respectively. Since all of these models are more stringent than basic utility maximization, one would expect their efficiency indices to be lower, and they are. Nonetheless, at an efficiency threshold of 0.9, around half of all subjects are consistent with the EU model, with the proportion distinctly higher under the symmetric treatment. In the symmetric case, the performance of the EU, RDU/DA,[25] and SMU models are very close; in fact, their efficiency index distributions are almost indistinguishable. In the asymmetric case, the distinctions between models are sharper. The RDU and SMU models appear to perform considerably better than the EU and DA models, with their distributions of efficiency indices close to the distribution for the locally nonsatiated utility model. We have not depicted any efficiency index distributions for the SMU, RDU, DA, or EU models using randomly generated data, but plainly these will be even lower than for basic utility maximization and therefore very different from the distributions for the experimental subjects.

4.2 Pass rates

In order to compare the pass rates for different models more closely, we now concentrate on their performance at the 0.9 and 0.95 efficiency thresholds. These efficiency levels seem like reasonable standards that one might set in order to consider whether a model is consistent with the data; exact rationalizability is too stringent and anything less than 0.9 may be too permissive. The pass rates at these thresholds are presented in Table 2, where the models are arranged according to their generality, with the most permissive at the top. Assuming that the experimental subjects are a random sample drawn from a larger population of decision makers, we can use the sample pass rate for a model to estimate its expected population pass rate; these confidence intervals can be calculated *exactly* using the Clopper-Pearson procedure and are displayed in the Online Appendix.

We see from Table 2 that, at the 0.9 threshold, around 80% of all subjects pass GARP. Among this group, about half in turn display behavior that is consistent with the EU model (and, in fact, significantly more than half under the symmetric treatment). There is some evidence that the RDU model explains a significant number of subjects not captured by the EU model. In particular, in the asymmetric case, almost 90% of subjects who pass

---

[25] Recall that in the symmetric case, the RDU and DA models are identical.

	$\pi_1 = 1/2$				$\pi_1 \neq 1/2$	
	$e = 0.90$	$e = 0.95$			$e = 0.90$	$e = 0.95$
GARP	38/47 (81%)	32/47 (68%)		GARP	37/46 (80%)	29/46 (63%)
F-GARP	30/47 (64%)	23/47 (49%)		F-GARP	33/46 (72%)	26/46 (57%)
RDU/DA	30/47 (64%)	23/47 (49%)		RDU	33/46 (72%)	24/46 (52%)
				DA	20/46 (43%)	12/46 (26%)
EU	30/47 (64%)	18/47 (38%)		EU	18/46 (39%)	12/46 (26%)

Table 2: Pass rates by efficiency level

GARP at the 0.9 threshold are also consistent with the RDU model, which appears to be a significantly better performance than for the EU model. Indeed, there is very little room for the RDU model to perform better, since it manages to accommodate almost every subject who passes F-GARP at the same threshold. On the other hand, the DA model, which lies strictly between the RDU and EU models under the asymmetric treatment, does not perform significantly better than the EU model.[26]

We should be more precise on what we mean by a model performing 'significantly better' than another. We do not simply mean that the difference between the true (population) pass rates is distinct from zero; such a statistical claim is not meaningful when two models are nested.[27] We adopt a more stringent notion of 'significant' difference: we test the null hypothesis that the difference in expected pass rates between model A and model B *is equal to 5%*, against an alternative hypothesis that this difference is greater than 5%; since model B is nested within A, we are checking whether the additional data sets which are *accommodated by model A but not B* significantly exceeds 5%. The findings of these tests are reported in Table 3. For example, at the 0.95 threshold under the symmetric treatment, Table 2 tells us that the sample proportion of subjects who *pass the test for RDU/DA but fail the test for EU is 5/47*; this gives a p-value of 0.085, which is not statistically significant, i.e., we fail to reject the null hypothesis of a 5% difference at the 0.05 significance level.

The performance of the RDU model under the asymmetric treatment is quite different.

---

[26] While the contexts and methods are very different, the relatively poor performance of the DA model has been noted in some other studies, for example, Hey and Orme (1994) and Barseghyan *et al.* (2013).

[27] Suppose model A contains model B. Denoting the expected pass rates of model A (B) by μ_A (μ_B), the null hypothesis that $\mu_A = \mu_B$ is rejected if there is one data set which passes A but not B. Indeed, given that B is a special case of A, we are effectively testing the proportion of data sets which pass A and fail B; we conclude that the proportion of data sets *of this type* is nonzero as long as one such data set is observed.

$\pi_1 = 1/2$					$\pi_1 \neq 1/2$					
		GARP	F-GARP	RDU/DA			GARP	F-GARP	RDU	DA
$e = 0.95$	F-GARP	*0.000*	·	·	$e = 0.95$	F-GARP	0.406	·	·	·
	RDU/DA	*0.000*	1.000	·		RDU	0.079	0.677	·	·
						DA	*0.000*	*0.000*	*0.000*	·
	EU	*0.000*	0.085	0.085		EU	*0.000*	*0.000*	*0.000*	1.000
		GARP	F-GARP	RDU/DA			GARP	F-GARP	RDU	DA
$e = 0.90$	F-GARP	*0.002*	·	·	$e = 0.90$	F-GARP	0.197	·	·	·
	RDU/DA	*0.002*	1.000	·		RDU	0.197	1.000	·	·
						DA	*0.000*	*0.000*	*0.000*	·
	EU	*0.002*	1.000	1.000		EU	*0.000*	*0.000*	*0.000*	0.677
Note: Each cell contains a *p*-value.										

Table 3: Pairwise 5%-differences in pass rates

First, while the stochastically monotone utility model is theoretically more general, its pass rate is not significantly higher than 5% of that for the RDU model (at both efficiency thresholds). On the other hand, the pass rate for the RDU model compared to the EU model *does* significantly exceed 5%. Another way of saying the same thing is that if we are to form a 90% confidence interval on the expected proportion of subjects who are RDU-rationalizable but not EU-rationalizable, the lower bound of that interval exceeds 5%. What is that lower bound? At the 0.9 and 0.95 efficiency thresholds it is, respectively, 21% and 16%, which is sizeable by any reckoning. The RDU model generalizes the EU model by permitting a distortion of the objective probabilities; the Online Appendix contains more information on the type of distortions which are consistent with the data.

There is a large empirical literature that evaluates the performance of different models of choice under risk using experimental or field data and our results appear to be broadly in line with the findings obtained in earlier studies, even though the very different empirical methods employed make formal comparisons difficult. In particular, other papers have also found that the RDU model performs well (see, for example, Bruhin, Fehr-Duda, and Epper (2010) and Barseghyan *et al.* (2013) and their references). We find that the EU model captures a significant portion of subjects, though by no means everyone, which is broadly consistent with the fairly common finding that the EU model puts in a respectable performance (see, for example, Hey and Orme (1994)). The pass rate that we report for the EU model is higher than that in some other papers (for example, Bruhin, Fehr-Duda, and Epper (2010) reports

a pass rate of 20% for the EU model), but it is worth bearing in mind that our formulation of the EU model is about as permissive as it could get. We require the Bernoulli function to be increasing in money, but it is estimated nonparametrically and with no curvature assumptions (such as concavity), so we have given the EU model the greatest possible scope to capture a subject's behavior.

