

# An Explicit Representation for Disappointment Aversion and Other Betweenness Preferences\*

Simone Cerreia-Vioglio<sup>a</sup>, David Dillenberger<sup>b</sup>, Pietro Ortoleva<sup>c</sup>

<sup>a</sup>Università Bocconi and Igieo, <sup>b</sup>University of Pennsylvania, <sup>c</sup>Princeton University

September 2018

## Abstract

One of the most well-known models of non-expected utility is Gul (1991)'s model of Disappointment Aversion. This model, however, is defined implicitly, as the solution to a functional equation; its explicit utility representation is unknown, which may limit its applicability. We show that an explicit representation can be easily constructed, using solely the components of the implicit one. We also provide a more general result: an explicit representation for preferences in the Betweenness class that also satisfy Negative Certainty Independence (Dillenberger, 2010).

JEL: D80, D81

KEYWORDS: Betweenness, Cautious Expected Utility, Disappointment Aversion, Utility representation.

---

\*We thank Soo Hong Chew, Faruk Gul, and Klaus Nehring for very useful comments. Cerreia-Vioglio acknowledges the financial support of ERC (grant SDDM-TEA). Dillenberger acknowledges the financial support of NSF (grant SES-1758653). Ortoleva acknowledges the financial support of NSF (grant SES-1559462). Part of this work was done while Dillenberger was visiting the Economics department at Princeton University. He is most grateful to this institution for its hospitality.

# 1 Introduction

One of the most well-known models of non-expected utility preferences is Gul (1991)’s model of Disappointment Aversion (henceforth DA). Its popularity is related both to the intuitive nature of the model, where the value of each outcome is determined relatively to an endogenously-defined “expected” payoff, capturing reference dependence; and to that it generalizes expected utility by adding only one parameter, potentially helping its applicability. DA preferences are behaviorally distinct from alternative models that use probability weighting, and as such can also accommodate violations of expected utility that are conceptually unrelated to distorted beliefs or pessimism — but are rather linked to a form of reference dependence.

Despite its behavioral appeal, there is one limitation to the applicability of the DA model: the value of each lottery is the solution of an equation that changes with the lottery – a so-called implicit representation. The (explicit) utility representation is instead unknown. This may be a concern if one wishes to apply this model to solve optimization problems, as is typically needed in Economics. The same concern applies to the broader class of Betweenness preferences, studied in Dekel (1986) and Chew (1989) and to which the DA model belongs: for such preferences only an implicit representation is known, while the explicit one is still elusive.<sup>1</sup>

The goal of this paper is to address these issues: we provide an explicit representation for DA preferences, showing that it can be easily obtained using solely the components of its implicit one. In fact, our results are more general: we provide an explicit representation for Betweenness preferences that satisfy Negative Certainty Independence (Dillenberger, 2010; Cerreia-Vioglio et al., 2015), of which the most prominent specification of the DA model is a special case (and which is consistent with almost all of the experimental evidence, such as the certainty effect and Allais paradox).

Let  $p$  be a lottery over monetary outcomes. Its value according to the DA model is the unique  $v$  that solves

$$v = \mathbb{E}_p(k_v) \tag{1}$$

where  $k_v$  is given by

$$k_v(x) = \begin{cases} u(x) & u(x) \leq v \\ \frac{u(x)+\beta v}{1+\beta} & u(x) > v \end{cases} .$$

Here  $u$  is a utility function over money, and  $\beta \in (-1, \infty)$  represents the coefficient of either disappointment aversion ( $\beta > 0$ ) or elation seeking ( $\beta < 0$ ). Note that this is an implicit equation, as the value  $v$  appears on both sides of Equation (1). In this model the value  $v$  is similar to expected utility, except that the individual gives an

---

<sup>1</sup>This is the case not only for the broad class, but also for most of its special cases. A notable exception is Chew and MacCrimmon (1979a,b)’s model of weighted-utility.

additional weight  $\beta$  to disappointing outcomes – those with a utility lower than the value of the lottery itself.<sup>2</sup> The DA model is thus a model of endogenous reference dependence: possible payoffs generate disappointment (or elation) depending on how their utilities compare to an endogenously-determined value – the utility of the lottery.<sup>3</sup> When  $\beta > 0$ , the disappointing outcomes receive greater weight, whereas the opposite is true for  $\beta < 0$ , justifying the terms disappointment aversion/elation seeking. If  $\beta = 0$ , the model reduces to expected utility.

In Section 3 we show that these preferences admit the following explicit representation. When  $\beta > 0$ , the case of disappointment aversion, preferences are represented by

$$V(p) = \min_v k_v^{-1}(\mathbb{E}_p(k_v)),$$

while when  $\beta \in (-1, 0)$ , the case of elation-seeking, they are represented by

$$V(p) = \max_v k_v^{-1}(\mathbb{E}_p(k_v)).$$

This means that one can easily construct an explicit representation for preferences in this class using solely the components of the implicit representation in Equation (1) – taking the min or the max of the certainty equivalents computed using each of the possible utilities involved. We also use our results to show additional properties of this model: for example, that it exhibits prudence only if it is expected utility.

There are at least two benefits of having an explicit representation. The first one is practical: it facilitates the applications of these models by simplifying optimization problems with these preferences. This is particularly relevant because DA preferences, while continuous, are not even Gateaux differentiable.<sup>4</sup> Therefore, one cannot apply differential methods via some ad hoc implicit function theorem, or Machina (1982)’s local utility approach and its extensions to the Gateaux case. But when  $\beta > 0$ , our results imply that the same optimization problem becomes a standard max min problem, for which one can apply the well-known Sion (1958)’s minimax theorem. Finding the optimum on a convex set with DA preferences then amounts to computing the optimum for some expected utility functional. We discuss this issue in Section 3.2.

---

<sup>2</sup>To see this, note that the value of a simple lottery  $p$  can equivalently be defined as the unique  $v$  that solves

$$v = \frac{\sum_{\{x:u(x)>v\}} u(x)p(x) + (1 + \beta) \sum_{\{x:u(x)\leq v\}} u(x)p(x)}{1 + \beta \sum_{\{x:u(x)\leq v\}} p(x)}.$$

<sup>3</sup>We should stress that this is different from other models of endogenous reference dependence under risk, e.g., Köszegi and Rabin (2006, 2007): both models are conceptually and behaviorally distinct (Masatlioglu and Raymond, 2016). For example, the DA model satisfies Betweenness, while both models above do not.

<sup>4</sup>See, e.g., Safra and Segal (2009).

The second advantage is instead conceptual: an explicit representation may help capturing the mental process adopted by the agent. While highly idealized, one can easily imagine a cautious (if  $\beta > 0$ ; optimistic otherwise) decision process that involves the max min criterion. It is instead less immediate to take the solution of an implicit equation as a descriptive decision making procedure. This argument is not behavioral, but relies on going beyond the ‘as if’ approach in interpreting representation theorems.<sup>5</sup>

After formally stating the result above for DA preferences (Theorem 3), we discuss an explicit representation for generic Betweenness preferences that also satisfy Negative Certainty Independence (Theorem 4), of which the previous result is a corollary.<sup>6</sup> Again the explicit representation is the min of the certainty equivalents using the functions (called local utilities) used in Dekel (1986)’s implicit representation. We also characterize, in terms of a notion of local risk aversion, the properties of these local utilities.

We conclude the paper by showing one potential application of our main result: how our explicit representation may simplify optimization problems using these preferences.

## 2 Preliminaries

Consider a nontrivial compact interval  $[w, b] \subseteq \mathbb{R}$  of monetary prizes. Let  $\Delta$  be the set of lotteries (Borel probability measures) over  $[w, b]$ , endowed with the topology of weak convergence. We denote by  $x, y, z$  generic elements of  $[w, b]$ ; by  $p, q, r$  generic elements of  $\Delta$ ; and by  $\delta_x \in \Delta$  the degenerate lottery (Dirac measure at  $x$ ) that gives the prize  $x \in [w, b]$  with certainty. The set  $\Delta_0$  denotes the subset of  $\Delta$  of all simple lotteries (convex linear combinations of Dirac measures). We denote by  $C([w, b])$  the space of continuous functions on  $[w, b]$  and we endow it with the topology induced by the supnorm. The set  $\mathcal{U}_{\text{nor}} \subseteq C([w, b])$  is the collection of all strictly increasing and continuous functions  $v : [w, b] \rightarrow \mathbb{R}$  such that  $v(w) = 0$  and  $v(b) = 1$ . Given  $p \in \Delta$  and a strictly increasing  $v \in C([w, b])$ , we define  $c(p, v) = v^{-1}(\mathbb{E}_p(v))$ . Lastly, we denote by  $p \succsim_{FSD} q$ , the case in which  $p$  first order stochastically dominates  $q$  (i.e.,  $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$  for all  $v \in \mathcal{U}_{\text{nor}}$ ).

