

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

**Working Paper**

## Revealed preferences over risk and uncertainty

IFS Working Papers, No. W15/25

**Provided in Cooperation with:**

Institute for Fiscal Studies (IFS), London

*Suggested Citation:* Polisson, Matthew; Quah, John K.-H.; Renou, Ludovic (2015) : Revealed preferences over risk and uncertainty, IFS Working Papers, No. W15/25, Institute for Fiscal Studies (IFS), London,  
<https://doi.org/10.1920/wp.ifs.2015.1525>

This Version is available at:

<http://hdl.handle.net/10419/145447>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Revealed preferences over risk and uncertainty

IFS Working Paper W15/25

Matthew Polisson  
John Quah  
Ludovic Renou



The Institute for Fiscal Studies (IFS) is an independent research institute whose remit is to carry out rigorous economic research into public policy and to disseminate the findings of this research. IFS receives generous support from the Economic and Social Research Council, in particular via the ESRC Centre for the Microeconomic Analysis of Public Policy (CPP). The content of our working papers is the work of their authors and does not necessarily represent the views of IFS research staff or affiliates.

# Revealed preferences over risk and uncertainty

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

September 12, 2015

**Abstract:** Consider a finite data set where each observation consists of a bundle of contingent consumption chosen by an agent from a constraint set of such bundles. We develop a general procedure for testing the consistency of this data set with a broad class of models of choice under risk and under uncertainty. Unlike previous work, we do not require that the agent has a convex preference, so we allow for risk loving and elation seeking behavior. Our procedure can also be extended to calculate the magnitude of violations from a particular model of choice, using an index first suggested by Afriat (1972, 1973). We then apply this index to evaluate different models (including expected utility and disappointment aversion) in the data collected by Choi *et al.* (2007). We show that among those subjects exhibiting choice behavior consistent with the maximization of some increasing utility function, more than half are consistent with models of expected utility and disappointment aversion.

**Keywords:** expected utility, rank dependent utility, maxmin expected utility, variational preferences, generalized axiom of revealed preference

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

---

\* Email addresses: matthew.polisson@le.ac.uk; john.quah@economics.ox.ac.uk; lrenou@essex.ac.uk.

This research made use of the ALICE High Performance Computing Facility at the University of Leicester. Part of this research was carried out while Matthew Polisson was a Visiting Associate at Caltech and while Ludovic Renou was a Visiting Professor at the University of Queensland, and they would like to thank these institutions for their hospitality and support. We are grateful to Don Brown, Ray Fisman, Felix Kubler, Shachar Kariv, and Collin Raymond for helpful discussions and comments, as well as to seminar audiences at Adelaide, Brown, Caltech, Leicester, NUS, NYU Stern, Oxford, PSE, Toulouse, UC Irvine, UCSD, Queensland, and QUT, and to conference participants at the 2014 Canadian Economic Association Annual Conference (Simon Fraser), the 2014 North American Summer Meeting of the Econometric Society (Minnesota), the 2014 SAET Conference (Waseda), the 2014 Workshop on Nonparametric Demand (IFS and CeMMAP), the 2014 Dauphine Workshop in Economic Theory, the 2015 Australasian Economic Theory Workshop (Deakin), the 2015 Bath Economic Theory Workshop, the 2015 Canadian Economic Theory Conference (Western Ontario), the 2015 ESRC Workshop on Decision Theory and Applications (Essex), the 2015 Cowles Summer Conference on Heterogeneous Agents and Microeconometrics (Yale), the 2015 SAET Conference (Cambridge), and the 2015 Econometric Society World Congress (Montreal).

## 1. INTRODUCTION

Let  $\mathcal{O} = \{(p^t, x^t)\}_{t=1}^T$  be a finite set, where  $p^t \in \mathbb{R}_{++}^{\bar{s}}$  and  $x^t \in \mathbb{R}_{+}^{\bar{s}}$ . We interpret  $\mathcal{O}$  as a set of observations, where  $x^t$  is the observed bundle of  $\bar{s}$  goods chosen by an agent (the *demand bundle*) at the price vector  $p^t$ . A utility function  $U : \mathbb{R}_{+}^{\bar{s}} \rightarrow \mathbb{R}$  is said to *rationalize*  $\mathcal{O}$  if, at every observation  $t$ ,  $x^t$  is the bundle that maximizes  $U$  in the budget set  $B^t = \{x \in \mathbb{R}_{+}^{\bar{s}} : p^t \cdot x \leq p^t \cdot x^t\}$ . For any data set that is rationalizable by a locally nonsatiated utility function, its revealed preference relations must satisfy a no-cycling condition called the generalized axiom of revealed preference (GARP). Afriat's (1967) Theorem shows that any data set that obeys GARP will in turn be rationalizable by a utility function that is continuous, strictly increasing, and concave. This result is very useful because it provides a nonparametric test of utility maximization that can be readily implemented in observational and experimental settings. It is computationally straightforward to check GARP directly; alternatively, GARP holds if and only if there is a solution to a set of linear inequalities constructed from the data and one could also check if this system has a solution.<sup>1</sup>

It is both useful and natural to develop tests, similar to the one developed by Afriat, for alternative hypotheses on agent behavior. Our objective in this paper is to develop a procedure that is useful for testing models of choice under risk and under uncertainty. Retaining the formal setting described in the previous paragraph, we can interpret  $\bar{s}$  as the number of states of the world, with  $x^t$  a bundle of contingent consumption, and  $p^t$  the state prices faced by the agent. In a setting like this, we can ask what conditions on the data set are necessary and sufficient for it to be consistent with an agent who is maximizing an expected utility function. This means that  $x^t$  maximizes the agent's expected utility, compared to other bundles in the budget set. Assuming that the probability of state  $s$  is commonly known to be  $\pi_s$ , this involves recovering a Bernoulli function  $u : \mathbb{R}_{+} \rightarrow \mathbb{R}$ , which we require to be continuous and strictly increasing, so that, for each  $t = 1, 2, \dots, T$ ,

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

---

<sup>1</sup> For proofs of Afriat's Theorem, see Afriat (1967), Diewert (1973), Varian (1982), and Fostel, Scarf, and Todd (2004). The term GARP is from Varian (1982); Afriat refers to the same property as *cyclical consistency*.

In the case where the state probabilities are subjective and unknown to the observer, it would be necessary to recover both  $u$  and  $\{\pi_s\}_{s=1}^{\bar{s}}$  so that (1) holds.

In fact, tests of this sort have already been developed by Varian (1983a, 1983b, 1988) and Green and Srivastava (1986). Such tests involve solving a set of inequalities that are derived from the data; there is consistency with expected utility maximization if and only if a solution to these inequalities exists. However, these results (and later generalizations and variations, including those on other models of choice under risk and under uncertainty<sup>2</sup>) rely on two crucial assumptions: (1) the agent’s utility function is concave, and (2) the budget set  $B^t$  takes the classical form defined above, where prices are linear and markets are complete. These two assumptions guarantee that the first order conditions are necessary *and* sufficient for optimality and can in turn be converted into a necessary and sufficient test. The use of concavity to simplify the formulation of revealed preference tests is well known and can be applied to models of choice in other contexts (see Diewert (2012)).

Our contribution in this paper is to develop a testing procedure that has the following features: **(i)** it is potentially adaptable to test for different models of choice under risk and under uncertainty, and not just the expected utility model; **(ii)** it is a ‘pure’ test of a given model as such and does not require the *a priori* exclusion of certain ‘behavioral’ phenomena, such as risk loving or elation seeking or reference point effects, that lead to a nonconcave Bernoulli utility function, or more generally, nonconvex preferences over contingent consumption; **(iii)** it is applicable to situations with complex budgetary constraints and can be employed even when there is market incompleteness or when there are nonconvexities in the budget set due to nonlinear pricing or other practices;<sup>3</sup> and **(iv)** it can be easily adapted to measure ‘near’ rationalizability (using the indices developed by Afriat (1972, 1973) and Varian (1990)) in cases where the data set is not exactly rationalizable by a particular model.

In the case of objective expected utility maximization, a data set is consistent with this model if and only if there is a solution to a set of linear inequalities. In the case of, for example, subjective expected utility, rank dependent utility, or maxmin expected utility, our test involves solving a finite set of bilinear inequalities that is constructed from

---

<sup>2</sup> See Diewert (2012), Bayer *et al.* (2013), Kubler, Selden, and Wei (2014), Echenique and Saito (2015), Chambers, Liu, and Martinez (2015), and Chambers, Echenique, and Saito (2015).

<sup>3</sup> For an extension Afriat’s Theorem to nonlinear budget constraints, see Forges and Minelli (2009).

the data. These problems are 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 certain cases can be computationally straightforward because of certain special features and/or when the number of observations is small.<sup>4</sup> 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.<sup>5</sup>

### 1.1 The lattice test procedure

We now give a brief description of our test. Given a data set  $\mathcal{O} = \{(p^t, x^t)\}_{t=1}^T$ , we define the discrete consumption set  $\mathcal{X} = \{x' \in \mathbb{R}_+ : x' = x_s^t \text{ for some } t, s\} \cup \{0\}$ . Besides zero,  $\mathcal{X}$  contains those levels of consumption that were chosen at some observation and at some state. Since  $\mathcal{O}$  is finite, so is  $\mathcal{X}$ , and its product  $\mathcal{L} = \mathcal{X}^{\bar{s}}$  forms a grid of points in  $\mathbb{R}_+^{\bar{s}}$ ; in formal terms,  $\mathcal{L}$  is a finite lattice. For example, Figure 1 depicts a data set where  $x^1 = (2, 5)$  at  $p^1 = (5, 2)$ ,  $x^2 = (6, 1)$  at  $p^2 = (1, 2)$ , and  $x^3 = (4, 3)$  at  $p^3 = (4, 3)$ . In this case,  $\mathcal{X} = \{0, 1, 2, 3, 4, 5, 6\}$  and  $\mathcal{L} = \mathcal{X} \times \mathcal{X}$ .

Suppose we would like to test whether  $\mathcal{O}$  is consistent with expected utility maximization given objective probabilities  $\{\pi_s\}_{s=1}^{\bar{s}}$  that are known to us. Clearly, a *necessary* condition for this to hold is that we can find a set of real numbers  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  with the following properties: (i)  $\bar{u}(r'') > \bar{u}(r')$  whenever  $r'' > r'$ , and (ii) for every  $t = 1, 2, \dots, T$ ,

$$\sum_{s=1}^{\bar{s}} \pi_s \bar{u}(x_s^t) \geq \sum_{s=1}^{\bar{s}} \pi_s \bar{u}(x_s) \quad \text{for all } x \in B^t \cap \mathcal{L}, \quad (2)$$

<sup>4</sup> It is not uncommon to perform tests on fewer than 20 observations. This is partly because revealed preference tests do not in general account for errors, which are unavoidable across many observations.

<sup>5</sup> 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)) that are potentially useful for implementing bilinear tests more generally.

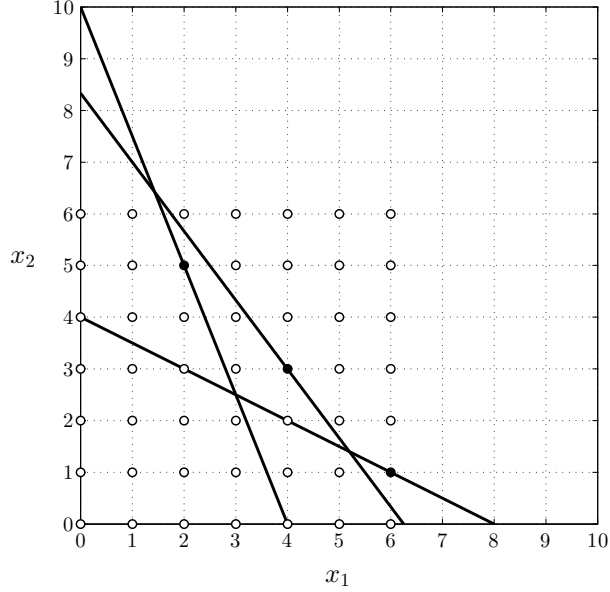


Figure 1: Constructing the finite lattice

with the inequality strict whenever  $x \in B^t \cap \mathcal{L}$  and  $x$  is in the interior of the budget set  $B^t$ . Since  $\mathcal{X}$  is finite, the existence of  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  with properties (i) and (ii) can be straightforwardly ascertained by solving a family of linear inequalities. Our main result says that *if a solution can be found, then there is a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}$  that extends  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  and satisfies (1)*. Returning to the example depicted in Figure 1, suppose we know that  $\pi_1 = \pi_2 = 0.5$ . Our test requires that we find  $\bar{u}(r)$ , for  $r = 0, 1, 2, \dots, 6$ , such that the expected utility of the chosen bundle  $(x_1^t, x_2^t)$  is (weakly) greater than that of any lattice points within the corresponding budget set  $B^t$ . One could check that these requirements are satisfied for  $\bar{u}(r) = r$ , for  $r = 0, 1, \dots, 6$ , so we conclude that the data set is consistent with expected utility maximization. A detailed explanation of our testing procedure and its application to the expected utility model is found in Section 2. Section 3 shows how this procedure can be applied to test for other models of choice under risk and under uncertainty, including rank dependent utility, choice acclimating personal equilibrium, maxmin expected utility, and variational preferences.

### 1.2 Empirical implementation

To illustrate the use of these tests, we implement them in Section 5 on a data set obtained from the portfolio choice experiment in Choi *et al.* (2007). 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 performed GARP tests on the data; subjects who passed or came very close to passing (and whose behavior was therefore broadly consistent with the maximization of an increasing utility function) were then fitted to a parameterized model of disappointment aversion.