4.3 Power

To examine more closely the power of different models, we adopt and then adapt the approach first suggested by Bronars (1987). We first generate data sets where on each budget set, the bundle is randomly chosen based on a uniform distribution on the budget frontier. The *power* of the model (or its *Bronars power*) is then given by the probability of such a data set being inconsistent with (say) utility maximization, which is synonymous with it failing GARP. As we have already pointed out in Section 4.1, when the data set consists of uniform bundles on 50 randomly chosen budget sets, the Bronars power is approximately 1 at both the 0.9 and 0.95 efficiency thresholds; in other words, the probability of such a randomly generated data set passing GARP at either threshold is vanishingly small. Obviously, the Bronars power of the other models we consider is also roughly equal to 1, since all of them imply locally nonsatiated preferences.

When we consider a model that is theoretically more stringent than basic utility maximization, it is natural to investigate the power of the model *in the context of observed behavior that is already consistent with GARP*. In other words, we would like to know the sharpness of the model's predictions relative to basic utility maximization. For example, to check the relative power of the EU model in this sense, we randomly generate a large number of data sets that pass GARP at a given efficiency threshold, and then test if they obey EU at the same threshold. (See the Online Appendix for details.) Since the EU, DA, and RDU models are consistent with stochastic monotonicity, it is also natural to investigate the power of these models, *relative to the SMU model*; this would give us a sense of how stringent are the restrictions imposed by (say) the EU model, over and above those imposed by F-GARP.

Table 4 presents the power of the different models, conditional on passing GARP. The most obvious and important feature in this table is the ubiquity of numbers close to 1: even

29

$\pi_1 = 1/2$			$\pi_1 \neq 1/2$		
	$e = 0.90$	$e = 0.95$		$e = 0.90$	$e = 0.95$
F-GARP	1.00	1.00	F-GARP	0.88	0.92
RDU/DA	1.00	1.00	RDU	1.00	1.00
			DA	1.00	1.00
EU	1.00	1.00	EU	1.00	1.00

Table 4: Power (conditional on GARP)

$\pi_1 = 1/2$			$\pi_1 \neq 1/2$		
	$e = 0.90$	$e = 0.95$		$e = 0.90$	$e = 0.95$
RDU/DA	0.75	0.99	RDU	0.99	1.00
			DA	1.00	1.00
EU	0.87	0.99	EU	1.00	1.00

Table 5: Power (conditional on F-GARP)

after conditioning on passing GARP, all of the models remain very precise. For example, the probability of a data set which obeys GARP also passing the EU test at the 0.9 or 0.95 threshold is effectively zero. The only partial exception is for the SMU model in the asymmetric case, where the power is around 90%.

Table 5 tells us that all the models remain very precise relative to the SMU model when the treatment is asymmetric. But in the symmetric case, the relative power of the RDU/DA and EU models is noticeably lower than 1; for example, a quarter of all subjects who pass F-GARP at the 0.9 threshold are also consistent with the RDU/DA model. A possible reason for the loss of relative power in the symmetric case is that stochastic monotonicity itself is very restrictive in this context since it is synonymous with an increasing *and symmetric* utility function. However, we should emphasize that the relative power of the RDU, DA, and EU models remains high.

4.4 Predictive success

The *index of predictive success* proposed by Selten (1991) (or the Selten index, for short) combines pass rates and power into a single measure. This index is defined as the difference between the pass rate and the size of the set of predicted outcomes (the imprecision), with the latter typically measured by a uniform measure on all outcomes. Selten provides an

axiomatic foundation for this index. The index varies between 1 and -1. It is close to 1 when the pass rate is close to 1 and the imprecision is close to zero; in other words, even though the model is very precise in its predictions (i.e., has a high power), the data collected are very often consistent with the model. On the other hand, an index close to -1 occurs when the pass rate is close to zero even though the model is very imprecise (i.e., has a low power). An index above zero indicates that the model has some predictive success. Our use of the Selten index to evaluate different models is not novel. In the context of consumer demand (which formally is very similar to ours), it has been used by Beatty and Crawford (2011); it has also been used by Harless and Camerer (1994) to compare the performance of different models of choice under risk.

As we have emphasized, with 50 observations on every subject, the locally nonsatiated utility model has a power that is almost indistinguishable from 1 at the 0.9 and 0.95 efficiency thresholds. The same is true of course of all the other models, since they are more restrictive than basic utility maximization. In other words, all the models have an imprecision of zero, so that the Selten indices for the different models are effectively given by their pass rates, as displayed in Table 2. Note also that because the different models are nested within one another, any observed differences (from zero) in the pass rates/Selten indices in Table 2 are all statistically significant. One conclusion to be drawn from this table is that the best model, as evaluated by the Selten index, is the locally nonsatiated utility model. This observation may be simple but it is not without interest: while a great deal of academic discussion is often focussed on comparing different models that have been tailor-made for decision making under risk, we should *not* take it for granted that such models are necessarily better than basic utility maximization in explaining choice behavior. In environments where state payoffs vary while state probabilities are fixed, one should not exclude the possibility that the locally nonsatiated utility model does a better job in explaining the data, even after accounting for its relative lack of specificity.[28]

---

[28] This observation is not an argument against the value of models of choice under risk, such as the SMU, RDU, and all other models considered here. In particular, these other models provide a theory of choice across *all* lotteries, allowing even for comparisons between lotteries where the same outcomes occur with different probabilities. In environments where agents are making choices among lotteries of this type, all the other models are still potentially applicable, but it is not clear how one could naturally generalize the locally nonsatiated utility model to accommodate such a context.

$\pi_1 = 1/2$			$\pi_1 \neq 1/2$		
	$e = 0.90$	$e = 0.95$		$e = 0.90$	$e = 0.95$
F-GARP	0.79	0.72	F-GARP	0.77	0.81
RDU/DA	0.79	0.72	RDU	0.89	0.83
			DA	0.54	0.41
EU	0.79	0.56	EU	0.49	0.41

Table 6: Predictive success (conditional on GARP)

Our next objective is to investigate *conditional predictive success*. We first turn to the case where we condition on basic rationalizability. The conditional pass rate of each model can be calculated from Table 2. A model's imprecision is simply 1 minus the Bronars power (conditional on GARP) and this is supplied in Table 4. The Selten indices, constructed by taking the difference between the conditional pass rate and the conditional imprecision, are displayed in Table 6. For the symmetric treatment, all of the models have a conditional power of approximately 1 (see Table 4), so the Selten indices are nearly completely determined by the conditional pass rates and, as such, the differences between the SMU, RDU/DA, and EU models are not large. For the asymmetric treatment, the best performing model is RDU; this is driven by two factors: its pass rate is higher than all models except SMU and its power is higher than SMU. In the Online Appendix, we show that this difference between RDU model and the more restrictive DA and EU models are statistically significant.