The primitive of our analysis is a binary relation  $\succsim$  over  $\Delta$ . The symmetric and asymmetric parts of  $\succsim$  are denoted by  $\sim$  and, respectively,  $\succ$ . A certainty equivalent

---

<sup>5</sup>A related argument appears in Chapter 17 of Gilboa (2009). Dekel and Lipman (2010) argued that “While the story need not be literally true for the model to be useful, it plays an important role. Confidence in the story of the model may lead us to trust the models predictions more. Perhaps more importantly, the story affects our intuitions about the model and hence whether and how we use and extend it.”

<sup>6</sup>More precisely, only the case of  $\beta \geq 0$  in Theorem 3 is a corollary of Theorem 4. The other case is obtained using specular techniques.

of a lottery  $p \in \Delta$  is a prize  $x_p \in [w, b]$  such that  $\delta_{x_p} \sim p$ . Throughout the paper, we focus on binary relations  $\succsim$  that satisfy the following three standard assumptions.

**A 1 (Weak Order)** *The relation  $\succsim$  is complete and transitive.*

**A 2 (Continuity)** *For each  $q \in \Delta$ , the sets  $\{p \in \Delta : p \succsim q\}$  and  $\{p \in \Delta : q \succsim p\}$  are closed.*

**A 3 (Strict First Order Stochastic Dominance)** *For each  $p, q \in \Delta$*

$$p \succ_{FSD} q \implies p \succ q.$$

### Betweenness Preferences

We study binary relations that satisfy the following assumption:

**A 4 (Betweenness)** *For each  $p, q \in \Delta$  and  $\lambda \in [0, 1]$*

$$p \sim q \implies p \sim \lambda p + (1 - \lambda) q \sim q.$$

Betweenness implies neutrality toward randomization among equally-good lotteries: if satisfied, then the agent has no preference for, or aversion to, randomization between indifferent lotteries. Binary relations satisfying this property were studied by Dekel (1986) and Chew (1989).

We say that a binary relation is a *Betweenness preference* if and only if it satisfies Weak Order, Continuity, Strict First Order Stochastic Dominance, and Betweenness. Dekel (1986) proves a version of the following result:<sup>7</sup>

**Theorem 1 (Dekel, 1986)** *If  $\succsim$  is a Betweenness preference, then there exists a function  $k : [w, b] \times [0, 1] \rightarrow \mathbb{R}$  such that:*

1.  $x \mapsto k(x, t)$  is strictly increasing and continuous on  $[w, b]$  for all  $t \in (0, 1)$ ,
2.  $t \mapsto k(x, t)$  is continuous on  $(0, 1)$  for all  $x \in [w, b]$ ,
3.  $k(w, t) = 0$  and  $k(b, t) = 1$  for all  $t \in [0, 1]$ ,

---

<sup>7</sup>Dekel's original result deals with a generic set of consequences and considers a weaker form of monotonicity. At the same time, it uses a stronger form of Betweenness. Given these differences, we prove Theorem 1 in Appendix C. Finally, we focus on the normalized representation of Dekel (that is,  $k$  satisfies the condition in point 3). Later we comment on how to use our results for non-normalized representations. Also observe that even though  $k(\cdot, 0)$  and  $k(\cdot, 1)$  are not assumed to be continuous, they are implicitly assumed to be integrable, given (2).

4.  $\succsim$  can be represented by a continuous utility function which strictly preserves first order stochastic dominance,  $\hat{V} : \Delta \rightarrow [0, 1]$ , where for each  $p \in \Delta$ ,  $\hat{V}(p)$  is the unique number in  $[0, 1]$  such that

$$\int_{[w,b]} k(x, \hat{V}(p)) dp = \hat{V}(p). \quad (2)$$

Fixing  $t$ , the function  $k(\cdot, t)$  is called the *local utility at  $t$* . The function  $k$  thus summarizes the collection of local utilities, one for each  $t \in [0, 1]$ . While the theorem above characterizes Betweenness preferences, it does not provide an explicit representation: indeed,  $\hat{V}(p)$  is the solution to (2), thus a fixed point of a functional equation.

An important class of Betweenness preferences is the one arising from Gul (1991)'s model of Disappointment Aversion (DA). These preferences admit a continuous utility function  $\tilde{V} : \Delta \rightarrow \mathbb{R}$  such that, for each  $p \in \Delta$ ,  $\tilde{V}(p)$  is the unique number that solves

$$\int_{[w,b]} \tilde{k}(x, \tilde{V}(p)) dp = \tilde{V}(p) \quad (3)$$

where  $\tilde{k} : [w, b] \times \text{Im } u \rightarrow \mathbb{R}$  is defined by

$$\tilde{k}(x, s) = \begin{cases} u(x) & \text{if } u(x) \leq s \\ \frac{u(x) + \beta s}{1 + \beta} & \text{if } u(x) > s \end{cases} \quad \forall x \in [w, b], \forall s \in \text{Im } u; \quad (4)$$

here  $u$  is a strictly increasing continuous utility function and  $\beta \in (-1, \infty)$ .<sup>8</sup> We discussed its interpretation in the Introduction. We say that a binary relation is a *DA preference* if and only if it admits a utility function  $\tilde{V}$  which satisfies (3) for some pair  $(u, \beta)$ .

## Negative Certainty Independence

As noted by Dillenberger (2010), a DA preference with  $\beta > 0$  satisfies the following axiom.

**A 5 (Negative Certainty Independence)** For each  $p, q \in \Delta$ ,  $x \in [w, b]$ , and  $\lambda \in [0, 1]$

$$p \succsim \delta_x \implies \lambda p + (1 - \lambda) q \succsim \lambda \delta_x + (1 - \lambda) q. \quad (5)$$

Negative Certainty Independence, initially suggested in Dillenberger (2010), is meant to capture the certainty effect. It states that if the sure outcome  $x$  is not enough to compensate the agent for the risky prospect  $p$ , then mixing it with any

---

<sup>8</sup>A careful inspection of (4) also suggests that two types of normalizations are due to link the implicit representation of Gul (1991) to the one of Dekel (1986) as in Theorem 1. In proving our results below, we also address these minor technical points.

other lottery, thus eliminating its certainty appeal, will not result in the mixture of  $\delta_x$  being more attractive than the corresponding mixture of  $p$ . The opposite condition, termed Positive Certainty Independence, simply inverts the role of  $p$  and  $\delta_x$  in (5).

We say that a binary relation is a *Cautious Expected Utility preference* if and only if it satisfies Weak Order, Continuity, Strict First Order Stochastic Dominance, and Negative Certainty Independence. Cerreia-Vioglio et al. (2015) prove the following:<sup>9</sup>

**Theorem 2 (Cerreia-Vioglio, Dillenberger, Ortoleva, 2015)** *If  $\succsim$  is a Cautious Expected Utility preference, then there exists  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  such that  $V : \Delta \rightarrow \mathbb{R}$ , defined by*

$$V(p) = \inf_{v \in \mathcal{W}} c(p, v) \quad \forall p \in \Delta, \quad (6)$$

*is a continuous utility representation of  $\succsim$ .*

### 3 Explicit Representations

We start by providing an explicit representation of DA preferences.

**Theorem 3** *Let  $\succsim$  be a DA preference and  $\mathcal{W}_{\text{da}} = \left\{ \tilde{k}(\cdot, z) \right\}_{z \in \text{Im } u}$ . The following statements are true:*

1. *If  $\beta > 0$ , then  $V : \Delta \rightarrow \mathbb{R}$ , defined by*

$$V(p) = \min_{v \in \mathcal{W}_{\text{da}}} c(p, v) \quad \forall p \in \Delta, \quad (7)$$

*is a continuous utility representation of  $\succsim$ .*

2. *If  $\beta = 0$ , then  $V : \Delta \rightarrow \mathbb{R}$ , defined by*

$$V(p) = c(p, u) \quad \forall p \in \Delta, \quad (8)$$

*is a continuous utility representation of  $\succsim$ .*

3. *If  $\beta < 0$ , then  $V : \Delta \rightarrow \mathbb{R}$ , defined by*

$$V(p) = \max_{v \in \mathcal{W}_{\text{da}}} c(p, v) \quad \forall p \in \Delta,$$

*is a continuous utility representation of  $\succsim$ .*

---

<sup>9</sup>More precisely, Cerreia-Vioglio et al. (2015) state the result below as an “if and only if” but using a weaker form of monotonicity. However, for ease of comparison with Theorem 1, we provide it using Strict First Order Stochastic Dominance.

In the case of disappointment aversion ( $\beta > 0$ ), our utility representation is the smallest of the certainty equivalents obtained using the local utilities. In the opposite case of elation seeking ( $\beta < 0$ ), it is instead the largest. Thus, the difference between the two behaviors is not only in the way in which disappointing/elating outcomes are weighted, but also in how they are aggregated – using the min or the max.