In our analysis, we test four utility-maximizing models on the data, using a nonparametric approach throughout. In decreasing order of generality, the models are increasing utility, stochastically monotone utility, disappointment aversion, and expected utility. Consistency with the increasing utility model can be verified by checking GARP. By *stochastically monotone* utility, we mean that the utility function is strictly increasing with respect to first order stochastic dominance; this hypothesis is stronger than saying that the utility function is strictly increasing and it can be tested via a stronger version of GARP recently formulated by Nishimura, Ok, and Quah (2015).<sup>6</sup> Lastly, the tools developed in this paper allow us to test the disappointment aversion and expected utility models.

Given that there are 50 observations on every subject, it is not empirically meaningful to simply carry out *exact* tests, because the overwhelming majority of subjects are likely to fail even the GARP test (which is the most permissive of the four). What is required is a way of measuring how close each subject is to being consistent with a particular model of behavior. In the case of the increasing utility model, Choi *et al.* (2007) measured this gap using the *critical cost efficiency index* (CCEI). This index was first proposed by Afriat (1972, 1973) who also showed how GARP can be modified to calculate the index. We extend this approach by calculating CCEIs for all four models of interest. Not all revealed preference tests can be straightforwardly adapted to perform CCEI calculations; the fact that it is possible to modify our tests of the disappointment aversion and expected utility models for this purpose is one of its important features and is due to the fact that these tests can be performed

---

<sup>6</sup> When the two states are equiprobable, a utility function is stochastically monotone if and only if it is strictly increasing and symmetric. If state 1 has a probability 2/3, then the utility function  $U$  is stochastically monotone if and only if it is strictly increasing and  $U(a, b) > U(b, a)$  whenever  $a > b$ . See Section 5 for details.

on nonconvex budget sets. We explain this in greater detail in Section 4, which discusses CCEIs. We also determine the *power* (Bronars, 1987) of each model, i.e. the probability of a random data set being inconsistent with that model (at a given efficiency threshold). This information allows us to rank the performance of each model using Selten’s (1991) index of predictive success. This index balances the success of a model in predicting observations (which favors the increasing utility model since it is the most permissive) with the specificity of its predictions (which favors expected utility since it is the most restrictive).

Our main findings are as follows. **(i)** In the context of the Choi *et al.* (2007) experiment, all four models have high power in the sense that the probability of a randomly drawn data set being consistent with the model is close to zero. **(ii)** Measured by the Selten index, the best performing model is increasing utility, followed by stochastically monotone utility, disappointment aversion, and expected utility. In other words, the greater success that the increasing utility model has in explaining a set of observations more than compensates for its relative lack of specificity. That said, all four models have considerable success in explaining the data; for example, at an efficiency level of 0.9, the pass rates of the four models are 81%, 68%, 54%, and 52%, respectively. **(iii)** Conditioning on agents who pass GARP (at or above some efficiency level, say 0.9), both disappointment aversion and expected utility remain very precise, i.e., the probability of a randomly drawn data set satisfying disappointment aversion (and hence expected utility), conditional on it passing GARP, is also very close to zero. **(iv)** On the other hand, more than half of the subjects who pass GARP are also consistent with the disappointment aversion and expected utility models, which gives clear support for these models in explaining the behavior of agents who are (in the first place) maximizing some increasing utility function.

## 2. TESTING THE MODEL ON A LATTICE

We assume that there is a finite set of states, denoted by  $S = \{1, 2, \dots, \bar{s}\}$ . The contingent consumption space is  $\mathbb{R}_+^{\bar{s}}$ ; for a typical consumption bundle  $x \in \mathbb{R}_+^{\bar{s}}$ , the  $s$ th entry,  $x_s$ , specifies the consumption level in state  $s$ .<sup>7</sup> We assume that there are  $T$  observations in the data set

---

<sup>7</sup> 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

$\mathcal{O}$ , where  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$ . This means that the agent is observed to have chosen the bundle  $x^t$  from the set  $B^t \subset \mathbb{R}_+^{\bar{s}}$ . We assume that  $B^t$  is compact and that  $x^t \in \partial B^t$ , where  $\partial B^t$  denotes the *upper boundary* of  $B^t$ . An element  $y \in B^t$  is in  $\partial B^t$  if there is no  $x \in B^t$  such that  $x > y$ .<sup>8,9</sup> The most important example of  $B^t$  is the standard budget set when markets are complete, i.e.,

$$B^t = \{x \in \mathbb{R}_+^{\bar{s}} : p^t \cdot x \leq p^t \cdot x^t\}, \quad (3)$$

with  $p^t \gg 0$  being the vector of state prices. Our formulation also allows for the market to be incomplete. Suppose that the agent's contingent consumption is achieved through a portfolio of securities and that the asset prices do not admit arbitrage; then it is well known that there exists a price vector  $p^t \gg 0$  such that

$$B^t = \{x \in \mathbb{R}_+^{\bar{s}} : p^t \cdot x \leq p^t \cdot x^t\} \cap \{Z + \omega\},$$

where  $Z$  is the span of assets available to the agent and  $\omega$  is his endowment of contingent consumption. Note that both  $B^t$  and  $x^t$  will be known to the observer so long as he can observe the asset prices and the agent's holding of securities, the asset payoffs in every state, and the agent's endowment of contingent consumption,  $\omega$ .

Let  $\{\phi(\cdot, t)\}_{t=1}^T$  be a collection of functions, where  $\phi(\cdot, t) : \mathbb{R}_+^{\bar{s}} \rightarrow \mathbb{R}$  is continuous and strictly increasing.<sup>10</sup> The data set  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is said to be *rationalizable* by  $\{\phi(\cdot, t)\}_{t=1}^T$  if there exists a continuous and strictly increasing function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , which we shall refer to as the *Bernoulli function*, such that

$$\phi(\mathbf{u}(x^t), t) \geq \phi(\mathbf{u}(x), t) \text{ for all } x \in B^t, \quad (4)$$

where  $\mathbf{u}(x) = (u(x_1), u(x_2), \dots, u(x_{\bar{s}}))$ . In other words,  $x^t$  is an optimal choice in  $B^t$ , assuming that the agent is maximizing  $\phi(\mathbf{u}(x), t)$ . 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  $\phi(\mathbf{u}(x), t)$  is continuous and strictly increasing in  $x$ , so that this model is a special case of

---

Wei (2014), Echenique and Saito (2015), and Chambers, Echenique, and Saito (2015). The papers by Varian (1983), Green and Srivastava (1986), Bayer *et al.* (2013), and Chambers, Liu, and Martinez (2015) allow for multi-dimensional realizations but they also require the convexity of the agent's preference over contingent consumption and linear budget sets.

<sup>8</sup>For vectors  $x, y \in \mathbb{R}_+^{\bar{s}}$ ,  $x > y$  if  $x \neq y$  and  $x_i \geq y_i$  for all  $i$ . If  $x_i > y_i$  for all  $i$ , we write  $x \gg y$ .

<sup>9</sup>For example, if  $B^t = \{(x, y) \in \mathbb{R}_+^2 : (x, y) \leq (1, 1)\}$ , then  $(1, 1) \in \partial B^t$  but  $(1, 1/2) \notin \partial B^t$ .

<sup>10</sup>By strictly increasing, we mean that  $\phi(z, t) > \phi(z', t)$  if  $z > z'$ .

the increasing utility model. The continuity of the utility function is an important property, because it guarantees that the agent's utility maximization problem always has a solution on a compact constraint set.<sup>11</sup> Many of the basic models of choice under risk and under uncertainty can be described within this framework, with different models leading to different functional forms for  $\phi(\cdot, t)$ . Of course, this includes expected utility, as we show in the example below. For some of these models (such as rank dependent utility (see Section 3.1)),  $\phi(\cdot, t)$  can be a nonconcave function, in which case the agent's preference over contingent consumption may be nonconvex, even if  $u$  is concave.

*Example:* 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,

$$\phi(u_1, u_2, \dots, u_{\bar{s}}, t) = \sum_{s=1}^{\bar{s}} \pi_s^t u_s, \quad (5)$$

and (4) requires that

$$\sum_{s=1}^{\bar{s}} \pi_s^t u(x_s^t) \geq \sum_{s=1}^{\bar{s}} \pi_s^t u(x_s) \text{ for all } x \in B^t, \quad (6)$$

i.e., the expected utility of  $x^t$  is greater than that of any other bundle in  $B^t$ . When there exists a Bernoulli function  $u$  such that (6) holds, we say that the data set is *EU-rationalizable with probability weights*  $\{\pi^t\}_{t=1}^T$ , where  $\pi^t = (\pi_1^t, \pi_2^t, \dots, \pi_{\bar{s}}^t)$ .

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  $x$ , we must have

$$\phi(\mathbf{u}(x^t), t) \geq \phi(\mathbf{u}(x), t) \text{ for all } x \in \underline{B}^t, \quad (7)$$

where  $\underline{B}^t = \{y \in \mathbb{R}_+^{\bar{s}} : y \leq x \text{ for some } x \in B^t\}$ . Furthermore, the inequality in (7) is strict whenever  $x \in \underline{B}^t \setminus \partial \underline{B}^t$  (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 at some state. Since the data set is finite, so is  $\mathcal{X}$ . Given  $\mathcal{X}$ , we may construct  $\mathcal{L} = \mathcal{X}^{\bar{s}}$ ,

---

<sup>11</sup> 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.

which consists of a finite grid of points in  $\mathbb{R}_+^{\bar{s}}$ ; in formal terms,  $\mathcal{L}$  is a finite lattice. Let  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  be the restriction of the Bernoulli function  $u$  to  $\mathcal{X}$ . Given our observations, the following must hold:

$$\phi(\bar{\mathbf{u}}(x^t), t) \geq \phi(\bar{\mathbf{u}}(x), t) \text{ for all } x \in \underline{B}^t \cap \mathcal{L} \text{ and} \quad (8)$$

$$\phi(\bar{\mathbf{u}}(x^t), t) > \phi(\bar{\mathbf{u}}(x), t) \text{ for all } x \in (\underline{B}^t \setminus \partial \underline{B}^t) \cap \mathcal{L}, \quad (9)$$

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

**THEOREM 1.** *Suppose that for some data set  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  and collection of continuous and strictly increasing functions  $\{\phi(\cdot, t)\}_{t=1}^T$ , there is a strictly increasing function  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  that satisfies conditions (8) and (9). Then there is a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  that extends  $\bar{u}$  and guarantees the rationalizability of  $\mathcal{O}$  by  $\{\phi(\cdot, t)\}_{t=1}^T$ .<sup>12</sup>*

The intuition for this result ought to be strong. Given  $\bar{u}$  satisfying (8) and (9), we can define the step function  $\hat{u} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  where  $\hat{u}(r) = \bar{u}([r])$ , with  $[r]$  being the largest element of  $\mathcal{X}$  weakly lower than  $r$ , i.e.,  $[r] = \max\{r' \in \mathcal{X} : r' \leq r\}$ . Notice that  $\phi(\hat{\mathbf{u}}(x^t), t) = \phi(\bar{\mathbf{u}}(x^t), t)$  and, for any  $x \in \underline{B}^t$ ,  $\phi(\hat{\mathbf{u}}(x), t) = \phi(\bar{\mathbf{u}}([x]), t)$ , where  $[x] = ([x_1], [x_2], \dots, [x_{\bar{s}}])$  in  $\underline{B}^t \cap \mathcal{L}$ . Clearly, if  $\bar{u}$  obeys (8) and (9) then  $\mathcal{O}$  is rationalized by  $\{\phi(\cdot, t)\}_{t=1}^T$  and  $\hat{u}$  (in the sense that (4) 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 added properties.

### 2.1 Testing the expected utility (EU) model

We wish to check whether  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is EU-rationalizable with probability weights  $\{\pi^t\}_{t=1}^T$ , in the sense defined in the previous example. By Theorem 1, EU-rationalizability holds if and only if there is a collection of real numbers  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  such that

$$0 \leq \bar{u}(r') < \bar{u}(r) \text{ whenever } r' < r, \quad (10)$$

and the inequalities (8) and (9) hold, where  $\phi(\cdot, t)$  is defined by (5). This is a linear program and it is both *solvable* (in the sense that there is an algorithm that can decide within a known number of steps whether or not a solution exists) and computationally feasible.

<sup>12</sup> The increasing assumptions on  $\phi$  and  $\bar{u}$  ensure that we may confine ourselves to checking (8) and (9) for undominated elements of  $\underline{B}^t \cap \mathcal{L}$ , i.e.,  $x \in \underline{B}^t \cap \mathcal{L}$  such that there does not exist  $x' \in \underline{B}^t \cap \mathcal{L}$  with  $x < x'$ .