Lastly, we investigate the predictive success of the EU, DA, and RDU models when we condition on passing F-GARP. The Selten indices displayed in Table 7 are obtained by taking the difference between the conditional pass rates (constructed from Table 2) and conditional power (from Table 5). Focussing firstly on the symmetric treatment, an interesting phenomenon is that, according to the Selten index, the EU model is now better than the RDU/DA model at the 0.9 efficiency threshold; this is entirely driven by the greater power of the EU model in this context. That said, the difference between the indices is not large and it is also reversed at the 0.95 threshold. For the asymmetric treatment, we notice that the RDU model performs well relative to the other models, because its pass rate is high *and* because the model continues to have high power, even after conditioning on passing F-GARP; we show in the Online Appendix that this difference is statistically significant.

	$\pi_1 = 1/2$			$\pi_1 \neq 1/2$	
	$e = 0.90$	$e = 0.95$		$e = 0.90$	$e = 0.95$
RDU/DA	0.75	0.99	RDU	0.99	0.92
			DA	0.61	0.46
EU	0.87	0.78	EU	0.55	0.46

Table 7: Predictive success (conditional on F-GARP)

Appendix

The proof of Theorem 1 uses the following lemma.

LEMMA 1. *Let* $\{C^t\}_{t=1}^T$ *be a finite collection of constraint sets in* $\mathbb{R}_+^{\bar{s}}$ *that are compact and downward closed (i.e., if* $\mathbf{x} \in C^t$ *then so is* $\mathbf{y} \in \mathbb{R}_+^{\bar{s}}$ *such that* $\mathbf{y} < \mathbf{x}$*) and let the functions* $\{\phi(\cdot, t)\}_{t=1}^T$ *be continuous and increasing in all dimensions. Suppose that there is a finite set* $\mathcal{X}$ *of* $\mathbb{R}_+$*, a strictly increasing function* $\bar{u} : \mathcal{X} \to \mathbb{R}_+$*, and* $\{M^t\}_{t=1}^T$ *such that the following holds:*

$$M^t \geqslant \phi(\bar{\mathbf{u}}(\mathbf{x}), t) \ \text{ for all } \mathbf{x} \in C^t \cap \mathcal{L} \text{ and} \tag{13}$$

$$M^t > \phi(\bar{\mathbf{u}}(\mathbf{x}), t) \ \text{ for all } \mathbf{x} \in (C^t \backslash \partial C^t) \cap \mathcal{L}, \tag{14}$$

where $\mathcal{L} = \mathcal{X}^{\bar{s}}$ *and* $\bar{\mathbf{u}}(\mathbf{x}) = (\bar{u}(x_1), \bar{u}(x_2), \dots, \bar{u}(x_{\bar{s}}))$*. Then there is a Bernoulli function* $u : \mathbb{R}_+ \to \mathbb{R}_+$ *that extends* $\bar{u}$ *such that*

$$M^t \geqslant \phi(\mathbf{u}(\mathbf{x}), t) \ \text{ for all } \mathbf{x} \in C^t \text{ and} \tag{15}$$

if $\mathbf{x} \in C^t$ *and* $M^t = \phi(\mathbf{u}(\mathbf{x}), t)$*, then* $\mathbf{x} \in \partial C^t \cap \mathcal{L}$ *and* $M^t = \phi(\bar{\mathbf{u}}(\mathbf{x}), t)$*. (16)*

REMARK: The property (16) needs some explanation. Conditions (13) and (14) allow for the possibility that $M^t = \phi(\bar{\mathbf{u}}(\mathbf{x}'), t)$ for some $\mathbf{x}' \in \partial C^t \cap \mathcal{L}$; we denote the set of points in $\partial C^t \cap \mathcal{L}$ with this property by X'. Clearly any extension u will preserve this property, i.e., $M^t = \phi(\mathbf{u}(\mathbf{x}'), t)$ for all $\mathbf{x}' \in X'$. Property (16) says that we can choose u such that for all $\mathbf{x} \in C^t \backslash X'$, we have $M^t > \phi(\mathbf{u}(\mathbf{x}), t)$.

Proof: We shall prove this result by induction on the dimension of the space containing the constraint sets. It is trivial to check that the claim is true if $\bar{s} = 1$. In this case, $\mathcal{L}$ consists of a finite set of points on $\mathbb{R}_+$ and each C^t is a closed interval with 0 as its minimum. Now

33

let us suppose that the claim holds for $\bar{s} = m$ and we shall prove it for $\bar{s} = m + 1$. If, for each t, there is a strictly increasing and continuous utility function $u^t : \mathbb{R}_+ \to \mathbb{R}_+$ extending $\bar{u}$ such that (15) and (16) hold, then the the same conditions will hold for the increasing and continuous function $u = \min_t u^t$. So we can focus our attention on constructing u^t for a single constraint set C^t.

Suppose $\mathcal{X} = \{0, r^1, r^2, r^3, \dots, r^I\}$, with $r^0 = 0 < r^i < r^{i+1}$, for $i = 1, 2, \dots, I - 1$. Let $\bar{r} = \max\{r \in \mathbb{R}_+ : (r, 0, 0, \dots, 0) \in C^t\}$ and suppose that $(r^i, 0, 0, \dots, 0) \in C^t$ if and only if $i \leqslant N$ (for some $N \leqslant I$). Consider the collection of sets of the form $D^i = \{\mathbf{y} \in \mathbb{R}_+^m : (r^i, \mathbf{y}) \in C^t\}$ (for $i = 1, 2, \dots, N$); this is a finite collection of compact and downward closed sets in $\mathbb{R}_+^m$. By the induction hypothesis applied to $\{D^i\}_{i=1}^N$, with $\{\phi(\bar{u}(r^i), \cdot, t)\}_{i=1}^N$ as the collection of functions, there is a strictly increasing function $u^* : \mathbb{R}_+ \to \mathbb{R}_+$ extending $\bar{u}$ such that

$$M^t \geqslant \phi(\bar{u}(r^i), \mathbf{u}^*(\mathbf{y}), t) \text{ for all } (r^i, \mathbf{y}) \in C^t \text{ and} \qquad (17)$$

if $(r^i, \mathbf{y}) \in C^t$ and $M^t = \phi(\bar{u}(r^i), \mathbf{u}^*(\mathbf{y}), t)$, then $(r^i, \mathbf{y}) \in \partial C^t \cap \mathcal{L}$ and $M^t = \phi(\bar{\mathbf{u}}(r^i, \mathbf{y}), t)$.

$$(18)$$

For each $r \in [0, \bar{r}]$, define

$$U(r) = \{u \leqslant u^*(r) : \max\{\phi(u, \mathbf{u}^*(\mathbf{y}), t) : (r, \mathbf{y}) \in C^t\} \leqslant M^t\}.$$