When  $\beta > 0$ , Gul’s model satisfies Negative Certainty Independence. We thus know that it must admit a Cautious Expected Utility representation. The content of Theorem 3 is then to show that this involves *precisely* the local utilities derived in the implicit representation. Thus, the explicit representation can be easily derived using solely the implicit one. When  $\beta < 0$ , the model does *not* satisfy Negative Certainty Independence. Instead, the opposite axiom holds, Positive Certainty Independence (Artstein-Avidan and Dillenberger, 2015). In this case the individual is elation seeking, and violates expected utility in a way opposite to the certainty effect.

We now use Theorem 3, and the machinery developed to prove it, to derive further properties of DA preferences. Recall the notion of prudence (also known as downside risk aversion): the preference for additional risk on the upside rather than the downside of a gamble (Eeckhoudt and Schlesinger, 2006).<sup>10</sup> Intuitively, one could think that risk aversion is to aversion to mean preserving spreads as prudence is to aversion to mean-variance preserving transformations (Menezes et al., 1980). This behavioral feature is often modeled as monotonicity with respect to the *third degree risk order*. Formally, define  $p \succ_{\text{pru}} q$  if and only if  $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$  for all  $v \in C([w, b])$  such that the derivative  $v'$  exists on  $(w, b)$  and is convex. A binary relation  $\succ$  on  $\Delta$  exhibits prudence if and only if  $p \succ_{\text{pru}} q \implies p \succ q$ . Our next result shows that the DA model is inconsistent with prudence unless it is expected utility.

**Proposition 1** *Let  $\succ$  be a DA preference. It exhibits prudence if and only if  $\beta = 0$  (i.e., it is expected utility), and  $u'$  exists on  $(w, b)$  and is convex.*

### 3.1 A General Result

We now show that any generic Betweenness preference that satisfies Negative Certainty Independence also admits an *explicit* representation of the Cautious Expected Utility form where the utilities in  $\mathcal{W}$  are the local ones obtained in Theorem 1, that is,  $\mathcal{W}_{\text{bet}} = \{k(\cdot, t)\}_{t \in (0,1)}$ .

**Theorem 4** *Let  $\succ$  be a Betweenness preference. The following statements are equivalent:*

---

<sup>10</sup>The name prudence and its relation with precautionary savings date back to Kimball (1990). In the case of expected utility preferences, prudence implies preference for skewness.



(i)  $\succsim$  satisfies Negative Certainty Independence;

(ii) The functional  $V : \Delta \rightarrow \mathbb{R}$ , defined by

$$V(p) = \min_{v \in \mathcal{W}_{\text{bet}}} c(p, v) \quad \forall p \in \Delta, \quad (9)$$

is a continuous utility representation of  $\succsim$ . In particular, for each  $p \in \Delta \setminus \{\delta_w, \delta_b\}$  the function  $v_p = k(\cdot, \hat{V}(p))$  is such that

$$v_p \in \operatorname{argmin}_{v \in \mathcal{W}_{\text{bet}}} c(p, v). \quad (10)$$

Like in the case of the previous result, Theorem 4 shows that these preferences admit an explicit representation of the Cautious Expected Utility class – which follows from Theorem 2. As before, the key contribution is to show that the utilities involved are exactly the local utilities identified in Theorem 1, included in  $\mathcal{W}_{\text{bet}}$ . Again, the explicit representation can be easily derived using solely the components of the implicit one. In addition, Equation (10) shows that the local utility giving the implicit representation of Dekel (1986) is also the one achieving the minimum in representation (9).

While Theorem 4 provides an explicit characterization for Betweenness preferences that satisfy Negative Certainty Independence, one may be interested in the following question: what are the characteristics of the local functions  $k(\cdot, t)$  that guarantee that preferences satisfy Negative Certainty Independence (and thus admit a representation as in Theorem 4)? Our next result provides an answer. In stating it, the following notation will be used: Given  $f : [0, 1] \rightarrow [0, 1]$ , we say that  $f$  is *convex at*  $t \in (0, 1)$  if and only if for each  $n \in \mathbb{N}$ ,  $\{t_i\}_{i=1}^n \subseteq [0, 1]$ , and  $\{\lambda_i\}_{i=1}^n \subseteq [0, 1]$  such that  $\sum_{i=1}^n \lambda_i = 1$

$$t = \sum_{i=1}^n \lambda_i t_i \implies f(t) \leq \sum_{i=1}^n \lambda_i f(t_i).$$

For each  $s, t \in (0, 1)$ , define  $f_{s,t}$  to be the transformation from  $k(\cdot, t)$  to  $k(\cdot, s)$ , that is,  $f_{s,t} : [0, 1] \rightarrow [0, 1]$  is such that  $k(x, s) = f_{s,t}(k(x, t))$  for all  $x \in [0, 1]$ . Note that  $f_{s,t}$  must exist since  $k(\cdot, t)$  and  $k(\cdot, s)$  are strictly increasing and continuous. Moreover,  $f_{s,t}$  is strictly increasing, continuous, and such that  $f_{s,t}(0) = 0$  and  $f_{s,t}(1) = 1$ .

**Proposition 2** *Let  $\succsim$  be a Betweenness preference. The following statements are equivalent:*

- (i) For each  $t \in (0, 1)$  and for each  $s \in (0, 1)$  the function  $f_{s,t}$  is convex at  $t$ ;
- (ii)  $\succsim$  satisfies Negative Certainty Independence.

Proposition 2 states that testing Negative Certainty Independence amounts to checking if for each  $t \in (0, 1)$  the transformations  $f_{s,t}$  are convex at  $t$  for all  $s \in (0, 1)$ . This is a handy tool because  $f_{s,t} = k(\cdot, s) \circ k^{-1}(\cdot, t)$  and is thus computable.

**Remark 1** This property of convexity is implied by the following sufficient condition: the subdifferential of  $f_{s,t}$  is nonempty at  $t$ ,  $\partial f_{s,t}(t) \neq \emptyset$ .<sup>11</sup> This takes a simple geometric interpretation, as it amounts to saying that the graph of  $f_{s,t}$  is supported by a line at the point  $(t, f_{s,t}(t))$ , that is, there exists a function  $g : [0, 1] \rightarrow \mathbb{R}$  such that  $g(t') = mt' + l$  for all  $t' \in [0, 1]$ , where  $l, m \in \mathbb{R}$  and

$$f_{s,t}(t) = g(t) \text{ as well as } g(t') \leq f_{s,t}(t') \quad \forall t' \in [0, 1]. \quad (11)$$

▽

Next, we discuss the possibility of obtaining more parsimonious explicit representations for the preferences we consider – for example, ones that involve finitely many utilities only. The following result shows that this is not the case: within the class of Cautious Expected Utility preferences, whenever Betweenness holds, either the preference is expected utility, or the set  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  must be infinite.<sup>12</sup>

**Proposition 3** *Let  $\succsim$  be a Cautious Expected Utility preference. If  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  satisfies (6) and  $|\mathcal{W}| < \infty$ , then either  $\succsim$  satisfies Independence or  $\succsim$  violates Betweenness.*

Finally, we show that there are models in the Betweenness class that satisfy Negative Certainty Independence but are not DA preferences. The following is an example that considers another form of disappointment aversion.

**Example 1** Consider a Betweenness preference with local utilities  $k : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$  defined as

$$k(x, t) = \begin{cases} x & \text{if } x \leq t \\ x^2 - tx + t & \text{if } x > t \end{cases} \quad \forall x \in [0, 1], \forall t \in [0, 1].$$

This retains the idea of disappointment aversion, but allows the weight to depend on the value  $x$ .<sup>13</sup> In Appendix C, relying on Proposition 2 and Remark 1, we show that these preferences satisfy Negative Certainty Independence and therefore admit an explicit representation. ▽

## 3.2 Explicit Representations and Applications

In this section we illustrate how our explicit representation can be useful for applications. In economic models agents need to pick the best action from a set of options,

---

<sup>11</sup>We provide a proof in Appendix C.

<sup>12</sup>Appendix C.1 provides further results and a discussion.

<sup>13</sup>This is a special case of Chew (1985)'s model of semi-implicit weighted utility, where  $[w, b]$  is set to be equal to  $[0, 1]$ .

which is often assumed to be convex and compact. As we pointed out in the Introduction, however, differential methods cannot be used to solve these optimization problems when preferences belong to the DA class, as these are not Gateaux differentiable. We will now show how our explicit representation results might facilitate solving this class of problems.

Consider a Betweenness preference  $\succsim$  that satisfies Negative Certainty Independence — for DA preferences, this amounts to assuming  $\beta \geq 0$ . By Theorem 4,  $\succsim$  is represented by

$$V(p) = \min_{v \in \mathcal{W}_{\text{bet}}} c(p, v) \quad \forall p \in \Delta \quad (12)$$

where  $\mathcal{W}_{\text{bet}} = \{k(\cdot, t)\}_{t \in (0,1)}$ . We further assume:

- $\alpha$ )  $k(\cdot, t)$  is strictly increasing and continuous on  $[w, b]$  for all  $t \in [0, 1]$ ; and
- $\beta$ )  $k$  is jointly continuous on  $[w, b] \times [0, 1]$ .