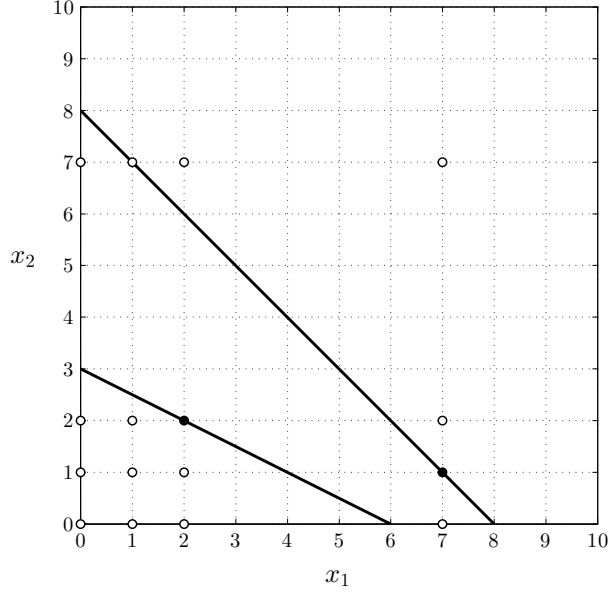


Figure 2: Violation of concave expected utility

Note that the Bernoulli function, whose existence is guaranteed by Theorem 1, need not be a concave function. Consider the example given in Figure 2 and suppose that  $\pi_1 = \pi_2 = 1/2$ . In this case,  $\mathcal{X} = \{0, 1, 2, 7\}$ , and one could check that (8) and (9) are satisfied (where  $\phi(\cdot, t)$  is defined by (5)) if  $\bar{u}(0) = 0$ ,  $\bar{u}(1) = 2$ ,  $\bar{u}(2) = 3$ , and  $\bar{u}(7) = 6$ . Thus the data set is EU-rationalizable. However, any Bernoulli function that rationalizes the data cannot also be concave. Indeed, since  $(3, 1)$  is strictly within the budget set when  $(2, 2)$  was chosen,  $2u(2) > u(1) + u(3)$ . By the concavity of  $u$ ,  $u(3) - u(2) \geq u(7) - u(6)$ , and thus we obtain  $u(6) + u(2) > u(7) + u(1)$ , contradicting the optimality of  $(7, 1)$ .

We now turn to a setting in which no objective probabilities can be attached to each state. The data set  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is said to be *rationalizable by subjective expected utility* (SEU) if there exist beliefs  $\pi = (\pi_1, \pi_2, \dots, \pi_{\bar{s}}) \gg 0$  and a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that, for all  $t = 1, 2, \dots, T$ ,

$$\sum_{s=1}^{\bar{s}} \pi_s u(x_s^t) \geq \sum_{s=1}^{\bar{s}} \pi_s u(x_s) \text{ for all } x \in B^t. \quad (11)$$

There is no loss of generality in requiring  $\sum_{s=1}^{\bar{s}} \pi_s = 1$ , in which case,  $\pi_s$  may be interpreted as the (subjective) probability of state  $s$ . In this case,  $\phi$  is independent of  $t$  and instead of being fixed, it is only required to belong to a particular family of functions. Formally, let  $\Phi_{SEU}$  be the family of functions  $\phi$  such that  $\phi(\mathbf{u}) = \sum_{s=1}^{\bar{s}} \pi_s u_s$  for some  $\pi \gg 0$ . By definition,

the data set  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is SEU-rationalizable if there is  $\phi \in \Phi_{SEU}$  such that (11) holds. The conditions (8) and (9) can be written as

$$\sum_{s=1}^{\bar{s}} \pi_s \bar{u}(x_s^t) \geq \sum_{s=1}^{\bar{s}} \pi_s \bar{u}(x_s) \quad \text{for all } x \in \underline{B}^t \cap \mathcal{L} \text{ and} \quad (12)$$

$$\sum_{s=1}^{\bar{s}} \pi_s \bar{u}(x_s^t) > \sum_{s=1}^{\bar{s}} \pi_s \bar{u}(x_s) \quad \text{for all } x \in (\underline{B}^t \setminus \partial \underline{B}^t) \cap \mathcal{L}. \quad (13)$$

Therefore, a necessary and (by Theorem 1) sufficient condition for SEU-rationalizability is that we can find real numbers  $\{\pi_s\}_{s=1}^{\bar{s}}$  and  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  such that  $\pi_s > 0$  for all  $s$ ,  $\sum_{s=1}^{\bar{s}} \pi_s = 1$ , and (10), (12), and (13) are satisfied. Notice that we are simultaneously searching for  $u$  and (through  $\pi_s > 0$ )  $\phi \in \Phi_{SEU}$  that rationalize the data. This set of conditions forms a finite system of inequalities that are bilinear in  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  and  $\{\pi_s\}_{s=1}^{\bar{s}}$ . The Tarski-Seidenberg Theorem tells us that such systems are decidable.

In the case when there are just two states, there is a straightforward way of implementing this test. Simply condition on the probability of state 1 (and hence the probability of state 2), and then perform a linear test to check if there is a solution to (10), (12), and (13). If not, choose another probability for state 1, implement the test, and repeat, if necessary. Even a search of up to two decimal places on state 1's probability will lead to no more than 99 linear tests, which can be implemented with little difficulty.

### 3. OTHER APPLICATIONS OF THE LATTICE TEST

Theorem 1 can also be used to test the rationalizability of other models of choice under risk and under uncertainty. Formally, this involves finding a Bernoulli function  $u$  and a function  $\phi$  belonging to some family  $\Phi$  (corresponding to that particular model) that rationalize the data. In the SEU case, we know that the 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 that 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 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 model. In particular, this is true for each of the models discussed in this section.

We illustrate the flexibility of our approach with a fairly long list of models, but readers who are more interested in the details and results of our specific empirical implementation need not (at least initially) read through this section. They should read Section 3.1 since it contains a discussion of the disappointment aversion model which we later implement; after this, they may skip the rest of the section and proceed straight to Section 4.

### 3.1 Rank dependent utility (RDU)

The rank dependent utility model (Quiggin, 1982) is a model of choice under risk where, for each state  $s$ , there is an objective probability  $\pi_s > 0$  that is known to the agent (and which we assume is also known to the observer). Given a vector  $x$ , we can rank the entries of  $x$  from the smallest to the largest, with ties broken by the rank of the state. We denote by  $r(x, s)$ , the rank of  $x_s$  in  $x$ . For example, if there are five states and  $x = (1, 4, 4, 3, 5)$ , we have  $r(x, 1) = 1$ ,  $r(x, 2) = 3$ ,  $r(x, 3) = 4$ ,  $r(x, 4) = 2$ , and  $r(x, 5) = 5$ . A rank dependent expected utility function gives to the bundle  $x$  the utility  $V(x, \pi) = \sum_{s=1}^{\bar{s}} \rho(x, s, \pi) u(x_s)$  where  $u : \mathbb{R}_+ \rightarrow \mathbb{R}$  is a Bernoulli function,

$$\rho(x, s, \pi) = g \left( \sum_{\{s' : r(x, s') \leq r(x, s)\}} \pi_{s'} \right) - g \left( \sum_{\{s' : r(x, s') < r(x, s)\}} \pi_{s'} \right), \quad (14)$$

and  $g : [0, 1] \rightarrow \mathbb{R}$  is a continuous and strictly increasing function. (If  $\{s' : r(x, s') < r(x, s)\}$  is empty, we let  $g \left( \sum_{\{s' : r(x, s') < r(x, s)\}} \pi_{s'} \right) = g(0)$ .) If  $g$  is the identity function (or, more generally when  $g$  is affine), we simply recover the expected utility model. When it is nonlinear, the function  $g$  distorts the cumulative distribution of the lottery  $x$ , 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,  $\rho(x, s, \pi) = \rho(\mathbf{u}(x), s, \pi)$ . It follows that we can write  $V(x, \pi) = \phi(\mathbf{u}(x), \pi)$ , where for any vector  $\mathbf{u} = (u_1, u_2, \dots, u_{\bar{s}})$ ,

$$\phi(\mathbf{u}, \pi) = \sum_{s=1}^{\bar{s}} \rho(\mathbf{u}, s, \pi) u_s. \quad (15)$$

The function  $\phi$  is continuous and strictly increasing in  $\mathbf{u}$ .

Suppose we wish to check whether  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is RDU-rationalizable with probability  $\pi = (\pi_1, \pi_2, \dots, \pi_{\bar{s}}) \gg 0$ ;<sup>13</sup> by this we mean that there is  $\phi$ , in the collection  $\Phi_{RDU}$

---

<sup>13</sup> To keep the notation light, we confine ourselves to the case where  $\pi$  does not vary across observations. There is no conceptual difficulty in allowing for this.



of functions of the form (15), and a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that, for each  $t$ ,  $\phi(\mathbf{u}(x^t), \pi^t) \geq \phi(\mathbf{u}(x), \pi^t)$  for all  $x \in B^t$ . To develop a necessary and sufficient test for this property, we first define

$$\Gamma = \left\{ \sum_{s \in S} \pi_s : S \subseteq \{1, 2, \dots, \bar{s}\} \right\},$$

i.e.,  $\Gamma$  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).

If  $\mathcal{O}$  is RDU-rationalizable, there must be increasing functions  $\bar{g} : \Gamma \rightarrow \mathbb{R}$  and  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  such that

$$\sum_{s=1}^{\bar{s}} \bar{\rho}(x^t, s, \pi^t) \bar{u}(x_s^t) \geq \sum_{s=1}^{\bar{s}} \bar{\rho}(x, s, \pi^t) u(x_s) \text{ for all } x \in \underline{B}^t \cap \mathcal{L} \text{ and} \quad (16)$$

$$\sum_{s=1}^{\bar{s}} \bar{\rho}(x^t, s, \pi^t) \bar{u}(x_s^t) > \sum_{s=1}^{\bar{s}} \bar{\rho}(x, s, \pi^t) \bar{u}(x_s) \text{ for all } x \in (\underline{B}^t \setminus \partial \underline{B}^t) \cap \mathcal{L}, \quad (17)$$

where

$$\bar{\rho}(x, s, \pi) = \bar{g} \left( \sum_{\{s': r(x, s') \leq r(x, s)\}} \pi_{s'} \right) - \bar{g} \left( \sum_{\{s': r(x, s') < r(x, s)\}} \pi_{s'} \right). \quad (18)$$

This is clear since we can simply take  $\bar{g}$  and  $\bar{u}$  to be the restriction of  $g$  and  $u$  respectively. Conversely, suppose there are strictly increasing functions  $\bar{g} : \Gamma \rightarrow \mathbb{R}$  and  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  such that (16), (17), and (18) are satisfied, and let  $g : [0, 1] \rightarrow \mathbb{R}$  be *any* continuous and strictly increasing extension of  $\bar{g}$ . With this  $g$ , and defining  $\phi(\mathbf{u}, \pi)$  by (15), the properties (16) and (17) may be re-written as

$$\phi(\mathbf{u}(x^t), \pi^t) \geq \phi(\mathbf{u}(x), \pi^t) \text{ for all } x \in \underline{B}^t \cap \mathcal{L} \text{ and}$$

$$\phi(\mathbf{u}(x^t), \pi^t) > \phi(\mathbf{u}(x), \pi^t) \text{ for all } x \in (\underline{B}^t \setminus \partial \underline{B}^t) \cap \mathcal{L}.$$

By Theorem 1, these properties guarantee that there exists a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  that extends  $\bar{u}$  such that the data set  $\mathcal{O}$  is RDU-rationalizable.

To recap, we have shown that  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is RDU-rationalizable with probability  $\pi \gg 0$  if and only if there are strictly increasing functions  $\bar{g} : \Gamma \rightarrow \mathbb{R}$  and  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  that satisfy (16), (17), and (18). As in the test for SEU-rationalizability, this test involves solving a set of bilinear inequalities; in this case, they are bilinear in  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  and  $\{\bar{g}(\gamma)\}_{\gamma \in \Gamma}$ .<sup>14</sup>

<sup>14</sup> It is also straightforward to modify the test to include restrictions on the shape of  $g$ . For example, to

In Section 5, we implement a test of Gul's (1991) model of disappointment aversion (DA). When there are just two states, disappointment aversion is a special case of rank dependent utility. Given a bundle  $x = (x_1, x_2)$ , let  $H$  denote the state with the higher payoff and  $\pi_H$  its objective probability. Under disappointment aversion, the probability of the higher outcome is distorted and becomes

$$\gamma(\pi_H) = \frac{\pi_H}{1 + (1 - \pi_H)\beta}, \quad (19)$$

where  $\beta \in (-1, \infty)$ . If  $\beta > 0$ , then  $\gamma(\pi_H) < \pi_H$  and the agent is said to be *disappointment averse*; if  $\beta < 0$ , then  $\gamma(\pi_H) > \pi_H$  and the agent is *elation seeking*; lastly, if  $\beta = 0$ , then there is no distortion and the agent simply maximizes expected utility. For a bundle  $(x_1, x_2) \in \mathbb{R}_+^2$ ,

$$\phi((u(x_1), u(x_2)), (\pi_1, \pi_2)) = (1 - \gamma(\pi_2))u(x_1) + \gamma(\pi_2)u(x_2) \text{ if } x_1 \leq x_2 \text{ and} \quad (20)$$

$$\phi((u(x_1), u(x_2)), (\pi_1, \pi_2)) = \gamma(\pi_1)u(x_1) + (1 - \gamma(\pi_1))u(x_2) \text{ if } x_2 < x_1. \quad (21)$$

When the agent is elation seeking,  $\phi$  is *not* concave in  $\mathbf{u} = (u_1, u_2)$ , so his preference over contingent consumption bundles need not be convex, even if  $u$  is concave. A data set is DA-rationalizable if and only if we can find  $\beta \in (-1, \infty)$  and a strictly increasing function  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  such that (16) and (17) are satisfied. Notice that, conditioning on the value of  $\beta$  (or, equivalently,  $\gamma(\pi_H)$ ), this test is linear in the variables  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  and so can be easily implemented. We use this feature when we test for disappointment aversion in Section 5.