This set is nonempty; indeed $\bar{u}(r^k) = u^*(r^k) \in U(r)$, where r^k is the largest element in $\mathcal{X}$ that is weakly smaller than r. This is because, if $(r, \mathbf{y}) \in C^t$ then so is $(r^k, \mathbf{y})$, and (17) guarantees that $\phi(\bar{u}(r^k), \mathbf{u}^*(\mathbf{y}), t) \leqslant M^t$. The downward closedness of C^t and the fact that u^* is increasing also guarantees that $U(r) \subseteq U(r')$ whenever $r < r'$. Now define $\tilde{u}(r) = \sup U(r)$; the function $\tilde{u}$ has a number of significant properties. (i) *For $r \in \mathcal{X}$, $\tilde{u}(r) = u^*(r) = \bar{u}(r)$* (by the induction hypothesis). (ii) *$\tilde{u}$ is a nondecreasing function* since U is nondecreasing. (iii) *$\tilde{u}(r) > \bar{u}(r^k)$ if $r > r^k$, where r^k is largest element in $\mathcal{X}$ smaller than r.* Indeed, because C^t is compact and ϕ continuous, $\phi(\tilde{u}(r), \mathbf{u}^*(\mathbf{y}), t) \leqslant M^t$ for all $(r, \mathbf{y}) \in C^t$. By way of contradiction, suppose $\tilde{u}(r) = \bar{u}(r^k)$ and hence $\tilde{u}(r) < u^*(r)$. It follows from the definition of $\tilde{u}(r)$ that, for any sequence u_n, with $\tilde{u}(r) < u_n < u^*(r)$ and $\lim_{n \to \infty} u_n = \tilde{u}(r)$, there is $(r, \mathbf{y}_n) \in C^t$ such that $\phi(u_n, \mathbf{u}^*(\mathbf{y}_n), t) > M^t$. Since C^t is compact, we may assume with no loss of generality that $\mathbf{y}_n \to \hat{\mathbf{y}}$ and $(r, \hat{\mathbf{y}}) \in C^t$, from which we obtain $\phi(\tilde{u}(r), \mathbf{u}^*(\hat{\mathbf{y}}), t) = M^t$. Since C^t is downward closed, $(r^k, \hat{\mathbf{y}}) \in C^t$ and, since $\bar{u}(r^k) = u^*(r^k)$, we have $\phi(\mathbf{u}^*(r^k, \hat{\mathbf{y}}), t) = M^t$. This

34

can only occur if $(r^k, \hat{\mathbf{y}}) \in \partial C^t \cap \mathcal{L}$ (because of (18)), but it is clear that $(r^k, \hat{\mathbf{y}}) \notin \partial C^t$ since $(r^k, \hat{\mathbf{y}}) < (r, \hat{\mathbf{y}})$. (iv) *If $r_n < r^i$ for all n and $r_n \to r^i \in \mathcal{X}$, then $\tilde{u}(r_n) \to u^*(r^i)$.* Suppose to the contrary, that the limit is $\hat{u} < u^*(r^i) = \bar{u}(r^i)$. Since u^* is continuous, we can assume, without loss of generality, that $\tilde{u}(r_n) < u^*(r_n)$. By the compactness of C^t, the continuity of ϕ, and the definition of $\tilde{u}$, there is $(r_n, \mathbf{y}_n) \in C^t$ such that $\phi(\tilde{u}(r_n), \mathbf{u}^*(\mathbf{y}_n), t) = M^t$. This leads to $\phi(\hat{u}, \mathbf{u}^*(\mathbf{y}'), t) = M^t$, where $\mathbf{y}'$ is an accumulation point of $\mathbf{y}_n$ and $(r^i, \mathbf{y}') \in C^t$. But since ϕ is strictly increasing, we obtain $\phi(u^*(r^i), \mathbf{u}^*(\mathbf{y}'), t) > M^t$, which contradicts (17).

Given the properties of $\tilde{u}$, we can find a continuous and strictly increasing function u^t such that u^t extends $\bar{u}$, i.e., $u^t(r) = \bar{u}(r)$ for $r \in \mathcal{X}$, $u^t(r) < u^*(r)$ for all $r \in \mathbb{R}_+ \backslash \mathcal{X}$ and $u^t(r) < \tilde{u}(r) \leqslant u^*(r)$ for all $r \in [0, \bar{r}] \backslash \mathcal{X}$. (In fact we can choose u^t to be smooth everywhere except possibly on $\mathcal{X}$.) We claim that (15) and (16) are satisfied for C^t. To see this, note that for $r \in \mathcal{X}$ and $(r, \mathbf{y}) \in C^t$, the induction hypothesis guarantees that (17) and (18) hold and they will continue to hold if u^* is replaced by u^t. In the case where $r \notin \mathcal{X}$ and $(r, \mathbf{y}) \in C^t$, since $u^t(r) < \tilde{u}(r)$ and ϕ is increasing, we obtain $M^t > \phi(\mathbf{u}^t(r, \mathbf{y}), t)$.   **QED**

Proof of Theorem 1: This follows immediately from Lemma 1 if we set $C^t = \underline{B}^t$, and $M^t = \phi(\bar{\mathbf{u}}(\mathbf{x}^t), t)$. If $\bar{u}$ obeys conditions (6) and (7) then it obeys conditions (13) and (14). The rationalizability of $\mathcal{O}$ by $\{\phi(\cdot, t)\}_{t \in T}$ then follows from (15).   **QED**

Description of the RDU-rationalizability test for multiple states: Suppose that $\pi_s > 0$ is the objective probability of state s. To develop a necessary and sufficient test for RDU-rationalizability, we first define $\Gamma = \{\sum_{s \in S} \pi_s : S \subseteq \{1, 2, \ldots, \bar{s}\}\}$, i.e., Γ is a finite subset of $[0, 1]$ that includes both 0 and 1 (corresponding to S equal to the empty set and the whole set, respectively). Suppose *there are strictly increasing functions $\bar{g} : \Gamma \to \mathbb{R}$ and $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ such that (6) and (7) are satisfied,* with $\phi(\mathbf{u}) = \sum_{s=1}^{\bar{s}} \delta(\mathbf{u}, s) u_s$ and

$$\delta(\mathbf{u}, s) = \bar{g}\left(\sum_{\{s' : r(\mathbf{u}, s') \leqslant r(\mathbf{u}, s)\}} \pi_{s'}\right) - \bar{g}\left(\sum_{\{s' : r(\mathbf{u}, s') < r(\mathbf{u}, s)\}} \pi_{s'}\right).$$

By Theorem 1, this guarantees that $\mathcal{O}$ is RDU-rationalizable, with $g : [0, 1] \to \mathbb{R}$ chosen to be any strictly increasing extension of $\bar{g}$. This test involves finding a solution to a set of inequalities that are bilinear in the unknowns $\{\bar{g}(\gamma)\}_{\gamma \in \Gamma}$ and $\{\bar{u}(r)\}_{r \in \mathcal{X}}$. It is also clear that these conditions are necessary for RDU-rationalizability since they will be satisfied if we simply let $\bar{g}$ and $\bar{u}$ be the restrictions of g and u to Γ and $\mathcal{X}$ respectively.   **QED**

35

<center>REFERENCES</center>

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

AFRIAT, S. N. (1972): "Efficiency Estimation of Production Functions," *International Economic Review*, 13(3), 568–598.