Note that both assumptions are satisfied by DA preferences as well as the preferences in Example 1.

**Proposition 4** *Let  $\succsim$  be a Betweenness preference that satisfies Negative Certainty Independence and such that  $k$  satisfies  $\alpha$  and  $\beta$ . If  $A$  is a convex and compact subset of  $\Delta$ , then*

$$\max_{p \in A} \min_{v \in \bar{\text{co}}(\mathcal{W}_{\text{bet}})} c(p, v) = \min_{v \in \bar{\text{co}}(\mathcal{W}_{\text{bet}})} \max_{p \in A} c(p, v).$$

*In particular, if  $\bar{p} \in A$  is such that  $V(\bar{p}) \geq V(p)$  for all  $p \in A$ , then there exists  $\hat{v} \in \bar{\text{co}}(\mathcal{W}_{\text{bet}})$  such that  $\mathbb{E}_{\bar{p}}(\hat{v}) \geq \mathbb{E}_p(\hat{v})$  for all  $p \in A$ .*

The result above says that any alternative that maximizes the original preference in  $A$  is also a maximizer of an expected utility preference with Bernoulli utility  $\hat{v}$ . The function  $\hat{v}$  is a “convex linear combination” of the utilities in  $\mathcal{W}_{\text{bet}}$  which are used to represent  $\succsim$ . The conceptual importance of Proposition 4 can be illustrated assuming that  $\succsim$  is risk averse (i.e.,  $k(\cdot, t)$  is concave for all  $t \in (0, 1)$ ). Assume that for the optimization problem at hand (say, a portfolio problem), it is known that under Expected Utility risk aversion leads to certain qualitative properties of the maximizers (like preference for diversification). If each function  $k(\cdot, t)$  is concave, then  $\hat{v}$  is concave. This, paired with Proposition 4, implies that the same qualitative properties are shared by the maximizers of the Betweenness preference  $\succsim$ .

# Appendix

## A The Expected Utility Core

We start by defining the expected utility core of  $\succsim$ , i.e., the subrelation  $\succsim'$  defined as:<sup>14</sup>

$$p \succsim' q \iff \lambda p + (1 - \lambda)r \succsim \lambda q + (1 - \lambda)r \quad \forall \lambda \in (0, 1], \forall r \in \Delta.$$

This notion is useful for two reasons. First, as Remark 2 below shows, in order to find a canonical representation of a Cautious Expected Utility preference, it is sufficient to find an expected multi-utility representation of  $\succsim'$ . This is instrumental in proving Theorem 4 (cf. Proposition 9). Second, as shown by Cerreia-Vioglio et al. (2017), in general  $\succsim'$  summarizes the risk attitudes of the decision maker irrespective of whether or not  $\succsim$  satisfies Negative Certainty Independence. In particular,  $\succsim$  is averse to Mean Preserving Spreads if and only if  $\succsim'$  is, which is equivalent to have all the utilities representing the latter being concave. Similar considerations hold for prudence, a fact we will exploit while proving Proposition 1.

**Remark 2** In addition to what is stated in Theorem 2, it can also be shown that the following are true:

1. There exists a set  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  such that

$$p \succsim' q \iff \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \quad \forall v \in \mathcal{W} \tag{13}$$

and  $V : \Delta \rightarrow \mathbb{R}$  defined as in Equation (6) is a continuous utility representation of  $\succsim$ .

2. If  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  satisfies (13), then it satisfies (6).
3. The set  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  can be chosen to be

$$\mathcal{W}_{\text{max-nor}} = \{v \in \mathcal{U}_{\text{nor}} : p \succsim' q \implies \mathbb{E}_p(v) \geq \mathbb{E}_q(v)\}.$$

4. If  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  satisfies (13), then

$$\mathcal{W} \subseteq \mathcal{W}_{\text{max-nor}} \text{ as well as } \bar{\text{co}}(\mathcal{W}) = \text{cl}(\mathcal{W}_{\text{max-nor}}).$$

▽

We call a set  $\mathcal{W}$  that satisfies (6) and (13) a canonical representation.

---

<sup>14</sup>Under Axioms A 1-2, one can show that  $\succsim'$  satisfies all the assumptions of expected utility with possibly the exception of completeness; and that it is the largest subrelation of  $\succsim$  satisfying these properties. See Cerreia-Vioglio (2009); Cerreia-Vioglio et al. (2015, 2017).

## B Compactness and Risk Aversion

In studying the intersection of Betweenness preferences and the Cautious Expected Utility model, an important role is played by the latter admitting a compact representation. In this appendix, we show that this feature is behaviorally related to two properties: Strict First Order Stochastic Dominance and a notion of risk aversion.

We start by introducing a notion of comparative risk aversion which is much weaker than, but in line with, the one used by Machina (1982) and which rests on the idea of Simple Compensated Spread (SCS). Given two lotteries  $p, q \in \Delta$ , the idea of SCS involves two elements: a) a notion of “being more dispersed” and b) a notion of having “the same value”. The former is objective, while the latter is subjective. More precisely, given  $r \in \Delta$ , denote by  $F_r : [w, b] \rightarrow \mathbb{R}$  the cumulative distribution function

$$F_r(x) = r([w, x]) \quad \forall x \in [w, b].$$

Given a binary relation  $\succsim$ ,  $q$  is a SCS of  $p$  for  $\succsim$  if and only if

a) there exists  $z \in [w, b]$  such that

$$\begin{cases} F_p(x) \leq F_q(x) & \forall x \in [w, z] \\ F_p(x) \geq F_q(x) & \forall x \in [z, b] \end{cases}. \quad (14)$$

b)  $p \sim q$ .

One of the reasons why Machina (1982) introduced the notion of Simple Compensated Spread is to define a notion of comparative risk aversion for non-expected utility models. Machina’s definition of comparative risk aversion indeed reads as follows:  $\succsim_1$  is more risk averse than  $\succsim_2$  if and only if whenever  $q$  is a SCS of  $p$  for  $\succsim_2$ , then  $p \succsim_1 q$ . Intuitively, if  $q$  is more dispersed than  $p$ , that is, it is riskier, but it is still good enough to compensate decision maker  $\succsim_2$ , then it is weakly less good for the more risk averse decision maker  $\succsim_1$ .

In what follows, we restrict ourselves to a particular class of SCSs and to a particular class of decision makers  $\succsim_2$ .

**Definition 1** *Let  $p, q \in \Delta_0$  and  $u \in \mathcal{U}_{\text{nor}}$ . We say that  $q$  is an Extreme Simple Compensated Spread of  $p$  if and only if there exist  $\bar{x} \in (w, b)$  and  $\gamma \in (0, 1)$  such that:*

1.  $p(x) = q(x)$  for all  $x \in (w, b) / \{\bar{x}\}$ ;
2.  $q(b) - p(b) = \gamma(p(\bar{x}) - q(\bar{x})) \geq 0$  and  $q(w) - p(w) = (1 - \gamma)(p(\bar{x}) - q(\bar{x})) \geq 0$ ;
3.  $c(p, u) = c(q, u)$ .

Assume  $q$  is an Extreme Simple Compensated Spread of  $p$ . Intuitively, conditions 1 and 2 capture the idea of  $q$  “being more dispersed” than  $p$ , since probability mass is shifted from an interior point  $\bar{x}$  to the extrema  $w$  and  $b$ . Condition 3 instead captures the idea that  $p$  and  $q$  “have the same value”. Indeed,  $p$  and  $q$  must have the same *quasi-arithmetic* mean with respect to  $u$ .

It is easy to verify that given  $p, q \in \Delta_0$  and  $u \in \mathcal{U}_{\text{nor}}$ , if  $q$  is an Extreme Simple Compensated Spread of  $p$ , then  $q$  is a SCS of  $p$  for the expected utility binary relation  $\succsim_2$  induced by  $u$ .<sup>15</sup>

**Remark 3** Given  $p, q \in \Delta_0$ , if  $q$  is an Extreme Simple Compensated Spread of  $p$ , we will denote it by  $p \succsim_{\text{ESCS}} q$ . Note that this latter notation is incomplete, since it does not refer to  $u$  explicitly. Nevertheless, in what follows, it will always be clear from the context what is  $u$ .  $\nabla$

**Definition 2** Let  $\succsim$  be a binary relation on  $\Delta$ . We say that  $\succsim$  is not infinitely risk loving if and only if there exists  $u \in \mathcal{U}_{\text{nor}}$  such that

$$p \succsim_{\text{ESCS}} q \implies p \succsim q. \quad (15)$$

Alternatively, we say that  $\succsim$  satisfies NIRL.