### 3.2 Choice acclimating personal equilibrium (CPE)

The choice acclimating personal equilibrium model of Koszegi and Rabin (2007) (with a piecewise linear gain-loss function) specifies the agent's utility as  $V(x) = \phi(\mathbf{u}(x), \pi)$ , where

$$\phi((u_1, u_2, \dots, u_{\bar{s}}), \pi) = \sum_{s=1}^{\bar{s}} \pi_s u_s + \frac{1}{2}(1 - \lambda) \sum_{r,s=1}^{\bar{s}} \pi_r \pi_s |u_r - u_s|, \quad (22)$$

test that  $\mathcal{O}$  is RDU-rationalizable with a convex function  $g$  we need to specify that  $\bar{g}$  obeys

$$\frac{\bar{g}(\gamma^j) - \bar{g}(\gamma^{j-1})}{\gamma^j - \gamma^{j-1}} \leq \frac{\bar{g}(\gamma^{j+1}) - \bar{g}(\gamma^j)}{\gamma^{j+1} - \gamma^j} \text{ for } j = 1, \dots, \bar{m} - 1.$$

It is clear that this condition is necessary for the convexity of  $g$ . It is also sufficient for the extension of  $\bar{g}$  to a convex and increasing function  $g : [0, 1] \rightarrow \mathbb{R}$ . Thus  $\mathcal{O}$  is RDU-rationalizable with probability weights  $\{\pi^t\}_{t=1}^T$  and a convex function  $g$  if and only if there exist real numbers  $\{\bar{g}(\gamma)\}_{\gamma \in \Gamma}$  and  $\{\bar{u}(r)\}_{r \in \mathcal{X}}$  that satisfy the above condition in addition to (10), (16), (17), and (18).

$\pi = \{\pi_s\}_{s=1}^{\bar{s}}$  are the objective probabilities, and  $\lambda \in [0, 2]$  is the coefficient of loss aversion.<sup>15</sup> We say that a data set  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is CPE-rationalizable with probability  $\pi = (\pi_1, \pi_2, \dots, \pi_{\bar{s}}) \gg 0$  if there is  $\phi$  in the collection  $\Phi_{DA}$  of functions of the form (22), and a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that, for each  $t$ ,  $\phi(\mathbf{u}(x^t), \pi^t) \geq \phi(\mathbf{u}(x), \pi^t)$  for all  $x \in B^t$ . Applying Theorem 1,  $\mathcal{O}$  is CPE-rationalizable if and only if there is  $\lambda \in [0, 2]$  and a strictly increasing function  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  that solve (8) and (9). It is notable that, irrespective of the number of states, this test is linear in the remaining variables for any given value of  $\lambda$ . Thus it is relatively straightforward to implement via a collection of linear tests (running over different values of  $\lambda \in [0, 2]$ ).

### 3.3 Maxmin expected utility (MEU)

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

$$\phi(\mathbf{u}) = \min_{\pi \in \Pi} \left\{ \sum_{s=1}^{\bar{s}} \pi_s u_s \right\}, \quad (23)$$

where  $\Pi \subset \Delta_{++} = \{\pi \in \mathbb{R}_{++}^{\bar{s}} : \sum_{s=1}^{\bar{s}} \pi_s = 1\}$  is nonempty, compact in  $\mathbb{R}^{\bar{s}}$ , and convex. ( $\Pi$  can be interpreted as a set of probability weights.) Given these restrictions on  $\Pi$ , the minimization problem in (23) always has a solution and  $\phi$  is strictly increasing.

A data set  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is said to be MEU-rationalizable if there is a function  $\phi$  in the collection  $\Phi_{MEU}$  of functions of the form (23), and a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that, for each  $t$ ,  $\phi(\mathbf{u}(x^t), \pi^t) \geq \phi(\mathbf{u}(x), \pi^t)$  for all  $x \in B^t$ . By Theorem 1, this holds if and only if there exist  $\Pi$  and  $\bar{u}$  that solve (8), (9), and (10). 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  $\pi_1^*$  and  $\pi_1^{**} \in (0, 1)$  such that  $\Pi = \{(\pi_1, 1 - \pi_1) : \pi_1^* \leq \pi_1 \leq \pi_1^{**}\}$ . Then it is clear that  $\phi(u_1, u_2) = \pi_1^* u_1 + (1 - \pi_1^*) u_2$  if  $u_1 \geq u_2$  and  $\phi(u_1, u_2) = \pi_1^{**} u_1 + (1 - \pi_1^{**}) u_2$  if

<sup>15</sup> Our presentation of CPE follows Masatlioglu and Raymond (2014). The restriction of  $\lambda$  to  $[0, 2]$  guarantees that  $V$  respects first order stochastic dominance but allows for loss-loving behavior (see Masatlioglu and Raymond (2014)).

$u_1 < u_2$ . 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  $x_1 \geq x_2$  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  $\pi_1^*$  and  $\pi_1^{**}$  in  $(0, 1)$ , with  $\pi_1^* \leq \pi_1^{**}$ , and an increasing function  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  that solve (8) and (9). 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  $\pi_1^*$  and  $\pi_1^{**}$ .

The result below (which we prove in the Appendix) covers the case with multiple states. Note that the test involves solving a system of bilinear inequalities in the variables  $\bar{\pi}_s(x)$  (for all  $s$  and  $x \in \mathcal{L}$ ) and  $\bar{u}(r)$  (for all  $r \in \mathcal{X}$ ).

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

$$\bar{\pi}(x^t) \cdot \bar{\mathbf{u}}(x^t) \geq \bar{\pi}(x) \cdot \bar{\mathbf{u}}(x) \text{ for all } x \in \mathcal{L} \cap \underline{B}^t, \quad (24)$$

$$\bar{\pi}(x^t) \cdot \bar{\mathbf{u}}(x^t) > \bar{\pi}(x) \cdot \bar{\mathbf{u}}(x) \text{ for all } x \in \mathcal{L} \cap (\underline{B}^t \setminus \partial \underline{B}^t), \text{ and} \quad (25)$$

$$\bar{\pi}(x) \cdot \bar{\mathbf{u}}(x) \leq \bar{\pi}(x') \cdot \bar{\mathbf{u}}(x) \text{ for all } (x, x') \in \mathcal{L} \times \mathcal{L}. \quad (26)$$

If these conditions hold,  $\mathcal{O}$  admits an MEU-rationalization where  $\Pi$  (in (23)) is the convex hull of  $\{\bar{\pi}(x)\}_{x \in \mathcal{L}}$  and  $V(x) = \min_{\pi \in \Pi} \{\pi \cdot \bar{\mathbf{u}}(x)\} = \bar{\pi}(x) \cdot \bar{\mathbf{u}}(x)$  for all  $x \in \mathcal{L}$ .

### 3.4 Variational preferences (VP)

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

$$\phi(\mathbf{u}) = \min_{\pi \in \Delta_{++}} \{\pi \cdot \mathbf{u} + c(\pi)\} \quad (27)$$

and  $c : \Delta_{++} \rightarrow \mathbb{R}_+$  is a continuous and convex function with the following boundary condition: for any sequence  $\pi^n \in \Delta_{++}$  tending to  $\tilde{\pi}$ , with  $\tilde{\pi}_s = 0$  for some  $s$ , we obtain  $c(\pi^n) \rightarrow \infty$ . This boundary condition, together with the continuity of  $c$ , guarantee that there is  $\pi^* \in \Delta_{++}$  that solves the minimization problem in (27).<sup>16</sup> Therefore,  $\phi$  is well-defined and strictly increasing.

<sup>16</sup> Indeed, pick any  $\tilde{\pi} \in \Delta_{++}$  and define  $S = \{\pi \in \Delta_{++} : \pi \cdot \mathbf{u} + c(\pi) \leq \tilde{\pi} \cdot \mathbf{u} + c(\tilde{\pi})\}$ . The boundary condition and continuity of  $c$  guarantee that  $S$  is compact in  $\mathbb{R}^{\bar{s}}$  and hence  $\arg \min_{\pi \in S} \{\pi \cdot \mathbf{u} + c(\pi)\} = \arg \min_{\pi \in \Delta_{++}} \{\pi \cdot \mathbf{u} + c(\pi)\}$  is nonempty.

We say that  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is VP-rationalizable if there is a function  $\phi$  in the collection  $\Phi_{VP}$  of functions of the form (27), and a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that, for each  $t$ ,  $\phi(\mathbf{u}(x^t), \pi^t) \geq \phi(\mathbf{u}(x), \pi^t)$  for all  $x \in B^t$ . By Theorem 1, this holds if and only if there exists a function  $c : \Delta_{++} \rightarrow \mathbb{R}_+$  that is continuous, convex, and has the boundary property, and an increasing function  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  that together solve (8) and (9), with  $\phi$  defined by (27). The following result (proved in the Appendix) is a reformulation of this characterization that has a similar flavor to Proposition 1; note that, once again, the necessary and sufficient conditions on  $\mathcal{O}$  are expressed as a set of bilinear inequalities.

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

$$\bar{\pi}(x^t) \cdot \bar{\mathbf{u}}(x^t) + \bar{c}(x^t) \geq \bar{\pi}(x) \cdot \bar{\mathbf{u}}(x) + \bar{c}(x) \text{ for all } x \in \mathcal{L} \cap \underline{B}^t, \quad (28)$$

$$\bar{\pi}(x^t) \cdot \bar{\mathbf{u}}(x^t) + \bar{c}(x^t) > \bar{\pi}(x) \cdot \bar{\mathbf{u}}(x) + \bar{c}(x) \text{ for all } x \in \mathcal{L} \cap (\underline{B}^t \setminus \partial \underline{B}^t), \text{ and} \quad (29)$$

$$\bar{\pi}(x) \cdot \bar{\mathbf{u}}(x) + \bar{c}(x) \leq \bar{\pi}(x') \cdot \bar{\mathbf{u}}(x) + \bar{c}(x') \text{ for all } (x, x') \in \mathcal{L} \times \mathcal{L}. \quad (30)$$

*If these conditions hold, then  $\mathcal{O}$  can be rationalized by a variational preference  $V$  such that  $V(x) = \bar{\pi}(x) \cdot \bar{\mathbf{u}}(x) + \bar{c}(x)$  for all  $x \in \mathcal{L}$ , with  $c$  obeying  $c(\bar{\pi}(x)) = \bar{c}(x)$  for all  $x \in \mathcal{L}$ .*

### 3.5 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  $x \in \mathbb{R}_+^{\bar{s}}$ . The subject is observed to choose  $x^t$  from  $B^t \subset \mathbb{R}_+^{\bar{s}}$ , so the data set is  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$ . 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  $\omega^t \in \mathbb{R}_+^{\bar{s}}$  and a classical budget set  $B^t = \{x \in \mathbb{R}_+^{\bar{s}} : p^t \cdot x \leq p^t \cdot \omega^t\}$ . 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  $e^t \in \mathbb{R}_+^{\bar{s}}$ , and that the subject chooses according to some utility function  $V : [-K, \infty)^{\bar{s}} \rightarrow \mathbb{R}_+$  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

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

Let  $\{\phi(\cdot, t)\}_{t=1}^T$  be a collection of functions, where  $\phi(\cdot, t) : [-K, \infty)^{\bar{s}} \rightarrow \mathbb{R}$  is increasing in all of its arguments. We say that  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  is *rationalizable by  $\{\phi(\cdot, t)\}_{t=1}^T$  and the reference points  $\{e^t\}_{t=1}^T$*  if there exists a Bernoulli function  $u : [-K, \infty) \rightarrow \mathbb{R}_+$  such that  $\phi(\mathbf{u}(x^t - e^t), t) \geq \phi(\mathbf{u}(x - e^t), t)$  for all  $x \in B^t$ . This is formally equivalent to saying that the modified data set  $\mathcal{O}' = \{(x^t - e^t, \tilde{B}^t)\}_{t=1}^T$  is rationalizable by  $\{\phi(\cdot, t)\}_{t=1}^T$ . Applying Theorem 1, rationalizability holds if and only if there is a strictly increasing function  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  that obeys (8) and (9), 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  $u : [-K, \infty) \rightarrow \mathbb{R}$  that is S-shaped (and hence nonconcave) around 0 and probabilities distorted by a weighting function.

#### 4. GOODNESS OF FIT

The revealed preference tests that we have presented in the previous two sections are ‘sharp’, in the sense that a data set either passes the test for a particular model or it fails. This either/or feature of the tests is not peculiar 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. We now give an account of the most common approach developed in the literature to address this issue (see, for example, Afriat (1972, 1973), Varian (1990), and Halevy, Persitz, and Zrill (2014)) and explain why implementing the same approach in our setting is possible (or at least no more difficult than implementing the exact test).