AFRIAT, S. N. (1973): "On a System of Inequalities in Demand Analysis: An Extension of the Classical Method," *International Economic Review*, 14(2), 460–472.

AHN, D., S. CHOI, D. GALE, S. KARIV (2014): "Estimating Ambiguity Aversion in a Portfolio Choice Experiment," *Quantitative Economics*, 5(2), 195–223.

ALLAIS, P. M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 21(4), 503–546.

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

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

BARSEGHYAN, L., F. MOLINARI, T. O'DONOGHUE, and J. C. TEITELBAUM (2013): "The Nature of Risk Preferences: Evidence from Insurance Choices," *American Economic Review*, 103(6), 2499–2529.

BAYER, R.-C., S. BOSE, M. POLISSON, and L. RENOU (2013): "Ambiguity Revealed," *IFS Working Papers*, W13/05, Institute for Fiscal Studies.

BEATTY, T. K. M., and I. A. CRAWFORD (2011): "How Demanding Is the Revealed Preference Approach to Demand?," *American Economic Review*, 101(6), 2782–2795.

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

BROWN, D. J., and R. L. MATZKIN (1996): "Testable Restrictions on the Equilibrium Manifold," *Econometrica*, 64(6), 1249–1262.

BRUHIN, A., H. FEHR-DUDA, and T. EPPER (2010): "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica*, 78(4), 1375–1412.

BUSSIECK, M. R., and S. VIGERSKE (2010): "MINLP solver software," *Wiley Encyclopedia of Operations Research and Management Science*.

CAPPELEN, A. W., S. KARIV, E. SØRENSEN, and B. TUNGODDEN (2015): "Is There a Development Gap in Rationality?," Unpublished manuscript, UC Berkeley.

CHAMBERS, C. P., C. LIU, and S.-K. MARTINEZ (2016): "A Test for Risk-Averse Expected Utility," *Journal of Economic Theory*, 163, 775–785.

CHAMBERS, C. P., F. ECHENIQUE, and K. SAITO (2016): "Testing Theories of Financial Decision Making," *Proceedings of the National Academy of Sciences*, 113(15), 4003–4008.

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.

<center>36</center>

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

DIEWERT, W. E. (1973): "Afriat and Revealed Preference Theory," *Review of Economic Studies*, 40(3), 419–425.

DIEWERT, W. E. (2012): "Afriat's Theorem and Some Extensions to Choice Under Uncertainty," *Economic Journal*, 122(560), 305–331.

DYBVIG, P., and H. POLEMARCHAKIS (1981): "Recovering Cardinal Utility," *Review of Economic Studies*, 48(1), 159–166.

ECHENIQUE, F., and K. SAITO (2015): "Savage in the Market," *Econometrica*, 83(4), 1467–1495.

ECHENIQUE, F., S. LEE, and M. SHUM (2011): "The money pump as a measure of revealed preference violations," *Journal of Political Economy*, 119(6), 1201–1223.

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

FRIEDMAN, M., and L. J. SAVAGE (1948): "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.

GNEEZY, U., and J. POTTERS (1997): "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics*, 112(2), 631–645.

GREEN, R. C., and S. SRIVASTAVA (1986): "Expected Utility Maximization and Demand Behavior," *Journal of Economic Theory*, 38(2), 313–323.

GUL, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59(3), 667–686.

HALEVY, Y., D. PERSITZ, and L. ZRILL (2016): "Parametric Recoverability of Preferences," *Journal of Political Economy*, Forthcoming.

HARBAUGH, W. T., K. KRAUSE, and T. R. BERRY (2001): "GARP for Kids: On the Development of Rational Choice Behavior," *American Economic Review*, 91(5), 1539–1545.

HARLESS, D. W., and C. CAMERER (1994): "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 62(6), 1251–1289.

HEY, J. D., and C. ORME (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62(6), 1291–1326.

HEY, J. D., and N. PACE (2014): "The Explanatory and Predictive Power of Non Two-Stage-Probability Theories of Decision Making under Ambiguity," *Journal of Risk and Uncertainty*, 49(1), 1–29.

IMAI, T., and C. F. CAMERER (2016): "Estimating Time Preferences from Budget Set Choices Using Optimal Adaptive Design," Unpublished manuscript, Caltech.

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

KUBLER, F., L. SELDEN, and X. WEI (2014): "Asset Demand Based Tests of Expected Utility Maximization," *American Economic Review*, 104(11), 3459–3480.

LOOMES, G. (1991): "Evidence of a New Violation of the Independence Axiom," *Journal of Risk and Uncertainty*, 4(1), 91–108.

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

MATTEI, A. (2000): "Full-Scale Real Tests of Consumer Behavior Using Experimental Data," *Journal of Economic Behavior and Organization*, 43(4), 487–497.

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.

PASTOR-BERNIER, A., C. R. PLOTT, and W. SCHULTZ (2017): "Monkeys choose as if maximizing utility compatible with basic principles of revealed preference theory," *Proceedings of the National Academy of Sciences*, 114(10), 1766–1775.

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

SELTEN, R. (1991): "Properties of a Measure of Predictive Success," *Mathematical Social Sciences*, 21(2), 153–167.

VARIAN, H. R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50(4), 945–973.

VARIAN, H. R. (1983a): "Non-Parametric Tests of Consumer Behaviour," *Review of Economic Studies*, 50(1), 99–110.

VARIAN, H. R. (1983b): "Nonparametric Tests of Models of Investor Behavior," *Journal of Financial and Quantitative Analysis*, 18(3), 269–278.

VARIAN, H. R. (1988): "Estimating Risk Aversion from Arrow-Debreu Portfolio Choice," *Econometrica*, 56(4): 973–979.

VARIAN, H. R. (1990): "Goodness-of-Fit in Optimizing Models," *Journal of Econometrics*, 46(1–2), 125–140.

Online Appendix

to

Revealed preferences over risk and uncertainty

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

A1. Introduction

This Online Appendix consists of six parts. In Section A2, we discuss further applications of the lattice method which are not covered in the Main Text. In particular, we cover the choice acclimating personal equilibrium (CPE) model (Köszegi and Rabin, 2007), the maxmin expected utility (MEU) model (Gilboa and Schmeidler, 1989), the variational preference (VP) model (Maccheroni, Marinacci, and Rustichini, 2006), and a model with budget-dependent reference points. The other sections provide additional information on the empirical implementation in the main paper.

A3. Description of GARP and F-GARP Tests

A4. Confidence intervals on pass rates

A5. Bronars power calculation

A6. Probability distortions in the RDU model

A7. Statistical analysis of Selten index differences

A2. Further applications of the lattice method

Theorem 1 in the Main Text can also be used to test the rationalizability of many other models of choice under risk and under uncertainty. Formally, this involves finding a Bernoulli function u and a function ϕ belonging to some family Φ (corresponding to the particular model at hand) which together rationalize the data. In the subjective expected utility (SEU) case that was discussed in the Main Text, the lattice 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 $\{\pi_s\}_{s=1}^{\bar{s}}$. 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 lattice method. In this section we further illustrate the flexibility of the lattice method by applying it to several prominent models of decision making under risk or uncertainty.