In light of Machina’s notion of comparative risk attitudes,  $\succsim$  satisfies NIRL if and only if it is more risk averse than some expected utility decision maker, where, in our case, aversion to Simple Compensated Spreads is imposed on the much *smaller* class of extreme spreads. We proceed by characterizing the NIRL property within the class of Cautious Expected Utility preferences. Before doing so, we introduce a property, Sensitivity, which will help our analysis and, given all the other assumptions, is equivalent to NIRL.

**A 6 (Sensitivity)** *The binary relation  $\succsim$  is such that:*

1. *For each  $\lambda \in (0, 1)$  there exists  $x \in (w, b)$  such that*

$$\delta_x \succsim' \lambda \delta_b + (1 - \lambda) \delta_w.$$

2. *For each  $x \in (w, b)$  there exists  $\lambda \in (0, 1)$  such that*

$$\delta_x \succsim' \lambda \delta_b + (1 - \lambda) \delta_w.$$

The next proposition elaborates on the relation between NIRL and Sensitivity.

---

<sup>15</sup>Let  $z$  in (14) be  $\bar{x}$  of Definition 1.

**Proposition 5** *Let  $\succcurlyeq$  be a Cautious Expected Utility preference and  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  satisfy (13). If  $\succcurlyeq$  satisfies NIRL, then the following statements are true:*

(a) *There exists  $u \in \mathcal{U}_{\text{nor}}$  such that for each  $p, q \in \Delta_0$*

$$p \succcurlyeq_{\text{ESCS}} q \implies p \succcurlyeq' q.$$

(b)  *$\succcurlyeq$  satisfies Sensitivity.*

**Proof.** Since  $\succcurlyeq$  satisfies NIRL, there exists  $u \in \mathcal{U}_{\text{nor}}$  such that for each  $p, q \in \Delta_0$

$$p \succcurlyeq_{\text{ESCS}} q \implies p \succcurlyeq q. \tag{16}$$

In the rest of the proof,  $u$  will be fixed.

(a). Consider  $p, q \in \Delta_0$ . It follows that

$$\begin{aligned} p \succcurlyeq_{\text{ESCS}} q &\implies \lambda p + (1 - \lambda)r \succcurlyeq_{\text{ESCS}} \lambda q + (1 - \lambda)r \quad \forall \lambda \in (0, 1], \forall r \in \Delta_0 \\ &\implies \lambda p + (1 - \lambda)r \succcurlyeq \lambda q + (1 - \lambda)r \quad \forall \lambda \in (0, 1], \forall r \in \Delta_0 \\ &\implies \lambda p + (1 - \lambda)r \succcurlyeq \lambda q + (1 - \lambda)r \quad \forall \lambda \in (0, 1], \forall r \in \Delta \implies p \succcurlyeq' q. \end{aligned}$$

The first implication follows from the definition of  $\succcurlyeq_{\text{ESCS}}$ . The second implication follows from (16). The third implication follows from the density of  $\Delta_0$  in  $\Delta$  and since  $\succcurlyeq$  satisfies Continuity. The last implication follows from the definition of  $\succcurlyeq'$ .

(b). Consider  $\lambda \in (0, 1)$ . Define  $x \in [w, b]$  to be such that  $u(x) = \mathbb{E}_{\lambda\delta_b + (1-\lambda)\delta_w}(u)$ . Since  $u \in \mathcal{U}_{\text{nor}}$ , note that  $x \in (w, b)$ . It is immediate to check that  $\delta_x \succcurlyeq_{\text{ESCS}} \lambda\delta_b + (1 - \lambda)\delta_w$ . By point (a), it follows that  $\delta_x \succcurlyeq' \lambda\delta_b + (1 - \lambda)\delta_w$ . Viceversa, consider  $x \in (w, b)$ . Let  $\lambda \in [0, 1]$  be such that  $u(x) = \mathbb{E}_{\lambda\delta_b + (1-\lambda)\delta_w}(u)$ . Since  $u \in \mathcal{U}_{\text{nor}}$ , note that  $\lambda \in (0, 1)$ . It is immediate to check that  $\delta_x \succcurlyeq_{\text{ESCS}} \lambda\delta_b + (1 - \lambda)\delta_w$ . By point (a), it follows that  $\delta_x \succcurlyeq' \lambda\delta_b + (1 - \lambda)\delta_w$ .  $\blacksquare$

The next result characterizes NIRL for Cautious Expected Utility preferences.

**Proposition 6** *Let  $\succcurlyeq$  be a Cautious Expected Utility preference and  $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$  satisfy (6) and (13). The following statements are equivalent:*

- (i)  *$\succcurlyeq$  satisfies NIRL;*
- (ii) *There exists  $u \in \mathcal{U}_{\text{nor}}$  for each  $v \in \mathcal{W}$  such that there exists  $f_v : [0, 1] \rightarrow [0, 1]$  where  $v = f_v \circ u$  and  $f_v(\gamma) \geq \gamma$  for all  $\gamma \in [0, 1]$ ;*
- (iii) *There exists  $u \in \mathcal{U}_{\text{nor}}$  such that  $v \geq u$  for all  $v \in \mathcal{W}$ .*

**Proof.** (i) implies (ii). Since  $\succ$  satisfies NIRL, there exists  $u \in \mathcal{U}_{\text{nor}}$  such that for each  $p, q \in \Delta_0$

$$p \succ_{ESCS} q \implies p \succ q.$$

Since each  $v \in \mathcal{W}$  is strictly increasing and continuous and so is  $u$ , for each  $v \in \mathcal{W}$  there exists  $f_v : [0, 1] \rightarrow \mathbb{R}$  which is strictly increasing, continuous, and such that  $v = f_v \circ u$ . Since  $u$  and  $v$  are normalized,  $f_v(0) = 0 = f_v(1) - 1$  for all  $v \in \mathcal{W}$ . Next, consider  $\gamma \in (0, 1)$  and define  $q = \gamma \delta_b + (1 - \gamma) \delta_w$ . Define  $x \in [w, b]$  to be such that  $u(x) = \mathbb{E}_q(u)$ . Since  $u \in \mathcal{U}_{\text{nor}}$ , note that  $x \in (w, b)$ . It is immediate to check that  $\delta_x \succ_{ESCS} q$ . By the proof of point (a) of Proposition 5 and since  $\mathcal{W}$  represents  $\succ'$ , it follows that  $\delta_x \succ' q$ , that is, for each  $v \in \mathcal{W}$

$$f_v(\gamma) = f_v(\mathbb{E}_q(u)) = f_v(u(x)) = v(x) = \mathbb{E}_{\delta_x}(v) \geq \mathbb{E}_q(v) = \gamma.$$

Since  $\gamma$  was arbitrarily chosen, the implication follows.

(ii) implies (iii). Since  $u \in \mathcal{U}_{\text{nor}}$ , for each  $x \in [w, b]$  we have that  $u(x) \in [0, 1]$ . Thus, we can conclude that

$$v(x) = f_v(u(x)) \geq u(x) \quad \forall x \in [w, b], \forall v \in \mathcal{W},$$

proving the implication.

(iii) implies (i). Consider  $p, q \in \Delta_0$  and assume that  $p \succ_{ESCS} q$  with respect to  $u$ . If  $p(\bar{x}) = q(\bar{x})$ , where  $\bar{x}$  is like in Definition 1, then  $p = q$  and  $p \succ q$ . Assume then that  $p(\bar{x}) > q(\bar{x})$ . Consider  $v \in \mathcal{U}_{\text{nor}}$ . Since  $p$  and  $q$  are in  $\Delta_0$ , it follows that

$$\mathbb{E}_p(v) = \sum_{x \in [w, b]} v(x) p(x) \quad \text{and} \quad \mathbb{E}_q(v) = \sum_{x \in [w, b]} v(x) q(x).$$

This implies that

$$\begin{aligned} \mathbb{E}_p(v) - \mathbb{E}_q(v) &= (p(b) - q(b))v(b) + (p(w) - q(w))v(w) + (p(\bar{x}) - q(\bar{x}))v(\bar{x}) \\ &= (p(b) - q(b))v(b) + (p(\bar{x}) - q(\bar{x}))v(\bar{x}). \end{aligned}$$

Since  $p \succ_{ESCS} q$  and  $v$  was arbitrarily chosen in  $\mathcal{U}_{\text{nor}}$ , we can conclude that

$$\begin{aligned} \mathbb{E}_p(v) - \mathbb{E}_q(v) &= -\gamma(p(\bar{x}) - q(\bar{x}))v(b) + (p(\bar{x}) - q(\bar{x}))v(\bar{x}) \\ &= (-\gamma v(b) + v(\bar{x}))(p(\bar{x}) - q(\bar{x})) \\ &= (-\gamma + v(\bar{x}))(p(\bar{x}) - q(\bar{x})) \quad \forall v \in \mathcal{U}_{\text{nor}} \end{aligned}$$

where  $\gamma \in (0, 1)$  is like in Definition 1. Since  $u \in \mathcal{U}_{\text{nor}}$ , we have that

$$\mathbb{E}_p(u) - \mathbb{E}_q(u) = (-\gamma + u(\bar{x}))(p(\bar{x}) - q(\bar{x})).$$