Suppose that the observer collects a data set  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$ ; following the earlier papers, we focus attention on the case where  $B^t$  is a classical linear budget set given by (3). With no loss of generality, we normalize the price vector  $p^t$  so that  $p^t \cdot x^t = 1$ . Given a number  $e^t \in [0, 1]$  we define

$$B^t(e^t) = \{x \in \mathbb{R}_+^{\bar{s}} : x \leq x^t\} \cup \{x \in \mathbb{R}_+^{\bar{s}} : p^t \cdot x \leq e^t\}. \quad (31)$$

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 continuous and strictly increasing utility functions. We define the set  $E(\mathcal{U})$  in the following manner: a vector  $\mathbf{e} = (e^1, e^2, \dots, e^T)$  is in  $E(\mathcal{U})$  if there is some function  $U \in \mathcal{U}$  that rationalizes the modified data set  $\mathcal{O}(\mathbf{e}) = \{(x^t, B^t(e^t))\}_{t=1}^T$ , i.e.,  $U(x^t) \geq U(x)$  for all  $x \in B^t(e^t)$ . Clearly, the data set  $\mathcal{O}$  is rationalizable by a utility function in  $\mathcal{U}$  if and only if the unit vector  $(1, 1, \dots, 1)$  is in  $E(\mathcal{U})$ . We also know that  $E(\mathcal{U})$  must be nonempty since it contains the vector 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  $D_A(\mathbf{e}) = 1 - \min_{1 \leq t \leq T} \{e^t\}$ , 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 more straightforward. Note that  $D_A(\mathbf{e}) = D_A(\tilde{\mathbf{e}})$  where  $\tilde{e}^t = \min \{e^1, e^2, \dots, e^T\}$  for all  $t$  and, since

$\tilde{\mathbf{e}} \leq \mathbf{e}$ , we obtain  $\tilde{\mathbf{e}} \in E(\mathcal{U})$  whenever  $\mathbf{e} \in E(\mathcal{U})$ . Therefore,

$$\min_{\mathbf{e} \in E(\mathcal{U})} D_A(\mathbf{e}) = \min_{\tilde{\mathbf{e}} \in \tilde{E}(\mathcal{U})} D_A(\tilde{\mathbf{e}}),$$

where  $\tilde{E}(\mathcal{U}) = \{\mathbf{e} \in E(\mathcal{U}) : e^t = e^1 \forall t\}$ , i.e., in searching for  $\mathbf{e} \in E(\mathcal{U})$  that minimizes the supnorm distance from  $(1, 1, \dots, 1)$ , we can focus our attention on those vectors in  $E(\mathcal{U})$  that shrink each observed budget set by the same proportion. Given a data set  $\mathcal{O}$ , Afriat refers to  $\sup\{e : (e, e, \dots, e) \in E(\mathcal{U})\}$  as the *critical cost efficiency index* (CCEI); we say that  $\mathcal{O}$  is *rationalizable in  $\mathcal{U}$  at the efficiency index/threshold  $e'$*  if  $(e', e', \dots, e') \in E(\mathcal{U})$ .

Calculating the CCEI (or an index based on the Euclidean metric or some other metric) will require checking whether a particular vector  $\mathbf{e} = (e^1, e^2, \dots, e^T)$  is in  $E(\mathcal{U})$ , i.e., whether  $\{(x^t, B^t(e^t))\}_{t=1}^T$  is rationalizable by a member of  $\mathcal{U}$ . In the case where  $\mathcal{U}$  is the family of *all* increasing and continuous utility functions, it is known that a modified version of GARP (that excludes strict revealed preference cycles based on the modified budget sets  $B^t(e^t)$ ) is both a necessary and sufficient condition for the rationalizability of  $\{(x^t, B^t(e^t))\}_{t=1}^T$  (see Afriat (1972, 1973)).<sup>17</sup>

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

To illustrate how CCEI is calculated, consider Figure 3 which depicts a data set with two observations, where the chosen bundles are  $(2, 3)$  and  $(4, 0)$  and the observed prices imply budget frontiers given by the dashed lines. Since each chosen bundle is strictly within the observed budget set at the *other* observation, the data violate GARP and cannot be rationalized by a nonsatiated utility function. If we shrink both budget sets by some factor (as depicted by the solid lines), then eventually  $(2, 3)$  is no longer contained in the shrunken

---

<sup>17</sup> Alternatively, consult Forges and Minelli (2009) for a generalization of Afriat's Theorem to nonlinear budget sets; the test developed by Forges and Minelli can be applied to  $\{(x^t, B^t(e^t))\}_{t=1}^T$ .



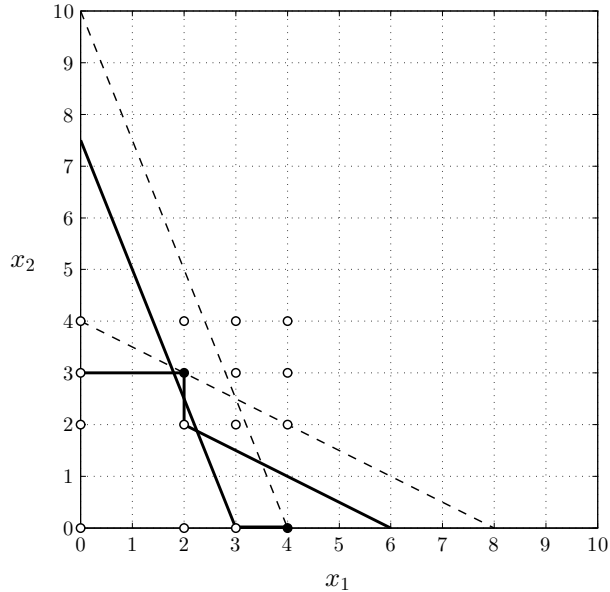


Figure 3: Establishing the CCEI

budget set containing  $(4, 0)$ ; at this efficiency threshold, the data set is rationalizable by some locally nonsatiated utility function (see Forges and Minelli (2009)). Whether the data set also passes a more stringent requirement such as EU-rationalizability at this efficiency threshold can be checked via the lattice procedure, performed on the finite lattice indicated in the figure.

#### 4.1 Approximate smooth rationalizability

While Theorem 1 guarantees that there is a Bernoulli function  $u$  that extends  $\bar{u} : \mathcal{X} \rightarrow \mathbb{R}_+$  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  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  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  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  rationalizes  $\mathcal{O} = \{(x^t, B^t)\}_{t=1}^T$  by  $\{\phi(\cdot, t)\}_{t=1}^T$ , then for any efficiency threshold  $e \in (0, 1)$ , there is a smooth Bernoulli function  $\tilde{u} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  that rationalizes  $\mathcal{O}' = \{(x^t, B^t(e))\}_{t=1}^T$  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.

## 5. IMPLEMENTATION

We implement our tests using data from the portfolio choice experiment in Choi *et al.* (2007), which was performed on 93 undergraduate subjects at the University of California, Berkeley. Every subject was asked to make consumption choices across 50 decision problems under risk. To be specific, he or she was asked to divide a 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 two asymmetric treatments applied to 17 and 29 subjects, the probability of the first state was  $1/3$  and  $2/3$ , respectively. These probabilities were objectively known. Income was normalized to one, and state prices were chosen at random and varied across subjects. Choi *et al.* (2007) analyzed the data by first implementing GARP tests on the observations for every subject; those subjects who passed, or came very close to passing (and were therefore consistent with the maximization of some strictly increasing utility function) were then fitted individually to a two-parameter version of the disappointment aversion model of Gul (1991).

### 5.1 GARP, FGARP, DA, and EU tests

Our principal objective is to assess how the data collected by Choi *et al.* (2007) perform against the nonparametric tests for DA- and EU-rationalizability that we have developed in this paper. We can check for EU-rationalizability using the test described in Section 2, which simply involves ascertaining whether or not there is a solution to a set of linear inequalities. As for DA-rationalizability, recall that the disappointment aversion model is a special case of the rank dependent utility model when there are two states; as we explained in Section 3.1, the test for DA-rationalizability is simply a linear test after conditioning on  $\beta$  (and hence the distorted probability of the favorable state,  $\gamma(\pi_H)$  (see (19))). We implement this test by letting  $\gamma(\pi_H)$  take up to 99 different values on  $(0, 1)$  and then performing the corresponding linear test. For example, in the symmetric case,  $\gamma(\pi_H)$  take on values 0.01, 0.02,  $\dots$ , 0.98, 0.99. Disappointment aversion is captured by  $\gamma(\pi_H) < 1/2$  (so  $\beta > 0$ ), while elation seeking behavior is captured by  $\gamma(\pi_H) > 1/2$  (so  $\beta < 0$ ).

It would be natural, before testing whether a subject satisfies EU- or DA-rationalizability, to take a few a steps back and ask a more fundamental question: is the subject's observed choice behavior consistent with the maximization of *any* well-behaved utility function? This can be answered by checking if the observations obey the *generalized axiom of revealed preference* (GARP), which is necessary and sufficient for rationalizability by some continuous and strictly increasing utility function; when a subject passes this test, we say that his or her behavior is consistent with the *increasing utility model*. The GARP test was carried out by Choi *et al.* (2007) and we repeat it here for all subjects. Given that we are in a contingent consumption context, it would also be natural to ask if a subject's behavior is rationalizable by a continuous utility function  $U : \mathbb{R}_+^{\bar{s}} \rightarrow \mathbb{R}$  with the stronger property of *stochastic monotonicity* (SCM); by this we mean that  $U(x) > U(y)$  whenever the contingent consumption plan  $x$  first order stochastically dominates  $y$  (where stochastic dominance is calculated with respect to the probabilities attached to each state). A test for rationalizability by continuous utility functions with this stronger property was recently developed by Nishimura, Ok, and Quah (2015) and has hitherto never been implemented. A subject who passes this test is said to be consistent with the *SCM utility model*.

In the experiment by Choi *et al.* (2007), the consumption space is  $\mathbb{R}_+^2$ . When  $\pi_1 =$

$\pi_2 = 1/2$ , it is straightforward to check that a utility function  $U$  is stochastically monotone if and only if it is increasing in both dimensions and symmetric. When  $\pi_1 > \pi_2$ , then  $U$  is stochastically monotone if and only if  $U$  is increasing in both dimensions and  $U(a, b) > U(b, a)$  whenever  $a > b$ . Data sets rationalizable by a utility function with this stronger property can be characterized by a stronger version of GARP, which we shall call FGARP (where ‘F’ stands for first order stochastic dominance).

To describe FGARP, it is useful to first recall GARP. Let  $\mathcal{D} = \{x^t : t = 1, 2, \dots, T\}$ ; in other words,  $\mathcal{D}$  consists of those bundles that have been observed somewhere in the data set. We say that  $x^t$  is *directly revealed preferred* to  $x^{t'}$  (for  $x^t$  and  $x^{t'}$  in  $\mathcal{D}$ ) if  $p^t \cdot x^t \geq p^t \cdot x^{t'}$ , or in other words, if  $x^{t'} \in B^t$ , where  $B^t$  is given by (3). This defines a reflexive binary relation on  $\mathcal{D}$ ; we call the transitive closure of this relation the *revealed preference relation* and say that  $x^t$  is *revealed preferred* to  $x^{t'}$  if  $x^t$  is related to  $x^{t'}$  by the revealed preference relation. Lastly, the bundle  $x^t$  is said to be *directly revealed strictly preferred* to  $x^{t'}$  if  $p^t \cdot x^t > p^t \cdot x^{t'}$ , or in other words,  $x^{t'} \in B^t \setminus \partial B^t$ , where  $\partial B^t$  is the upper boundary of  $B^t$ . The data set  $\mathcal{O}$  obeys GARP if, whenever  $x^t$  is revealed preferred to  $x^{t'}$  for any two bundles  $x^t$  and  $x^{t'}$  in  $\mathcal{D}$ , then  $x^{t'}$  is not directly revealed strictly preferred to  $x^t$ .

FGARP is a stronger version of GARP. Its general formulation can be found in Nishimura, Ok, and Quah (2015); we shall confine our explanation of FGARP to the two special cases relevant to our implementation. First, suppose that  $\pi_1 = \pi_2$ . In this case, we say that  $x^t$  is *directly revealed preferred* to  $x^{t'}$  if  $x^{t'} \in B^t$  or  $\tilde{x}^{t'} \in B^t$ , where  $\tilde{x}_1^{t'} = x_2^{t'}$  and  $\tilde{x}_2^{t'} = x_1^{t'}$ . (Note that while we are using the same term, the concept here is weaker than the one in the formulation of (standard) GARP.) We say that  $x^t$  is *revealed preferred* to  $x^{t'}$  if they are related by transitive closure. We say that  $x^t$  is *directly revealed strictly preferred* to  $x^{t'}$  if  $x^{t'} \in B^t \setminus \partial B^t$  or  $\tilde{x}^{t'} \in B^t \setminus \partial B^t$ . A data set  $\mathcal{O}$  obeys FGARP if, for any  $x^t$  and  $x^{t'}$  in  $\mathcal{D}$  such that  $x^t$  is revealed preferred to  $x^{t'}$ , then  $x^{t'}$  is not directly revealed strictly preferred to  $x^t$ . It is straightforward to check that if a subject is maximizing some SCM utility function  $U$ , then  $U(x^t) \geq U(x^{t'})$  whenever  $x^t$  is revealed preferred to  $x^{t'}$ , and  $U(x^t) > U(x^{t'})$  whenever  $x^t$  is directly revealed strictly preferred to  $x^{t'}$ ; from this it follows immediately that FGARP is a necessary property on  $\mathcal{O}$ . It turns out that this property is also sufficient to guarantee the rationalizability of the data by some continuous and SCM utility function (see Nishimura,

Ok, and Quah (2015)).

In the case where  $\pi_1 > \pi_2$ , we say that  $x^t$  is *directly revealed preferred* to  $x^{t'}$  if (i)  $x^{t'} \in B^t$  or (ii)  $\tilde{x}^{t'} \in B^t$  and  $x_2^{t'} > x_1^{t'}$ . As usual, we say that  $x^t$  is *revealed preferred* to  $x^{t'}$  if they are related via the transitive closure. We say that  $x^t$  is *directly revealed strictly preferred* to  $x^{t'}$  if (i)  $x^{t'} \in B^t \setminus \partial B^t$  or (ii)  $\tilde{x}^{t'} \in B^t$  and  $x_2^{t'} > x_1^{t'}$ . The data set obeys FGARP if whenever  $x^t$  is revealed preferred to  $x^{t'}$ , then  $x^{t'}$  is not directly revealed strictly preferred to  $x^t$ . Once again, FGARP is necessary and sufficient for the rationalizability of the data set by some continuous and SCM utility function.