A2.1 Choice acclimating personal equilibrium

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

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

$\boldsymbol{\pi} = \{\pi_s\}_{s=1}^{\bar{s}}$ are the objective probabilities, and $\lambda \in [0, 2]$ is the coefficient of loss aversion.[1] We say that a data set $\mathcal{O} = \{(\mathbf{x}^t, B^t)\}_{t=1}^T$ is CPE-rationalizable with the probability weights $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_{\bar{s}}) \gg 0$ if there is ϕ in the collection Φ_{CPE} of functions of the form (A.1), 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 in the Main Text, $\mathcal{O}$ *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 $\lambda \in [0, 2]$).

A2.2 Maxmin expected utility

---

[1] 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 again consider a setting where no objective probabilities can be attached to each state. An agent with maxmin expected utility (MEU), first presented by 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

$$\phi(\mathbf{u}) = \min_{\boldsymbol{\pi} \in \Pi} \left\{ \sum_{s=1}^{\bar{s}} \pi_s u_s \right\}, \tag{A.2}$$

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.2) 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.2), 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 in the Main Text, this holds if and only if there exist Π and $\bar{u}$ that solve (6), (7), and (8) in the Main Text. We claim that this requirement can be reformulated 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 $(x_1, x_2) \in \mathcal{L}$, 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 $\mathbf{x} \in \mathcal{L}$) 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.2)) and that $\bar{\boldsymbol{\pi}}(\mathbf{x})$ is the distribution in Π that minimizes the expected utility of the bundle $\mathbf{x}$ (see (A.6)).

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

$$\bar{\boldsymbol{\pi}}(\mathbf{x}^t) \cdot \bar{\mathbf{u}}(\mathbf{x}^t) \geqslant \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathcal{L} \cap \underline{B}^t, \tag{A.3}$$

3

$$\bar{\boldsymbol{\pi}}(\mathbf{x}^t) \cdot \bar{\mathbf{u}}(\mathbf{x}^t) > \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) \ \textit{for all } \mathbf{x} \in \mathcal{L} \cap (\underline{B}^t \backslash \partial \underline{B}^t), \textit{ and} \tag{A.4}$$

$$\bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) \leqslant \bar{\boldsymbol{\pi}}(\mathbf{x}') \cdot \bar{\mathbf{u}}(\mathbf{x}) \ \textit{for all } (\mathbf{x}, \mathbf{x}') \in \mathcal{L} \times \mathcal{L}. \tag{A.5}$$

If these conditions hold, $\mathcal{O}$ admits an MEU-rationalization where Π (in (A.2)) is the convex hull of $\{\bar{\boldsymbol{\pi}}(\mathbf{x})\}_{\mathbf{x} \in \mathcal{L}}$, the Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}$ extends $\bar{u}$, and

$$V(\mathbf{x}) = \min_{\boldsymbol{\pi} \in \Pi} \{\boldsymbol{\pi} \cdot \bar{\mathbf{u}}(\mathbf{x})\} = \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) \ \textit{for all } \mathbf{x} \in \mathcal{L}. \tag{A.6}$$

Proof: Suppose that $\mathcal{O}$ is rationalizable by ϕ as defined by (A.2). For any $\mathbf{x}$ in the finite lattice $\mathcal{L}$, 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.3)–(A.5) hold.

Conversely, suppose that there is a function $\bar{\boldsymbol{\pi}}$ and a strictly increasing function $\bar{u}$ obeying the conditions (A.3)–(A.5). Define Π as the convex hull of $\{\bar{\boldsymbol{\pi}}(\mathbf{x}) : \mathbf{x} \in \mathcal{L}\}$; Π is a nonempty and convex subset of Δ_{++} and it is compact in $\mathbb{R}^{\bar{s}}$ since $\mathcal{L}$ is finite. Suppose that there exists $\mathbf{x} \in \mathcal{L}$ 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 $\{\bar{\boldsymbol{\pi}}(\mathbf{x}) : \mathbf{x} \in \mathcal{L}\}$, there must exist $\mathbf{x}' \in \mathcal{L}$ 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.5). 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 $\mathbf{x} \in \mathcal{L}$. 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.3) and (A.4) are just versions of (6) and (7) in the Main Text, and so Theorem 1 in the Main Text guarantees that there is 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**

A2.3 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

$$\phi(\mathbf{u}) = \min_{\boldsymbol{\pi} \in \Delta_{++}} \{\boldsymbol{\pi} \cdot \mathbf{u} + c(\boldsymbol{\pi})\} \tag{A.7}$$

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 minimization problem in (A.7).[2] 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.7), 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 in the Main Text, 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 an increasing function $\bar{u} : \mathcal{X} \to \mathbb{R}_+$ that together solve (6) and (7) in the Main Text, with ϕ defined by (A.7). The following result is a reformulation of this characterization that has a similar flavor to Proposition A.1; crucially, the necessary and sufficient conditions on $\mathcal{O}$ are formulated as a finite set of bilinear inequalities.

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

$$\bar{\boldsymbol{\pi}}(\mathbf{x}^t) \cdot \bar{\mathbf{u}}(\mathbf{x}^t) + \bar{c}(\mathbf{x}^t) \geqslant \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}) \ \ \text{for all } \mathbf{x} \in \mathcal{L} \cap \underline{B}^t, \tag{A.8}$$

$$\bar{\boldsymbol{\pi}}(\mathbf{x}^t) \cdot \bar{\mathbf{u}}(\mathbf{x}^t) + \bar{c}(\mathbf{x}^t) > \bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}) \ \ \text{for all } \mathbf{x} \in \mathcal{L} \cap (\underline{B}^t \backslash \partial \underline{B}^t), \text{ and} \tag{A.9}$$

$$\bar{\boldsymbol{\pi}}(\mathbf{x}) \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}) \leqslant \bar{\boldsymbol{\pi}}(\mathbf{x}') \cdot \bar{\mathbf{u}}(\mathbf{x}) + \bar{c}(\mathbf{x}') \ \ \text{for all } (\mathbf{x}, \mathbf{x}') \in \mathcal{L} \times \mathcal{L}. \tag{A.10}$$

If these conditions hold, then $\mathcal{O}$ can be rationalized by a variational preference V, with ϕ given by (A.7), such that the following holds:

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

(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 $\mathbf{x} \in \mathcal{L}$.*

Proof: Suppose $\mathcal{O}$ is rationalizable by ϕ as defined by (A.7). Let $\bar{u}$ be the restriction of u to $\mathcal{X}$. For any $\mathbf{x}$ in $\mathcal{L}$, 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.8)–(A.10) hold.

---

[2] 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.