Since  $\mathbb{E}_p(u) = \mathbb{E}_q(u)$  and  $p(\bar{x}) > q(\bar{x})$ , this implies that  $-\gamma + u(\bar{x}) = 0$ . Since  $v \geq u$  for all  $v \in \mathcal{W}$ , this implies that  $-\gamma + v(\bar{x}) \geq -\gamma + u(\bar{x}) = 0$  for all  $v \in \mathcal{W}$ . In turn, this yields that  $0 \leq (-\gamma + v(\bar{x}))(p(\bar{x}) - q(\bar{x})) = \mathbb{E}_p(v) - \mathbb{E}_q(v)$  for all  $v \in \mathcal{W}$ . We can conclude that  $c(p, v) \geq c(q, v)$  for all  $v \in \mathcal{W}$ , yielding that  $p \succcurlyeq q$  and proving NIRL. ■

**Remark 4** Note that (iii) implies (i) holds also if  $\mathcal{W}$  only satisfies (6). ▽

The next corollary shows that if  $\succcurlyeq$  is risk averse, that is averse to Mean Preserving Spreads, then  $\succcurlyeq$  is not infinitely risk loving according to Definition 2.

**Corollary 1** *Let  $\succcurlyeq$  be a Cautious Expected Utility preference. If  $\succcurlyeq$  is risk averse, then  $\succcurlyeq$  satisfies NIRL.*

**Proof.** Since  $\succcurlyeq$  is a Cautious Expected Utility preference, consider a set  $\mathcal{W}$  that satisfies (6) and (13). By (Cerreia-Vioglio et al., 2015, Theorem 3), if  $\succcurlyeq$  is risk averse, then  $\mathcal{W}$  is such that each  $v \in \mathcal{W}$  is concave. Consider  $u \in \mathcal{U}_{\text{nor}}$  to be such that  $u(x) = \frac{x-w}{b-w}$  for all  $x \in [w, b]$ . Since each  $v \in \mathcal{W}$  is concave and normalized, it is immediate to see that  $v \geq u$ . By Proposition 6, the statement follows. ■

In order to characterize the compactness of the set  $\mathcal{W}$ , we are going to need the following two ancillary results. Lemma 1 is routine we report a proof for ease of reference.

**Lemma 1** *Let  $\mathcal{W}$  be a subset of  $\mathcal{U}_{\text{nor}}$ . The following statements are equivalent:*

- (i)  $\mathcal{W}$  is sequentially compact with respect to the pointwise convergence topology;
- (ii)  $\mathcal{W}$  is norm compact.

**Proof.** (ii) implies (i). It is trivial.

(i) implies (ii). Consider  $\{v_n\}_{n \in \mathbb{N}} \subseteq \mathcal{W}$ . Observe that, by construction,  $\{v_n\}_{n \in \mathbb{N}}$  is uniformly bounded. By assumption, there exists  $\{v_{n_k}\}_{k \in \mathbb{N}} \subseteq \{v_n\}_{n \in \mathbb{N}}$  and  $v \in \mathcal{W}$  such that  $v_{n_k}(x) \rightarrow v(x)$  for all  $x \in [w, b]$ . By (Aliprantis and Burkinshaw, 1998, p. 79) and since  $v$  is a continuous function and each  $v_{n_k}$  is increasing, it follows that this convergence is uniform, proving the statement. ■

**Theorem 5** *Let  $\succcurlyeq$  be a Cautious Expected Utility preference and let  $V : \Delta \rightarrow \mathbb{R}$  be a continuous utility representation of  $\succcurlyeq$  such that*

$$V(p) = \inf_{v \in \mathcal{W}} c(p, v) \quad \forall p \in \Delta$$

where  $\mathcal{W} = \mathcal{W}_{\text{max-nor}}$ . *If  $\succcurlyeq$  satisfies Sensitivity, then  $\mathcal{W}$  is sequentially compact with respect to the topology of pointwise convergence.*

**Proof.** Since  $\succcurlyeq$  satisfies Strict First Order Stochastic Dominance, it follows that  $V$  strictly preserves first order stochastic dominance. We next prove a few ancillary claims.

*Claim 1: For each  $\varepsilon > 0$  there exists  $\delta \in (0, b - w)$  such that for each  $v \in \mathcal{W}$*

$$v(w + \delta) < \varepsilon. \quad (17)$$

*Proof of the Claim.* By contradiction, assume that there exists  $\bar{\varepsilon} > 0$  such that for each  $\delta \in (0, b - w)$  there exists  $v_\delta \in \mathcal{W}$  such that  $v_\delta(w + \delta) \geq \bar{\varepsilon}$ . In particular, for each  $k \in \mathbb{N}$  such that  $\frac{1}{k} < b - w$  there exists  $v_k \in \mathcal{W}$  such that  $v_k(w + \frac{1}{k}) \geq \bar{\varepsilon}$ . Define  $\lambda_k \in [0, 1]$  for each  $k > \frac{1}{b-w}$  to be such that

$$\lambda_k = v_k\left(w + \frac{1}{k}\right) \geq \bar{\varepsilon} > 0. \quad (18)$$

Define  $p_k = \lambda_k \delta_b + (1 - \lambda_k) \delta_w$  for all  $k > \frac{1}{b-w}$ . Without loss of generality, we can assume that  $\lambda_k \rightarrow \lambda$ . Note that  $\lambda \geq \bar{\varepsilon}$ . Define  $p = \lambda \delta_b + (1 - \lambda) \delta_w$ . It is immediate to verify that  $p_k \rightarrow p$  and  $p \succ_{FSD} \delta_w$ . Since  $\mathbb{E}_{p_k}(v_k) = \lambda_k$  and by (18) and by definition of  $V$ , it follows that

$$w \leq V(p_k) \leq v_k^{-1}(\mathbb{E}_{p_k}(v_k)) = w + \frac{1}{k} \quad \forall k > \frac{1}{b-w}.$$

Since  $V$  is continuous and by passing to the limit, we have that  $V(p) = w = V(\delta_w)$ , a contradiction with  $V$  strictly preserving first order stochastic dominance.  $\square$

*Claim 2: For each  $x \in (w, b)$  and for each  $\varepsilon > 0$  there exists  $\delta \in (0, \min\{x - w, b - x\})$  such that for each  $v \in \mathcal{W}$*

$$v(x + \delta) - v(x - \delta) < \varepsilon. \quad (19)$$

*Proof of the Claim.* By contradiction, assume that there exist  $\bar{x} \in (w, b)$  and  $\bar{\varepsilon} > 0$  such that for each  $\delta \in (0, \min\{\bar{x} - w, b - \bar{x}\})$  there exists  $v_\delta \in \mathcal{W}$  such that  $v_\delta(\bar{x} + \delta) - v_\delta(\bar{x} - \delta) \geq \bar{\varepsilon}$ . In particular, for each  $k \in \mathbb{N}$  such that  $\frac{1}{k} < \min\{\bar{x} - w, b - \bar{x}\}$  there exists  $v_k \in \mathcal{W}$  such that  $v_k(\bar{x} + \frac{1}{k}) - v_k(\bar{x} - \frac{1}{k}) \geq \bar{\varepsilon}$ . Define  $\lambda_k \in [0, 1]$  for each  $k > \frac{1}{\min\{\bar{x} - w, b - \bar{x}\}}$  to be such that

$$\lambda_k = v_k\left(\bar{x} + \frac{1}{k}\right) - v_k\left(\bar{x} - \frac{1}{k}\right) \geq \bar{\varepsilon} > 0. \quad (20)$$

Define  $p_k = \lambda_k \delta_b + (1 - \lambda_k) \delta_{\bar{x} - \frac{1}{k}}$  for all  $k > \frac{1}{\min\{\bar{x} - w, b - \bar{x}\}}$ . Without loss of generality, we can assume that  $\lambda_k \rightarrow \lambda$ . Note that  $\lambda \geq \bar{\varepsilon}$ . Define  $p = \lambda \delta_b + (1 - \lambda) \delta_{\bar{x}}$ . It is immediate

to verify that  $p_k \rightarrow p$  and  $p \succ_{FSD} \delta_{\bar{x}}$ . By (20), it follows that for each  $k > \frac{1}{\min\{\bar{x}-w, b-\bar{x}\}}$

$$\begin{aligned}\mathbb{E}_{p_k}(v_k) &= \lambda_k v_k(b) + (1 - \lambda_k) v_k\left(\bar{x} - \frac{1}{k}\right) \\ &= v_k\left(\bar{x} + \frac{1}{k}\right) - v_k\left(\bar{x} - \frac{1}{k}\right) + (1 - \lambda_k) v_k\left(\bar{x} - \frac{1}{k}\right) \\ &= v_k\left(\bar{x} + \frac{1}{k}\right) - \lambda_k v_k\left(\bar{x} - \frac{1}{k}\right) \leq v_k\left(\bar{x} + \frac{1}{k}\right).\end{aligned}$$