To illustrate how FGARP works and to highlight its difference from GARP, suppose that  $\pi_1 = \pi_2 = 1/2$  and consider a data set with just one observation, where the bundle (1, 2) is purchased at the prices (3, 4). Given that there is just one observation, this data set will obviously obey GARP. However, it is clearly not compatible with a subject maximizing an SCM utility function, since the agent is buying more of the more expensive good, even though both states are equiprobable. This singleton data set violates FGARP. First, (1, 2) is directly revealed preferred to itself. On the other hand,  $(3, 4) \cdot (1, 2) > (3, 4) \cdot (2, 1)$  so (1, 2) is also directly revealed strictly preferred to itself.

In cases where a data set  $\mathcal{O}$  fails GARP (or FGARP), we may wish to find the efficiency level at which it is rationalizable by a continuous and strictly increasing (stochastically monotone) utility function. This can be done by a simple modification of GARP (FGARP). For example, to check whether the data set is rationalizable at an efficiency threshold of 0.9, we need only modify the revealed preference and revealed strict preference relations by replacing  $B^t$  and its upper boundary  $\partial B^t$  with  $B^t(0.9)$  (as defined by (31)) and its upper boundary  $\partial(B^t(0.9))$ . A data set  $\mathcal{O}$  then obeys GARP (FGARP) if, whenever  $x^t$  is revealed preferred to  $x^{t'}$  for any two bundles  $x^t$  and  $x^{t'}$  in  $\mathcal{D}$ , then  $x^{t'}$  is not directly revealed strictly preferred to  $x^t$ .

## 5.2 Pass rates and goodness of fit

We first test all four models on the data and the results are displayed in Table 1. Across 50 decision problems, 16 out of 93 subjects obey GARP and are therefore consistent with the increasing utility model; subjects in the symmetric treatment perform distinctly better

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

Table 1: Pass rates

than those in the asymmetric treatment. Of the 16 who pass GARP, only 4 pass FGARP and so display behavior consistent with the SCM utility model. Hardly any subjects pass the EU test and, while the DA model is in principle more permissive than EU, it does not fare any better on these data.

Given that we observe 50 decisions for every subject, it may not be intuitively surprising that there are so many violations of GARP (let alone more stringent conditions). Following Choi *et al.* (2007), we next investigate the efficiency thresholds at which subjects pass the different tests. We first calculate the CCEI associated with the increasing utility model, for each of the 93 subjects. This distribution is depicted in Figure 4. (Note that this figure is a replication of Figure 4 in Choi *et al.* (2007).) We can see that more than 80% of subjects have a CCEI above 0.9, and more than 90% have a CCEI above 0.8. A first glance at these results suggests that the data are largely consistent with the increasing utility model.

To better understand whether a CCEI of 0.9 implies the relative success or failure of a model to explain a given data set, it is useful to postulate an alternative hypothesis of some other form of behavior against which a comparison can be made. 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 lines. The Bronars (1987) approach has become common practice in the revealed preference literature as a way of assessing the ‘power’ 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 of whom is choosing randomly uniformly from 50 budget lines that are selected in the same random fashion as in the experimental setting. The dashed gray line in Figure 4 corresponds to the CCEI distribution for our simulated subjects. The experimental and simulated distributions are starkly different. For example, while 80% of subjects have a CCEI of 0.9 or higher, the chance of a randomly drawn sample passing GARP at an efficiency threshold of 0.9 is

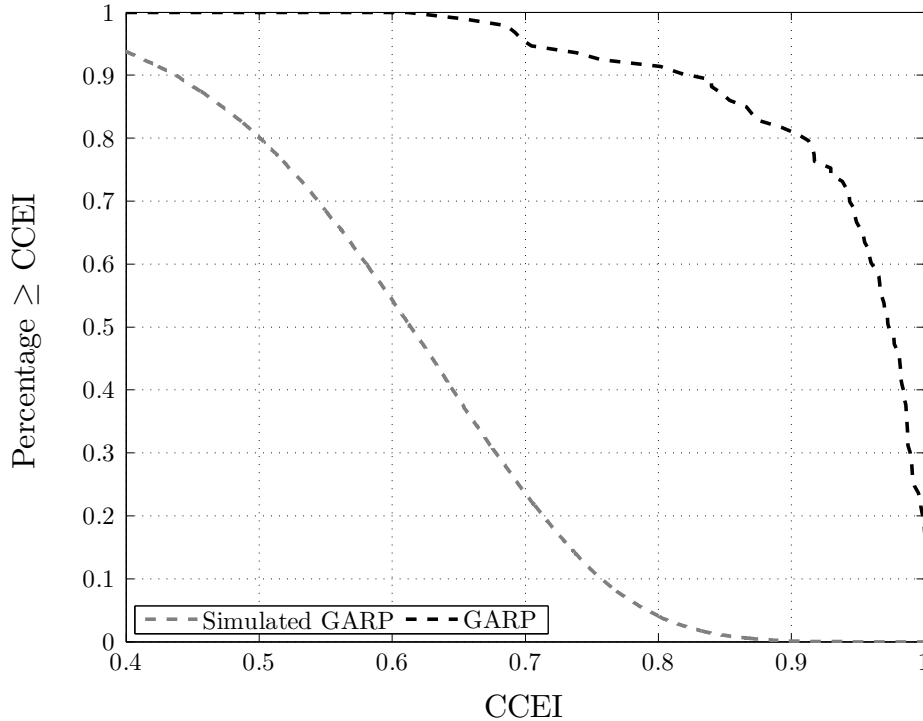
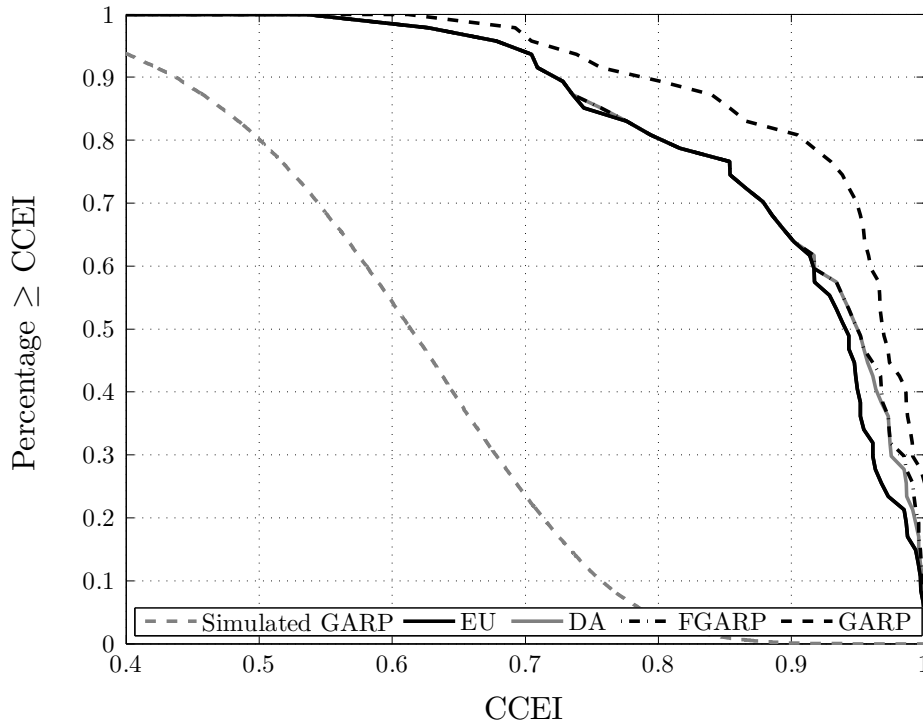


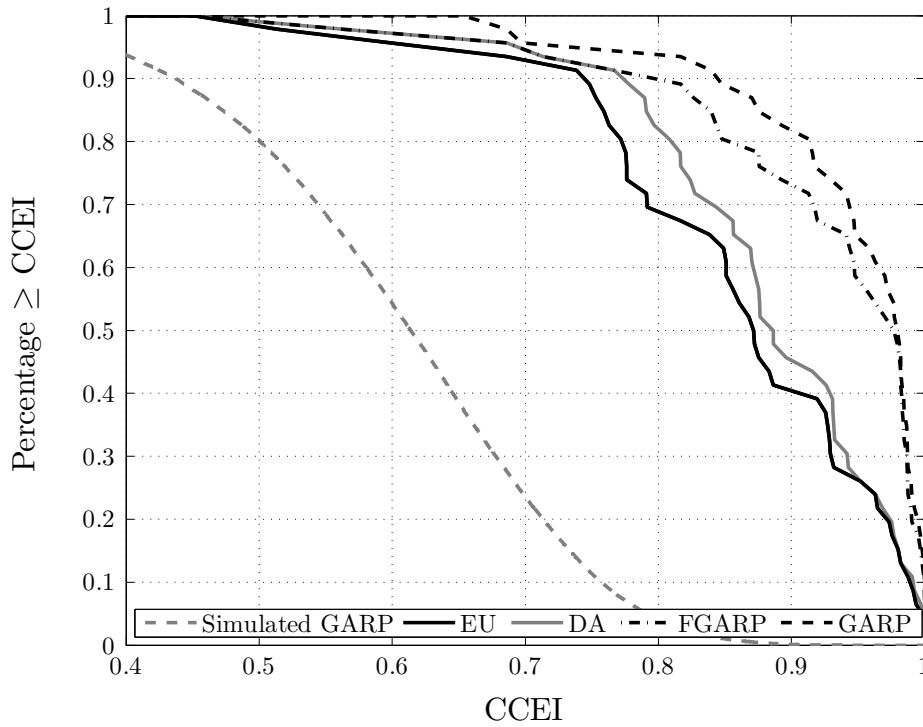
Figure 4: CCEI distribution for the increasing utility model

negligible, which lends support to increasing utility maximization as a model of choice among contingent consumption bundles.

Going beyond Choi *et al.* (2007), we then calculate CCEIs for the FGARP, DA, and EU tests 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 the increasing utility model, one would expect their CCEIs to be lower than for GARP, and they are. Nonetheless, at an efficiency index of 0.9, around half of all subjects are consistent with the EU model, with the proportion distinctly higher under the symmetric treatment. For the symmetric case, the performance for EU, DA, and FGARP are very close; in fact, the CCEI distributions for DA and FGARP are almost indistinguishable. For the asymmetric case, the performances of the different models are more distinct: FGARP does considerably better than EU or DA and its CCEI distribution is close to that for GARP. This is not altogether surprising and reflects the fact that FGARP is effectively a more stringent test in the symmetric case than in the asymmetric case (see Section 5.1). We have not depicted the CCEI distributions for the randomly generated



(a)  $\pi_1 = 1/2$



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

Figure 5: CCEI distributions for different models



data (under FGARP, DA, or EU), but plainly they will be even lower than that for GARP and therefore very different from the CCEI distributions for the experimental subjects. We conclude that a large proportion of the subjects behave in a way that is nearly consistent with the EU or DA models, a group which is too sizable to be dismissed as occurring naturally in random behavior.

### 5.3 Predictive success

While these results are highly suggestive, we would like a more formal way of comparing across different candidate models of behavior. The four models being compared are, in increasing order of stringency, increasing utility, SCM utility, disappointment aversion, and expected utility. What is needed in comparing these models is a way of trading off a model's frequency of correct predictions (which favors increasing utility) with the precision of its predictions (which favors expected utility). To do this, we make use of an axiomatic measure of predictive success proposed by Selten (1991). Selten's *index of predictive success* (which we shall refer to simply as the *Selten index*) is defined as the difference between the relative frequency of correct predictions (the 'hit rate') and the relative size of the set of predicted outcomes (the 'precision'). Our use of this index to evaluate different consumption models is not novel; see, in particular, Beatty and Crawford (2011).

To calculate the Selten index, we need the empirical frequency of correct predictions and the relative size of the set of predicted outcomes. To measure the latter, we use the frequency of hitting the set of predicted outcomes with uniform random draws. Specifically, for each subject, we generate 1,000 synthetic data sets containing consumption bundles chosen randomly uniformly from the actual budget sets facing that subject. (Recall that each subject in Choi *et al.* (2007) faced a different collection of 50 budget sets.) For a given efficiency threshold and for each model, we calculate the Selten index for every subject, which is either 1 (pass) or 0 (fail) minus the fraction of the 1,000 randomly simulated subject-specific data sets that pass the test (for that model).<sup>18</sup> The index ranges from  $-1$  to  $1$ , where  $-1$

---

<sup>18</sup> Each of the 1,000 synthetic data sets is subjected to a test of GARP, FGARP, DA, and EU. We test for GARP and FGARP using Warshall's algorithm and the EU test is linear, so these tests are computationally undemanding. The test for DA-rationalizability is more computationally intensive, since each test involves performing 99 linear tests (see the earlier discussion in this section).