Conversely, suppose that there is a strictly increasing function $\bar{u}$ and functions $\bar{\boldsymbol{\pi}}$ and $\bar{c}$ obeying conditions (A.8)–(A.10). For every $\boldsymbol{\pi} \in \Delta_{++}$, define $\tilde{c}(\boldsymbol{\pi}) = \max_{\mathbf{x} \in \mathcal{L}} \{\bar{c}(\mathbf{x}) - (\boldsymbol{\pi} - \bar{\boldsymbol{\pi}}(\mathbf{x})) \cdot \bar{\mathbf{u}}(\mathbf{x})\}$. It follows from (A.10) 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 $\mathbf{x} \in \mathcal{L}$. Therefore, $\tilde{c}(\bar{\boldsymbol{\pi}}(\mathbf{x}')) = \bar{c}(\mathbf{x}')$ for any $\mathbf{x}' \in \mathcal{L}$. 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 $\boldsymbol{\pi} \in \{\bar{\boldsymbol{\pi}}(\mathbf{x}) : \mathbf{x} \in \mathcal{L}\}$. 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 $\mathbf{x} \in \mathcal{L}$. Indeed, for any $\boldsymbol{\pi} \in \Delta_{++}$,

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

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.7); then (A.8) and (A.9) are just versions of (6) and (7) in the Main Text, and so Theorem 1 in the Main Text 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**

A2.4 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 an experiment in which subjects are asked to choose from a constraint set of state contingent monetary prizes. Assuming that there are $\bar{s}$ states and that the subject never suffers a loss, we can represent each prize by a vector $\mathbf{x} \in \mathbb{R}_+^{\bar{s}}$. The subject is observed to choose $\mathbf{x}^t$ from $B^t \subset \mathbb{R}_+^{\bar{s}}$, 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 his choice from B^t is governed by a preference defined on the prizes, which implies that the situation where he never receives a prize (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 $\mathbf{e}^t \in \mathbb{R}^{\bar{s}}_+$, 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

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

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* $\{\mathbf{e}^t\}_{t=1}^T$ if there exists a Bernoulli function $u : [-K, \infty) \to \mathbb{R}_+$ such that $\phi(\mathbf{u}(\mathbf{x}^t - \mathbf{e}^t), t) \geqslant \phi(\mathbf{u}(\mathbf{x} - \mathbf{e}^t), t)$ for all $\mathbf{x} \in B^t$. This is formally equivalent to saying that the modified data set $\mathcal{O}' = \{(\mathbf{x}^t - \mathbf{e}^t, \tilde{B}^t)\}_{t=1}^T$ is rationalizable by $\{\phi(\cdot, t)\}_{t=1}^T$. Applying Theorem 1 in the Main Text, 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

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

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. Description of GARP and F-GARP tests

In addition to the expected utility (EU), disappointment aversion (DA), and rank dependent utility (RDU) models which we implement in the Main Text, there are other more

basic notions of rationalizabiity that we also test. In this section, we describe these models and their revealed preference tests.

<center>*A3.1 Locally nonsatiated utility*</center>

The locally nonsatiated utility model is the most permissive of the models that we consider in the sense that all others are special cases. 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 say that a data set $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^T$ is *rationalizable* if it can be rationalized by a continuous and locally nonsatiated utility function.

Afriat's (1967) Theorem tells us that $\mathcal{O}$ is rationalizable if and only if it obeys a consistency condition known as the *generalized axiom of revealed preference* (GARP).[3] Afriat (1972, 1973) also shows that there is natural generalization of GARP that characterizes rationalizability at some efficiency index e, which we now describe. Let $\mathcal{D} = \{\mathbf{x}^t : t = 1, 2, \ldots, T\}$; in other words, $\mathcal{D}$ consists of those bundles that have been observed somewhere 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 index (or threshold) e (we denote this by $\mathbf{x}^t \succcurlyeq_e^* \mathbf{x}^{t'}$) if $\mathbf{x}^{t'} \in B^t(e)$, where $B^t(e)$ is given by (12) in the Main Text;[4] $\mathbf{x}^t$ is said to be *strictly revealed preferred to* $\mathbf{x}^{t'}$ (and we denote this by $\mathbf{x}^t \succ_e^* \mathbf{x}^{t'}$) if $\mathbf{x}^{t'} \in B^t(e)$ and $\mathbf{p}^t \cdot \mathbf{x}^{t'} < e \, \mathbf{p}^t \cdot \mathbf{x}^t$. $\mathcal{O}$ is rationalizable at the efficiency index e 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

$$\mathbf{x}^{t_1} \succcurlyeq_e^* \mathbf{x}^{t_2}, \ \ \mathbf{x}^{t_2} \succcurlyeq_e^* \mathbf{x}^{t_3}, \ \ldots \ , \ \mathbf{x}^{t_{n-1}} \succcurlyeq_e^* \mathbf{x}^{t_n}, \text{ and } \mathbf{x}^{t_n} \succcurlyeq_e^* \mathbf{x}^{t_1}, \qquad (A.11)$$

then we cannot replace $\succcurlyeq_e^*$ with $\succ_e^*$ anywhere in this chain; in other words, while there can be revealed preference cycles in $\mathcal{O}$, they cannot contain a strict revealed preference. This property is a generalization of GARP, which is the special case where $e = 1$. Checking for this property is computationally undemanding: the (strict) revealed preference relations on $\mathcal{D}$ can be easily constructed; once this has been established, we can apply Warshall's

---

[3] This term and its acronym were coined by Varian (1982), who also provides a proof of Afriat's Theorem. To be specific, the theorem says that GARP is necessary whenever the data set is rationalizable by a locally nonsatiated utility function (continuity is not needed); conversely, when a data set obeys GARP, then it is rationalizable by a continuous, strictly increasing, and concave utility function.

[4] Our terminology differs a little from the standard, which refers to $\succcurlyeq_e^*$ 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.

algorithm to compute the transitive closure of the revealed preference relations and then check for the absence of cycles containing strict revealed preferences.

A3.2 Stochastically monotone utility

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 $U : \mathbb{R}_+^{\bar{s}} \to \mathbb{R}$ is *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 RDU, DA, and EU 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 $\pi_2 > \pi_1$, a utility function U is stochastically monotone if and only if it is strictly increasing and $U(a, b) > U(b, a)$ whenever $b > a$.

A data set $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t=1}^T$ is said to be rationalizable by the stochastically monotone utility (SMU) model if there is a continuous and stochastically monotone utility function U that rationalizes the observations. Since a utility function U that is stochastically monotone will be strictly increasing, it is also locally nonsatiated. Hence any SMU-rationalizable data set is also rationalizable by utility maximization but the converse is not true. Indeed, the single observation given in Example 1 passes GARP trivially, but it cannot be rationalized by any symmetric and strictly increasing utility function.