By definition of  $V$  and since  $V$  (strictly) preserves first order stochastic dominance, this implies that

$$\bar{x} - \frac{1}{k} \leq V(p_k) \leq v_k^{-1}(\mathbb{E}_{p_k}(v_k)) \leq \bar{x} + \frac{1}{k} \quad \forall k > \frac{1}{\min\{\bar{x}-w, b-\bar{x}\}}.$$

Since  $V$  is continuous and by passing to the limit, we have that  $V(p) = \bar{x} = V(\delta_{\bar{x}})$ , a contradiction with  $V$  strictly preserving first order stochastic dominance.  $\square$

*Claim 3:* For each  $\varepsilon \in (0, 1)$  there exists  $\delta \in (0, b - w)$  such that for each  $v \in \mathcal{W}$

$$1 - v(b - \delta) \leq \varepsilon. \quad (21)$$

*Proof of the Claim.* Given  $\varepsilon \in (0, 1)$ , define  $\lambda_\varepsilon \in (0, 1)$  by  $\lambda_\varepsilon = 1 - \varepsilon$ . Since  $\succ$  satisfies Sensitivity, we have that there exists  $x \in (w, b)$  such that  $\delta_x \succ' \lambda_\varepsilon \delta_b + (1 - \lambda_\varepsilon) \delta_w$ . Define  $\delta = b - x$ . Note that  $\delta \in (0, b - w)$ . Since  $\mathcal{W} = \mathcal{W}_{\max - \text{nor}}$  represents  $\succ'$ , this implies that

$$v(b - \delta) = v(x) \geq \lambda_\varepsilon v(b) + (1 - \lambda_\varepsilon) v(w) = \lambda_\varepsilon \quad \forall v \in \mathcal{W}_{\max - \text{nor}} = \mathcal{W},$$

proving the statement.  $\square$

*Claim 4:* For each  $x \in (w, b)$  we have  $\inf_{v \in \mathcal{W}} v(x) > 0$ .

*Proof of the Claim.* Since  $\succ$  satisfies Sensitivity, we have that for each  $x \in (w, b)$  there exists  $\lambda \in (0, 1)$  such that  $\delta_x \succ' \lambda \delta_b + (1 - \lambda) \delta_w$ . Since  $\mathcal{W} = \mathcal{W}_{\max - \text{nor}}$  represents  $\succ'$ , this implies that

$$v(x) \geq \lambda v(b) + (1 - \lambda) v(w) = \lambda \quad \forall v \in \mathcal{W}_{\max - \text{nor}} = \mathcal{W},$$

yielding that  $\inf_{v \in \mathcal{W}} v(x) \geq \lambda > 0$ .  $\square$

Consider  $\{v_n\}_{n \in \mathbb{N}} \subseteq \mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$ . By Helly's Theorem (see, e.g., (Carothers, 2000, Lemma 13.15)), there exists  $\{v_{n_k}\}_{k \in \mathbb{N}} \subseteq \{v_n\}_{n \in \mathbb{N}}$  and  $\bar{v} \in \mathbb{R}^{[w, b]}$  such that  $v_{n_k}(x) \rightarrow \bar{v}(x)$  for all  $x \in [w, b]$  and  $\bar{v}$  is increasing. It is immediate to see that  $\bar{v}$  is such that  $\bar{v}(w) = 0 = \bar{v}(b) - 1$ . We are left to show that  $\bar{v} \in \mathcal{W}_{\max - \text{nor}}$ , that is,  $\bar{v}$  is continuous, strictly increasing, and such that  $p \succ' q$  implies  $\mathbb{E}_p(\bar{v}) \geq \mathbb{E}_q(\bar{v})$ . By Claims 1, 2, and 3,

it follows that  $\bar{v}$  satisfies (17), (19), and (21) with weak inequalities. This implies that  $\bar{v}$  is continuous at each point of  $[w, b]$ . We next show that  $\bar{v}$  is strictly increasing. We argue by contradiction. Assume that  $\bar{v}$  is not strictly increasing. Since  $\bar{v}$  is increasing, there exist  $x, y \in [w, b]$  such that  $y > x$  and  $\bar{v}(y) = \bar{v}(x)$ . By Claim 4 and since  $y > x \geq w$ , we have that  $\bar{v}(y), v_{n_k}(y) \geq \inf_{v \in \mathcal{W}} v(y) > 0$  for all  $k \in \mathbb{N}$ . Define  $\{\lambda_k\}_{k \in \mathbb{N}} \subseteq [0, 1]$  by

$$\lambda_k = \frac{v_{n_k}(x)}{v_{n_k}(y)} \quad \forall k \in \mathbb{N}.$$

Define also  $p_k = \lambda_k \delta_y + (1 - \lambda_k) \delta_w$  for all  $k \in \mathbb{N}$ . Since  $\lim_k \lambda_k = \frac{\bar{v}(x)}{\bar{v}(y)} = 1$ , it is immediate to see that  $p_k \rightarrow \delta_y$  and that

$$\mathbb{E}_{p_k}(v_{n_k}) = \lambda_k v_{n_k}(y) + (1 - \lambda_k) v_{n_k}(w) = v_{n_k}(x) \quad \forall k \in \mathbb{N}.$$

Thus, we also have that

$$V(p_k) \leq v_{n_k}^{-1}(\mathbb{E}_{p_k}(v_{n_k})) \leq x \quad \forall k \in \mathbb{N}.$$

Since  $V$  is continuous and by passing to the limit, we obtain that  $x < y = V(\delta_y) = \lim_k V(p_k) \leq x$ , a contradiction. Finally, assume that  $p \not\approx' q$ . Since  $\{v_{n_k}\}_{k \in \mathbb{N}} \subseteq \mathcal{W}$ , it follows that  $\mathbb{E}_p(v_{n_k}) \geq \mathbb{E}_q(v_{n_k})$  for all  $k \in \mathbb{N}$ . By the Lebesgue Dominated Convergence Theorem and since  $\{v_{n_k}\}_{k \in \mathbb{N}}$  is a uniformly bounded sequence which converges pointwise to  $\bar{v}$ , it follows that  $\mathbb{E}_p(\bar{v}) \geq \mathbb{E}_q(\bar{v})$ , proving that  $\bar{v} \in \mathcal{W}_{\max - \text{nor}}$ . ■

We are ready to characterize the compactness of the representing set  $\mathcal{W}$ . In a nutshell, the next result shows that the class of Cautious Expected Utility preferences that admit a compact representation is the subset that further satisfies NIRL.

**Theorem 6** *Let  $\succsim$  be a binary relation on  $\Delta$ . The following statements are equivalent:*

- (i)  $\succsim$  satisfies Weak Order, Continuity, Strict First Order Stochastic Dominance, Negative Certainty Independence, and NIRL;
- (ii) There exists a compact set  $\mathcal{W}$  in  $\mathcal{U}_{\text{nor}}$  such that

$$p \succsim q \iff \min_{v \in \mathcal{W}} c(p, v) \geq \min_{v \in \mathcal{W}} c(q, v).$$

*In particular,  $\mathcal{W}$  can be chosen to be  $\mathcal{W}_{\max - \text{nor}}$  and the latter is compact.*

**Proof.** Before starting, we add an intermediate point:

(iii)  $\succsim$  satisfies Weak Order, Continuity, Strict First Order Stochastic Dominance, Negative Certainty Independence, and Sensitivity.

(i) implies (iii). It follows from point (b) of Proposition 5.

(iii) implies (ii). By Theorem 2 and Remark 2, we have that the set  $\mathcal{W}_{\max-\text{nor}} \subseteq \mathcal{U}_{\text{nor}}$  is such that the function  $V : \Delta \rightarrow \mathbb{R}$ , defined by

$$V(p) = \inf_{v \in \mathcal{W}_{\max-\text{nor}}} c(p, v) \quad \forall p \in \Delta, \quad (22)$$

is a continuous utility representation of  $\succcurlyeq$ . By Theorem 5 and since  $\succcurlyeq$  satisfies Sensitivity, it follows that  $\mathcal{W}_{\max-\text{nor}}$  is sequentially compact with respect to the topology of pointwise convergence. By Lemma 1, this implies that  $\mathcal{W}_{\max-\text{nor}}$  is also compact with respect to the topology induced by the supnorm. We can conclude that the inf in (22) is attained and so the statement follows.