		Efficiency Level				
		0.80	0.85	0.90	0.95	1.00
GARP	$\pi_1 = 1/2$	42/47 (89%) 0.83 [0.94]	40/47 (85%) 0.84 [0.99]	38/47 (81%) 0.81 [1.00]	32/47 (68%) 0.68 [1.00]	12/47 (26%) 0.26 [1.00]
	$\pi_1 \neq 1/2$	43/46 (93%) 0.88 [0.94]	40/46 (87%) 0.86 [0.99]	37/46 (80%) 0.80 [1.00]	29/46 (63%) 0.63 [1.00]	4/46 (9%) 0.09 [1.00]
	Total	85/93 (91%) <b>0.86</b> [0.94]	80/93 (86%) <b>0.85</b> [0.99]	75/93 (81%) <b>0.81</b> [1.00]	61/93 (66%) <b>0.66</b> [1.00]	16/93 (17%) <b>0.17</b> [1.00]
FGARP	$\pi_1 = 1/2$	37/47 (79%) 0.79 [1.00]	36/47 (77%) 0.77 [1.00]	30/47 (64%) 0.64 [1.00]	23/47 (49%) 0.49 [1.00]	1/47 (2%) 0.02 [1.00]
	$\pi_1 \neq 1/2$	41/46 (89%) 0.88 [0.99]	36/46 (78%) 0.78 [1.00]	33/46 (72%) 0.72 [1.00]	26/46 (57%) 0.57 [1.00]	3/46 (7%) 0.07 [1.00]
	Total	78/93 (84%) <b>0.83</b> [0.99]	72/93 (77%) <b>0.77</b> [1.00]	63/93 (68%) <b>0.68</b> [1.00]	49/93 (53%) <b>0.53</b> [1.00]	4/93 (4%) <b>0.04</b> [1.00]
DA	$\pi_1 = 1/2$	37/47 (79%) 0.78 [1.00]	36/47 (77%) 0.77 [1.00]	30/47 (64%) 0.64 [1.00]	23/47 (49%) 0.49 [1.00]	1/47 (2%) 0.02 [1.00]
	$\pi_1 \neq 1/2$	37/46 (80%) 0.80 [1.00]	31/46 (67%) 0.67 [1.00]	20/46 (43%) 0.43 [1.00]	12/46 (26%) 0.26 [1.00]	1/46 (2%) 0.02 [1.00]
	Total	74/93 (80%) <b>0.79</b> [1.00]	67/93 (72%) <b>0.72</b> [1.00]	50/93 (54%) <b>0.54</b> [1.00]	35/93 (38%) <b>0.38</b> [1.00]	2/93 (2%) <b>0.02</b> [1.00]
EU	$\pi_1 = 1/2$	37/47 (79%) 0.79 [1.00]	36/47 (77%) 0.77 [1.00]	30/47 (64%) 0.64 [1.00]	18/47 (38%) 0.38 [1.00]	1/47 (2%) 0.02 [1.00]
	$\pi_1 \neq 1/2$	31/46 (67%) 0.67 [1.00]	28/46 (61%) 0.61 [1.00]	18/46 (39%) 0.39 [1.00]	12/46 (26%) 0.26 [1.00]	1/46 (2%) 0.02 [1.00]
	Total	68/93 (73%) <b>0.73</b> [1.00]	64/93 (69%) <b>0.69</b> [1.00]	48/93 (52%) <b>0.52</b> [1.00]	30/93 (32%) <b>0.32</b> [1.00]	2/93 (2%) <b>0.02</b> [1.00]

Table 2: Pass rates, power, and predictive success

corresponds to failing a lenient test and 1 to passing a stringent test. Lastly, we take the arithmetic average of these indices across subjects in order to obtain an aggregate Selten index.<sup>19</sup> Equivalently, the Selten index is the difference between the empirical frequency of a correct prediction and the precision, which is the relative size of predicted outcomes (averaged across agents). Following Bronars (1987), we refer to 1 minus the average relative size of predicted outcomes as a model's *power*; a power index very close to 1 means that the model's predictions are highly specific.

The results are shown in Table 2. Within each cell in the table are four numbers, with

<sup>19</sup> This aggregation is supported by the axiomatization in Selten (1991). Note that our procedure for calculating the Selten index is essentially the same as that in Beatty and Crawford (2011).

the pass rate in the top row and, in the lower row, the Selten index and the power, with the latter in brackets. For example, with  $\pi_1 = 1/2$  and at an efficiency threshold of 0.8, 42 out of 47 subjects (89%) pass GARP, the power is 0.94, and so the Selten index is  $0.83 = 0.89 - (1 - 0.94)$ . What emerges immediately is that, at any efficiency level and for any model, the Selten indices of predictive success are almost completely determined by the pass rates. This is because all of the models have uniformly high power (in fact, very close to 1). It turns out that with 50 observations for each subject, even the increasing utility model has very high power. Given this, *the Selten index ranks the increasing utility model as the best, followed by SCM utility, disappointment aversion, and expected utility*; this ranking holds at any efficiency level and across both treatments. All four models have indices well within the positive range, indicating that they are clearly superior to the hypothesis of uniform random choice (which has a Selten index of 0). While academic discussion is often focussed on comparing different models that have been tailor-made for decisions under risk and uncertainty, these findings suggest that we should *not* take it for granted that such models are necessarily better than the standard increasing utility model. At least in the data analyzed here, one could argue that this model does a better job in explaining the data, even after accounting for its relative lack of specificity.

#### 5.4 Conditional predictive success

Our next objective is to investigate the success of the EU and DA models in explaining agent behavior, conditional on the agent maximizing some utility function. To do this, we first identify, at a given efficiency threshold, those subjects who pass GARP at that threshold. For each of these subjects, we then generate 1,000 synthetic data sets that obey GARP at the same efficiency threshold and are thus consistent with the increasing utility model.<sup>20</sup> (Note that since we focus on five different efficiency thresholds, this implies that we

---

<sup>20</sup> The procedure for creating a synthetic data set is as follows. We randomly select a budget line (out of the 50 budget lines) and then randomly choose a bundle on that line. We next randomly select a second budget line, and then randomly choose from that part of the line which guarantees that this observation, along with the first, obeys GARP (or, in the case where the efficiency index is lower than 1, a modified version of GARP). We then randomly select a third budget line, and choose a random bundle on that part of the line which guarantees that the three observations together obey GARP (or modified GARP). Note that there must exist such a bundle on the third budget line; indeed it can be chosen to be the demand (on the third budget line) of any utility function rationalizing the first two observations. We then select a fourth budget line, and so on.

		Efficiency Level				
		0.80	0.85	0.90	0.95	1.00
FGARP	$\pi_1 = 1/2$	37/42 (88%) 0.87 [0.99]	36/40 (90%) 0.90 [1.00]	30/38 (79%) 0.79 [1.00]	23/32 (72%) 0.72 [1.00]	1/12 (8%) 0.08 [1.00]
	$\pi_1 \neq 1/2$	41/43 (95%) 0.74 [0.79]	36/40 (90%) 0.73 [0.83]	33/37 (89%) 0.77 [0.88]	26/29 (90%) 0.82 [0.93]	3/4 (75%) 0.73 [0.98]
	Total	78/85 (92%) <b>0.81</b> [0.89]	72/80 (90%) <b>0.82</b> [0.92]	63/75 (84%) <b>0.78</b> [0.94]	49/61 (80%) <b>0.77</b> [0.97]	4/16 (25%) <b>0.25</b> [1.00]
DA	$\pi_1 = 1/2$	37/42 (88%) 0.87 [0.99]	36/40 (90%) 0.90 [1.00]	30/38 (79%) 0.79 [1.00]	23/32 (72%) 0.72 [1.00]	1/12 (8%) 0.08 [1.00]
	$\pi_1 \neq 1/2$	37/43 (86%) 0.85 [0.99]	31/40 (78%) 0.78 [1.00]	20/37 (54%) 0.54 [1.00]	12/29 (41%) 0.41 [1.00]	1/4 (25%) 0.25 [1.00]
	Total	74/85 (87%) <b>0.86</b> [0.99]	67/80 (84%) <b>0.84</b> [1.00]	50/75 (67%) <b>0.67</b> [1.00]	35/61 (57%) <b>0.57</b> [1.00]	2/16 (12%) <b>0.12</b> [1.00]
EU	$\pi_1 = 1/2$	37/42 (88%) 0.87 [0.99]	36/40 (90%) 0.90 [1.00]	30/38 (79%) 0.79 [1.00]	18/32 (56%) 0.56 [1.00]	1/12 (8%) 0.08 [1.00]
	$\pi_1 \neq 1/2$	31/43 (72%) 0.72 [1.00]	28/40 (70%) 0.70 [1.00]	18/37 (49%) 0.49 [1.00]	12/29 (41%) 0.41 [1.00]	1/4 (25%) 0.25 [1.00]
	Total	68/85 (80%) <b>0.80</b> [1.00]	64/80 (80%) <b>0.80</b> [1.00]	48/75 (64%) <b>0.64</b> [1.00]	30/61 (49%) <b>0.49</b> [1.00]	2/16 (12%) <b>0.12</b> [1.00]

Table 3: Pass rates, power, and predictive success (conditional on GARP)

have to generate a total of 5,000 synthetic data sets for every subject.) We then implement, on each of these data sets, the FGARP, DA, and EU tests (at the corresponding efficiency threshold). This gives us the precision or predictive power of these models *within the context of behavior that is consistent with the (basic) increasing utility model*. We then calculate the Selten index; in this case, it is the difference between the empirical frequency of passing the model (for example, the number who pass EU amongst those who pass GARP) and the precision, i.e., the relative size of predicted outcomes (averaged across subjects).

The results are displayed in Table 3, where the numbers in each cell have interpretations analogous to those in Table 2. For example, with  $\pi_1 = 1/2$  and an efficiency threshold of 0.9, 38 subjects pass GARP, of which 30 (79%) pass DA. Amongst the data sets that pass GARP at the 0.9 threshold, the power of the DA is (effectively) 1; in other words, the model has a precision close to zero. It follows that the Selten index is 0.79. (Notice that the number of subjects with the symmetric treatment passing GARP shrinks from 42 to 12 as the efficiency threshold is raised.)

		Efficiency Level				
		0.80	0.85	0.90	0.95	1.00
DA	$\pi_1 = 1/2$	37/37 (100%) 0.11 [0.11]	36/36 (100%) 0.29 [0.29]	30/30 (100%) 0.65 [0.65]	23/23 (100%) 0.97 [0.97]	1/1 (100%) 1.00 [1.00]
	$\pi_1 \neq 1/2$	37/41 (90%) 0.89 [0.99]	31/36 (86%) 0.86 [1.00]	20/33 (61%) 0.61 [1.00]	12/26 (46%) 0.46 [1.00]	1/3 (33%) 0.33 [1.00]
	Total	74/78 (95%) <b>0.52</b> [0.57]	67/72 (93%) <b>0.57</b> [0.64]	50/63 (79%) <b>0.63</b> [0.83]	35/49 (71%) <b>0.70</b> [0.99]	2/4 (50%) <b>0.50</b> [1.00]
EU	$\pi_1 = 1/2$	37/37 (100%) 0.23 [0.23]	36/36 (100%) 0.50 [0.50]	30/30 (100%) 0.81 [0.81]	18/23 (78%) 0.77 [0.99]	1/1 (100%) 1.00 [1.00]
	$\pi_1 \neq 1/2$	31/41 (76%) 0.76 [1.00]	28/36 (78%) 0.78 [1.00]	18/33 (55%) 0.55 [1.00]	12/26 (46%) 0.46 [1.00]	1/3 (33%) 0.33 [1.00]
	Total	68/78 (87%) <b>0.51</b> [0.63]	64/72 (89%) <b>0.64</b> [0.75]	48/63 (76%) <b>0.67</b> [0.91]	30/49 (61%) <b>0.61</b> [0.99]	2/4 (50%) <b>0.50</b> [1.00]

Table 4: Pass rates, power, and predictive success (conditional on FGARP)

Perhaps the first feature worth highlighting in this table is that the EU and DA models remain very precise, with power indices that are very close to 1. In other words, even if we are choosing randomly amongst data sets that obey GARP, the probability of choosing a data set obeying EU or DA is close to zero. Nonetheless, the DA and EU models *do* in fact capture a large proportion of those subjects who display utility-maximizing behavior. Since both DA and EU models are almost equally precise, the Selten index favors the DA model and this is true at all efficiency thresholds. Closer examination suggests that the main difference in performance between the two models arises from the asymmetric treatment, where DA is better at explaining the subjects' behavior.

Under the symmetric treatment, the performance of the SCM utility model is also very similar to that for EU and DA. But for the case of asymmetric treatment, there is an interesting departure: while the pass rate for FGARP is significantly higher than that for DA, the power is also significantly lower. This leads to the DA model having a higher Selten index at the 0.8 and 0.85 efficiency thresholds, while the SCM utility model performs better at higher efficiency levels.

An interesting question left largely unanswered by Table 3 is the precision of the DA or EU models, *relative to the SCM utility model*. Put another way, does the DA model impose significantly greater restrictions on the data, beyond that imposed by an SCM utility

function? To examine this issue systematically, we build a table similar to Table 3 except that now we confine our attention to those subjects who pass FGARP (rather than GARP) at a given threshold. For each of these subjects, we then generate 1,000 synthetic data sets that obey FGARP at that threshold and are thus consistent with the SCM utility model (using a procedure analogous to that used for generating random data sets obeying GARP). We then implement the EU and DA tests on these data sets. This gives us the precision of these models relative to the SCM utility model; based on this we may then calculate Selten's index of predictive success.

The results are displayed in Table 4, where the numbers in each cell have interpretations analogous to those in Table 3. Perhaps the first observation to make is that the EU and DA models are no longer almost perfectly precise for the symmetric case. For example, at an efficiency threshold of 0.9, the power of the DA model is 0.65; in other words, the probability of a data set obeying the DA model, when it is drawn from data sets obeying FGARP is roughly 0.35. The EU model is more precise but, even in that case, the power is clearly below 1. This suggests that the symmetry of the utility function (which is essentially what the SCM utility model requires in this case) sharply narrows the collection of permissible data sets, so much so that the relative power of the DA model is no longer close to 1. On the other hand, the DA model's relative precision remains close to 1 in the asymmetric case, where the FGARP test is significantly less stringent. The picture is mixed if we compare the Selten indices of the DA and EU models, with the DA model performing worse at the 0.9 threshold but better at 0.95. Whenever DA ranks below EU, it is driven by the relative imprecision of the DA model in the symmetric case. Significantly, even in the symmetric case, both models retain their predictive power, with positive Selten indices.