Nishimura, Ok, and Quah (2017) have recently developed a test for rationalizability by the SMU model. The test can be thought of as a version of GARP, but with suitably modified revealed preference relations. We say that the bundle $\mathbf{x}^t$ *is SMU-revealed preferred to* $\mathbf{x}^{t'}$ at the efficiency threshold e (for $\mathbf{x}^t$ and $\mathbf{x}^{t'}$ in $\mathcal{D}$) if there is a bundle $\mathbf{y}$ such that $\mathbf{y} \in B^t(e)$ and $\mathbf{y} \geqslant_{FSD} \mathbf{x}^{t'}$; this revealed preference is *strict* if $\mathbf{y}$ can be chosen to satisfy $\mathbf{y} >_{FSD} \mathbf{x}^t$. Nishimura, Ok, and Quah (2017) show that a data set is rationalizable by the SMU model at a threshold e if and only if it does not admit SMU-revealed preference cycles (such as (A.11)) containing strict SMU-revealed preferences; we call the latter property F-GARP (at the efficiency threshold e), where 'F' stands for first order stochastic dominance. Clearly this

9

result is analogous to the characterization for basic rationalizability, except that the revealed preferences are defined differently. With two states, the SMU-revealed preference relations are easily obtained, and therefore checking F-GARP is also easy to implement.

A4. Confidence intervals on pass rates

	$\pi_1 = 1/2$			$\pi_1 \neq 1/2$	
	$e = 0.90$	$e = 0.95$		$e = 0.90$	$e = 0.95$
GARP	0.81 [0.67, 0.91]	0.68 [0.53, 0.81]	GARP	0.80 [0.66, 0.91]	0.63 [0.48, 0.77]
F-GARP	0.64 [0.49, 0.77]	0.49 [0.34, 0.64]	F-GARP	0.72 [0.57, 0.84]	0.57 [0.41, 0.71]
RDU/DA	0.64 [0.49, 0.77]	0.49 [0.34, 0.64]	RDU	0.72 [0.57, 0.84]	0.52 [0.37, 0.67]
			DA	0.43 [0.29, 0.59]	0.26 [0.14, 0.41]
EU	0.64 [0.49, 0.77]	0.38 [0.25, 0.54]	EU	0.39 [0.25, 0.55]	0.26 [0.14, 0.41]
Note: Each cell contains a pass rate and *exact* 95% confidence interval (in square brackets), where the latter is obtained using the Clopper-Pearson procedure.					

Table A.1: Pass rates and 95% confidence intervals

	$\pi_1 = 1/2$			$\pi_1 \neq 1/2$	
	$e = 0.90$	$e = 0.95$		$e = 0.90$	$e = 0.95$
F-GARP	0.79 [0.63, 0.90]	0.72 [0.53, 0.86]	F-GARP	0.89 [0.75, 0.97]	0.90 [0.73, 0.98]
RDU/DA	0.79 [0.63, 0.90]	0.72 [0.53, 0.86]	RDU	0.89 [0.75, 0.97]	0.83 [0.64, 0.94]
			DA	0.54 [0.37, 0.71]	0.41 [0.24, 0.61]
EU	0.79 [0.63, 0.90]	0.56 [0.38, 0.74]	EU	0.49 [0.32, 0.66]	0.41 [0.24, 0.61]
Note: Each cell contains a pass rate and *exact* 95% confidence interval (in square brackets), where the latter is obtained using the Clopper-Pearson procedure.					

Table A.2: Pass rates and 95% confidence intervals (conditional on GARP)

A5. Bronars power calculation

In order to calculate the Bronars (1987) power, we need to generate random data sets. As we described in the Main Text, this involves first producing budget sets in the same random fashion as in the Choi *et al.* (2007) experiment itself, and then randomly selecting bundles from these budget sets. In the case where we are interested in power unconditionally, our algorithm simply selects bundles randomly uniformly from the frontiers of these budget sets. While the unconditional procedure needs little explanation, the method for calculating conditional power, i.e., power conditional on passing the generalized axiom of revealed preference (GARP) or F-GARP, requires further explanation.

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 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,[5] before subjecting each data set to a test for a given model. 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 our estimate of the Bronars power is within 0.01 of the true value.

A6. 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) analyzed in the Main Text, 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 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.

---

[5] So there are four distinct collections of data sets, with each collection containing 30,000 data sets.

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 2 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 $[0.45, 0.5)$.

A7. STATISTICAL ANALYSIS OF SELTEN INDEX DIFFERENCES

A statistical analysis of the differences in the Selten indices conditional on GARP is provided in Table A.3. (Table A.4 provides the same analysis after conditioning on F-GARP.) Each entry in the table gives the p-value of the test of null hypothesis that the two models have the same Selten index, with the alternative hypothesis that they do not. To illustrate how this test is carried out, consider a test of the equality of the Selten indices between the EU and SMU models, under the asymmetric treatment and at the efficiency threshold 0.9. As shown in Table 4, the SMU model has a power of 0.88 and the EU model has a power of 1; we take this as given. The null hypothesis that the Selten indices are equal is equivalent to the hypothesis that $\mu_{SMU} - \mu_{EU} = 1 - 0.88 = 0.12$, where μ_{EU} and μ_{SMU} are the expected conditional pass rates for the EU and SMU models. The alternative hypothesis is $\mu_{SMU} - \mu_{EU} \neq 0.12$. The sample estimate of $\mu_{SMU} - \mu_{EU}$ has a binomial distribution and its realized value is $(33 - 18)/37$; according to Table A.3, the probability of obtaining this sample estimate or something more extreme if $\mu_{SMU} - \mu_{EU} = 0.12$ is effectively zero.

12

	π₁ = 1/2				π₁ ≠ 1/2		
$e = 0.95$		F-GARP	RDU/DA	$e = 0.95$		F-GARP	RDU	DA
	RDU/DA	1.000	.		RDU	0.432	.	.
					DA	*0.000*	*0.000*	.
	EU	*0.000*	*0.000*		EU	*0.000*	*0.000*	1.000
$e = 0.90$		F-GARP	RDU/DA	$e = 0.90$		F-GARP	RDU	DA
	RDU/DA	1.000	.		RDU	*0.005*	.	.
					DA	*0.002*	*0.000*	.
	EU	1.000	1.000		EU	*0.000*	*0.000*	*0.000*

Note: Each cell contains a *p*-value, with values below 0.05 italicized.

Table A.3: Pairwise differences in predictive success
(conditional on GARP)

	π₁ = 1/2			π₁ ≠ 1/2	
$e = 0.95$		RDU/DA	$e = 0.95$		RDU	DA
				DA	*0.000*	.
	EU	*0.000*		EU	*0.000*	1.000
$e = 0.90$		RDU/DA	$e = 0.90$		RDU	DA
				DA	*0.000*	.
	EU	*0.003*		EU	*0.000*	*0.000*

Note: Each cell contains a *p*-value.

Table A.4: Pairwise differences in predictive success
(conditional on F-GARP)

Notice that, in the asymmetric case, the differences in Selten indices between the RDU model and the more restrictive DA and EU models are all statistically significant. This is true after conditioning on GARP, and also after conditioning on F-GARP (see Tables A.3 and A.4 respectively).

REFERENCES

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

AFRIAT, S. N. (1972): "Efficiency Estimation of Production Functions," *International Economic Review*, 13(3), 568–598.

AFRIAT, S. N. (1973): "On a System of Inequalities in Demand Analysis: An Extension of the Classical Method," *International Economic Review*, 14(2), 460–472.

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.

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

KÖSZEGI, 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 (2014): "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.

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

VARIAN, H. R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50(4), 945–973.