## 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  $R_+^{\bar{s}}$  that are compact and downward closed (i.e., if  $x \in C^t$  then so is  $y \in R_+^{\bar{s}}$  such that  $y < x$ ) 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  $R_+$ , a strictly increasing function  $\bar{u} : \mathcal{X} \rightarrow R_+$ , and  $\{M^t\}_{t \in T}$  such that the following*

holds:

$$M^t \geq \phi(\bar{\mathbf{u}}(x), t) \text{ for all } x \in C^t \cap \mathcal{L} \text{ and} \quad (32)$$

$$M^t > \phi(\bar{\mathbf{u}}(x), t) \text{ for all } x \in (C^t \setminus \partial C^t) \cap \mathcal{L}, \quad (33)$$

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

$$M^t \geq \phi(\mathbf{u}(x), t) \text{ for all } x \in C^t \text{ and} \quad (34)$$

$$\text{if } x \in C^t \text{ and } M^t = \phi(\mathbf{u}(x), t), \text{ then } x \in \partial C^t \cap \mathcal{L} \text{ and } M^t = \phi(\bar{\mathbf{u}}(x), t). \quad (35)$$

REMARK: The property (35) needs some explanation. Conditions (32) and (33) allow for the possibility that  $M^t = \phi(\bar{\mathbf{u}}(x'), t)$  for some  $x' \in \partial C^t \cap \mathcal{L}$ ; 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.,  $M^t = \phi(\mathbf{u}(x'), t)$  for all  $x' \in X'$ . Property (35) says that we can choose  $u$  such that for all  $x \in C^t \setminus X'$ , we have  $M^t > \phi(\bar{\mathbf{u}}(x), t)$ .

*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  $R_+$  and each  $C^t$  is a closed interval with 0 as its minimum. Now 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  $u^t : R_+ \rightarrow R_+$  extending  $\bar{u}$  such that (34) and (35) hold, then the the same conditions will hold for the increasing and continuous function  $u = \min_{t \in T} u^t$ . 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  $\bar{r} = \max\{r \in R_+ : (r, 0, 0, \dots, 0) \in C^t\}$  and suppose that  $(r^i, 0, 0, \dots, 0) \in C^t$  if and only if  $i \leq N$  (for some  $N \leq I$ ). Consider the collection of sets of the form  $D^i = \{y \in R_+^m : (r^i, y) \in C^t\}$  (for  $i = 1, 2, \dots, N$ ); this is a finite collection of compact and downward closed sets in  $R_+^m$ . 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  $u^* : R_+ \rightarrow R_+$  extending  $\bar{u}$  such that

$$M^t \geq \phi(\bar{u}(r^i), \mathbf{u}^*(y), t) \text{ for all } (r^i, y) \in C^t \text{ and} \quad (36)$$

if  $(r^i, y) \in C^t$  and  $M^t = \phi(\bar{u}(r^i), \mathbf{u}^*(y), t)$ , then  $(r^i, y) \in \partial C^t \cap \mathcal{L}$  and  $M^t = \phi(\bar{\mathbf{u}}(r^i, y), t)$ . (37)

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

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

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  $(r, y) \in C^t$  then so is  $(r^k, y)$ , and (36) guarantees that  $\phi(\bar{u}(r^k), \mathbf{u}^*(y), t) \leq M^t$ . 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  $\phi$  continuous,  $\phi(\tilde{u}(r), \mathbf{u}^*(y), t) \leq M^t$  for all  $(r, y) \in C^t$ . 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  $\lim_{n \rightarrow \infty} u_n = \tilde{u}(r)$ , there is  $(r, y_n) \in C^t$  such that  $\phi(u_n, \mathbf{u}^*(y_n), t) > M^t$ . Since  $C^t$  is compact, we may assume with no loss of generality that  $y_n \rightarrow \hat{y}$  and  $(r, \hat{y}) \in C^t$ , from which we obtain  $\phi(\tilde{u}(r), \mathbf{u}^*(\hat{y}), t) = M^t$ . Since  $C^t$  is downward closed,  $(r^k, \hat{y}) \in C^t$  and, since  $\bar{u}(r^k) = u^*(r^k)$ , we have  $\phi(\mathbf{u}^*(r^k), \hat{y}, t) = M^t$ . This can only occur if  $(r^k, \hat{y}) \in \partial C^t \cap \mathcal{L}$  (because of (37)), but it is clear that  $(r^k, \hat{y}) \notin \partial C^t$  since  $(r^k, \hat{y}) < (r, \hat{y})$ . (iv) If  $r_n < r^i$  for all  $n$  and  $r_n \rightarrow r^i \in \mathcal{X}$ , then  $\tilde{u}(r_n) \rightarrow u^*(r^i)$ . 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  $\phi$ , and the definition of  $\tilde{u}$ , there is  $(r_n, y_n) \in C^t$  such that  $\phi(\tilde{u}(r_n), \mathbf{u}^*(y_n), t) = M^t$ . This leads to  $\phi(\hat{u}, \mathbf{u}^*(y'), t) = M^t$ , where  $y'$  is an accumulation point of  $y_n$  and  $(r^i, y') \in C^t$ . But since  $\phi$  is strictly increasing, we obtain  $\phi(u^*(r^i), \mathbf{u}^*(y'), t) > M^t$ , which contradicts (36).

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  $r \in R_+ \setminus \mathcal{X}$  and  $u^t(r) < \tilde{u}(r) \leq u^*(r)$  for all  $r \in [0, \bar{r}] \setminus \mathcal{X}$ . (In fact we can choose  $u^t$  to be smooth everywhere except possibly on  $\mathcal{X}$ .) We claim that (34) and (35) are satisfied for  $C^t$ . To see this, note that for  $r \in \mathcal{X}$  and  $(r, y) \in C^t$ , the induction hypothesis guarantees that (36) and (37) hold and they will continue to hold if  $u^*$  is replaced by  $u^t$ . In the case where  $r \notin \mathcal{X}$  and  $(r, y) \in C^t$ ,



since  $u^t(r) < \tilde{u}(r)$  and  $\phi$  is increasing, we obtain  $M^t > \phi(\mathbf{u}^t(r, y), t)$ . QED

*Proof of Theorem 1:* This follows immediately from Lemma 1 if we set  $C^t = \underline{B}^t$ , and  $M^t = \phi(\bar{\mathbf{u}}(x^t), t)$ . If  $\bar{u}$  obeys conditions (8) and (9) then it obeys conditions (32) and (33). The rationalizability of  $\mathcal{O}$  by  $\{\phi(\cdot, t)\}_{t \in T}$  then follows from (34). QED

*Proof of Proposition 1:* Suppose that  $\mathcal{O}$  is rationalizable by  $\phi$  as defined by (23). For any  $x$  in the finite lattice  $\mathcal{L}$ , let  $\bar{\pi}(x)$  be an element in  $\arg \min_{\pi \in \Pi} \pi \cdot \mathbf{u}(x)$  and let  $\bar{u}$  be the restriction of  $u$  to  $\mathcal{X}$ . Then it is clear that the conditions (24) – (26) hold.

Conversely, suppose that there is a function  $\bar{\pi}$  and a strictly increasing function  $\bar{u}$  obeying the conditions (24) – (26). Define  $\Pi$  as the convex hull of  $\{\bar{\pi}(x) : x \in \mathcal{L}\}$ ;  $\Pi$  is a nonempty and convex subset of  $\Delta_{++}$  and it is compact in  $\mathbb{R}^{\bar{s}}$  since  $\mathcal{L}$  is finite. Suppose that there exists  $x \in \mathcal{L}$  and  $\pi \in \Pi$  such that  $\pi \cdot \bar{\mathbf{u}}(x) < \bar{\pi}(x) \cdot \bar{\mathbf{u}}(x)$ . Since  $\pi$  is a convex combination of elements in  $\{\bar{\pi}(x) : x \in \mathcal{L}\}$ , there must exist  $x' \in \mathcal{L}$  such that  $\bar{\pi}(x') \cdot \bar{\mathbf{u}}(x) < \bar{\pi}(x) \cdot \bar{\mathbf{u}}(x)$ , which contradicts (26). We conclude that  $\bar{\pi}(x) \cdot \bar{\mathbf{u}}(x) = \min_{\pi \in \Pi} \pi \cdot \bar{\mathbf{u}}(x)$  for all  $x \in \mathcal{L}$ . We define  $\phi : \mathbb{R}_+^{\bar{s}} \rightarrow \mathbb{R}$  by  $\phi(\mathbf{u}) = \min_{\pi \in \Pi} \pi \cdot \mathbf{u}$ . Then the conditions (24) and (25) are just versions of (8) and (9) and so Theorem 1 guarantees that there is Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  extending  $\bar{u}$  such that  $\mathcal{O}$  is rationalizable by  $V(x) = \phi(\mathbf{u}(x))$ . QED

*Proof of Proposition 2:* Suppose  $\mathcal{O}$  is rationalizable by  $\phi$  as defined by (27). Let  $\bar{u}$  be the restriction of  $u$  to  $\mathcal{X}$ . For any  $x$  in  $\mathcal{L}$ , let  $\bar{\pi}(x)$  be an element in  $\arg \min_{\pi \in \Delta_{++}} \{\pi \cdot \mathbf{u}(x) + c(\pi)\}$ , and let  $\bar{c}(x) = c(\bar{\pi}(x))$ . Then it is clear that the conditions (28) – (30) hold.

Conversely, suppose that there is a strictly increasing function  $\bar{u}$  and functions  $\bar{\pi}$  and  $\bar{c}$  obeying conditions (28) – (30). For every  $\pi \in \Delta_{++}$ , define  $\tilde{c}(\pi) = \max_{x \in \mathcal{L}} \{\bar{c}(x) - (\pi - \bar{\pi}(x)) \cdot \bar{\mathbf{u}}(x)\}$ . It follows from (30) that  $\bar{c}(x') \geq \bar{c}(x) - (\bar{\pi}(x') - \bar{\pi}(x)) \cdot \bar{\mathbf{u}}(x)$  for all  $x \in \mathcal{L}$ . Therefore,  $\tilde{c}(\bar{\pi}(x')) = \bar{c}(x')$  for any  $x' \in \mathcal{L}$ . 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  $\Delta_{++}$  that is convex, continuous, obeys the boundary condition, with  $c(\pi) \geq \tilde{c}(\pi)$  for all  $\pi \in \Delta_{++}$  and  $c(\pi) = \tilde{c}(\pi)$  for  $\pi \in \{\bar{\pi}(x) : x \in \mathcal{L}\}$ . We claim that, with  $c$  so defined,  $\min_{\pi \in \Delta_{++}} \{\pi \cdot \bar{\mathbf{u}}(x) + c(\pi)\} = \pi(x) \cdot \bar{\mathbf{u}}(x) + \bar{c}(x)$  for all  $x \in \mathcal{L}$ . Indeed, for any  $\pi \in \Delta_{++}$ ,

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

On the other hand,  $\bar{\pi}(x) \cdot \mathbf{u}(x) + c(\bar{\pi}(x)) = \bar{\pi}(x) \cdot \mathbf{u}(x) + \bar{c}(x)$ , which establishes the claim. We define  $\phi : \mathbb{R}_+^{\bar{s}} \rightarrow \mathbb{R}$  by (27); then (28) and (29) are just versions of (8) and (9) and so Theorem 1 guarantees that there is a Bernoulli function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  extending  $\bar{u}$  such that  $\mathcal{O}$  is rationalizable by  $V(x) = \phi(\mathbf{u}(x))$ . *QED*

## 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.
- BAYER, R.-C., S. BOSE, M. POLISSON, and L. RENO (2013): “Ambiguity Revealed”, *IFS Working Papers*, W13/05.
- 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.
- BUSSIECK, M. R., and S. VIGERSKE (2010): “MINLP solver software,” *Wiley Encyclopedia of Operations Research and Management Science*.
- CHAMBERS, C. P., C. LIU, and S.-K. MARTINEZ (2015): “A Test for Risk-Averse Expected Utility,” Mimeo.
- CHAMBERS, C. P., F. ECHENIQUE, and K. SAITO (2015): “Testable Implications of Translation Invariance and Homotheticity: Variational, Maxmin, CARA and CRRA Preferences,” Mimeo.
- 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.
- 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.
- ECHENIQUE, F., and K. SAITO (2015): “Savage in the Market,” *Econometrica*, 83(4), 1467–1495.

- FORGES, F., and E. MINELLI (2009): “Afriat’s Theorem for General Budget Sets,” *Journal of Economic Theory*, 144(1), 135–145.
- FOSTEL, A., H. E. SCARF, and M. J. TODD (2004): “Two New Proofs of Afriat’s Theorem,” *Economic Theory*, 24(1), 211–219.
- GILBOA, I., and D. SCHMEIDLER (1989): “Maxmin Expected Utility with Non-Unique Prior,” *Journal of Mathematical Economics*, 18(2), 141–153.
- 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 (2014): “Parametric Recoverability of Preferences,” Mimeo.
- KOSZEGI, 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.
- 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,” Mimeo.
- NISHIMURA, H., E. A. OK, and J. K.-H. QUAH (2015): “A Comprehensive Approach to Revealed Preference Theory,” *Discussion Paper, Department of Economics, University of Oxford*, 752.
- 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.
- 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.
- 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.