(ii) implies (i). Consider  $V : \Delta \rightarrow \mathbb{R}$  defined by

$$V(p) = \min_{v \in \mathcal{W}} c(p, v) \quad \forall p \in \Delta.$$

By hypothesis,  $V$  is well defined and it represents  $\succcurlyeq$ . Since  $\mathcal{W}$  is compact, we have that  $V$  is continuous. By (Cerrea-Vioglio et al., 2015, Theorem 1), this implies that  $\succcurlyeq$  satisfies Weak Order, Continuity, and Negative Certainty Independence. Next, consider  $p, q \in \Delta$  such that  $p \succ_{FSD} q$ . Consider also  $v \in \mathcal{W}$  such that  $V(p) = c(p, v)$ . Since  $v$  is strictly increasing, we have that  $V(p) = v^{-1}(\mathbb{E}_p(v)) > v^{-1}(\mathbb{E}_q(v)) \geq V(q)$ , proving that  $\succcurlyeq$  satisfies Strict First Order Stochastic Dominance. We are left to show that  $\succcurlyeq$  satisfies NIRL. Define  $u : [w, b] \rightarrow [0, 1]$  by  $u(x) = \min_{v \in \mathcal{W}} v(x)$ . Since  $\mathcal{W}$  is compact, it is immediate to verify that  $u \in \mathcal{U}_{\text{nor}}$ . By Proposition 6 and Remark 4,  $\succcurlyeq$  satisfies NIRL. ■

## C Betweenness

**Proof of Theorem 1.** Compared to (Dekel, 1986, Proposition 2), we only need to prove that the following form of Betweenness holds

$$p \succcurlyeq q \implies p \succcurlyeq \lambda p + (1 - \lambda) q \succcurlyeq q \quad \forall \lambda \in (0, 1)$$

and

$$p \succ q \implies p \succ \lambda p + (1 - \lambda) q \succ q \quad \forall \lambda \in (0, 1).$$

The proof of the first implication is routine.<sup>16</sup> As for the second, suppose  $p \succ q$ . By the first implication, we have that  $p \succcurlyeq \lambda p + (1 - \lambda) q \succcurlyeq q$  for all  $\lambda \in (0, 1)$ . By contradiction, assume that there exists  $\bar{\lambda} \in (0, 1)$  such that  $p \sim \bar{\lambda} p + (1 - \bar{\lambda}) q$ . We have two cases:

---

<sup>16</sup>For example, it can be proved by using the techniques of (Cerrea-Vioglio et al., 2011, Lemma 56).

1.  $p = \delta_b$ . Since  $p \succ q$ , we have that  $\delta_b = p \neq q$ , yielding that  $p \succ_{FSD} \bar{\lambda}p + (1 - \bar{\lambda})q$ . Since  $\succ$  satisfies Strict First Order Stochastic Dominance, we can conclude that  $p \succ \bar{\lambda}p + (1 - \bar{\lambda})q$ , a contradiction.
2.  $p \neq \delta_b$ . Since  $\succ$  satisfies Betweenness, we have that

$$1 \geq \lambda \geq \bar{\lambda} \Rightarrow \lambda p + (1 - \lambda)q \sim p. \quad (23)$$

Since  $\succ$  satisfies Strict First Order Stochastic Dominance, we have that  $\gamma p + (1 - \gamma)\delta_b \succ p$  for all  $\gamma \in (0, 1)$ . By (23) and since  $\succ$  satisfies Strict First Order Stochastic Dominance, we also have that

$$1 \geq \lambda \geq \bar{\lambda} \implies \lambda(\gamma p + (1 - \gamma)\delta_b) + (1 - \lambda)q \succ p \quad \forall \gamma \in (0, 1).$$

Next, we are going to define an ancillary object  $r_{\eta, \gamma} = \eta(\gamma p + (1 - \gamma)\delta_b) + (1 - \eta)q$  for all  $\eta, \gamma \in (0, 1)$ . Note that for each  $\eta, \gamma \in (0, 1)$  and for each  $\lambda \in (\bar{\lambda}, 1) \subseteq (0, 1)$ , we have that

$$\begin{aligned} \lambda p + (1 - \lambda)r_{\eta, \gamma} &= \\ &= \lambda p + (1 - \lambda)[\eta(\gamma p + (1 - \gamma)\delta_b) + (1 - \eta)q] \\ &= (\lambda + (1 - \lambda)\eta\gamma)p \\ &\quad + (1 - \lambda - (1 - \lambda)\eta\gamma) \left[ \frac{(1 - \lambda)\eta(1 - \gamma)}{(1 - \lambda - (1 - \lambda)\eta\gamma)}\delta_b + \frac{(1 - \lambda)(1 - \eta)}{(1 - \lambda - (1 - \lambda)\eta\gamma)}q \right]. \end{aligned}$$

Since  $\gamma p + (1 - \gamma)\delta_b \succ p \succ q$  for all  $\gamma \in (0, 1)$  and  $\succ$  satisfies Continuity, for each  $\gamma \in (0, 1)$  there exists  $\bar{\eta}_\gamma \in (0, 1)$  such that  $r_{\bar{\eta}_\gamma, \gamma} = \bar{\eta}_\gamma(\gamma p + (1 - \gamma)\delta_b) + (1 - \bar{\eta}_\gamma)q \sim p$ . Since  $\succ$  satisfies Betweenness,  $\lambda p + (1 - \lambda)r_{\bar{\eta}_\gamma, \gamma} \sim p$  for all  $\lambda \in (\bar{\lambda}, 1)$  and for all  $\gamma \in (0, 1)$ . Fix a generic  $\gamma \in (0, 1)$ . Choose  $\lambda \in (\bar{\lambda}, 1)$  close enough to 1, so that  $\hat{\lambda} = \lambda + (1 - \lambda)\bar{\eta}_\gamma \in (\bar{\lambda}, 1)$ . Note that

$$\hat{r} \stackrel{def}{=} \frac{(1 - \lambda)\bar{\eta}_\gamma(1 - \gamma)}{(1 - \lambda - (1 - \lambda)\bar{\eta}_\gamma\gamma)}\delta_b + \frac{(1 - \lambda)(1 - \bar{\eta}_\gamma)}{(1 - \lambda - (1 - \lambda)\bar{\eta}_\gamma\gamma)}q \succ_{FSD} q.$$

By the characterization of  $\lambda p + (1 - \lambda)r_{\bar{\eta}_\gamma, \gamma}$ , we can also conclude that

$$(\lambda + (1 - \lambda)\bar{\eta}_\gamma\gamma)p + (1 - \lambda - (1 - \lambda)\bar{\eta}_\gamma\gamma)\hat{r} \sim p. \quad (24)$$

By (23) and (24), we can conclude that  $\hat{\lambda} \in (\bar{\lambda}, 1)$ ,

$$\hat{\lambda}p + (1 - \hat{\lambda})\hat{r} \sim p \sim \hat{\lambda}p + (1 - \hat{\lambda})q \text{ and } \hat{\lambda}p + (1 - \hat{\lambda})\hat{r} \succ_{FSD} \hat{\lambda}p + (1 - \hat{\lambda})q.$$

Since  $\succ$  satisfies Strict First Order Stochastic Dominance, it follows that  $\hat{\lambda}p + (1 - \hat{\lambda})\hat{r} \succ \hat{\lambda}p + (1 - \hat{\lambda})q$ , a contradiction. A similar proof yields that  $\lambda p + (1 - \lambda)q \succ q$  for all  $\lambda \in (0, 1)$ .  $\blacksquare$

We next prove a few results pertaining to the expected utility core of a Betweenness preference. We start with a definition and an observation. Define  $K : \Delta \times [0, 1] \rightarrow \mathbb{R}$  by

$$K(r, t) = \int_{[w, b]} k(x, t) dr \quad \forall r \in \Delta, \forall t \in [0, 1].$$

It is immediate to see that  $K$  is affine wrt the first component. Note that for each  $r \in \Delta$  and for each  $t \in [0, 1]$

$$\begin{aligned} K\left(r, \hat{V}(r)\right) &= \int_{[w, b]} k\left(x, \hat{V}(r)\right) dr = \hat{V}(r) = \hat{V}(r)k(b, t) + \left(1 - \hat{V}(r)\right)k(w, t) \\ &= \int_{[w, b]} k(x, t) d\left(\hat{V}(r)\delta_b + \left(1 - \hat{V}(r)\right)\delta_w\right) \\ &= K\left(\hat{V}(r)\delta_b + \left(1 - \hat{V}(r)\right)\delta_w, t\right). \end{aligned}$$

Finally, we have that for each  $p \in \Delta$  the number  $\hat{V}(p) \in [0, 1]$  is the unique number such that

$$\hat{V}(p) = K\left(p, \hat{V}(p)\right).$$

**Proposition 7** *Let  $\succsim$  be a Betweenness preference. If  $K(p, t) \geq K(q, t)$  for all  $t \in (0, 1)$ , then  $p \succsim q$ .*

**Proof.** Consider  $p, q \in \Delta$ . By contradiction, assume that  $K(p, t) \geq K(q, t)$  for all  $t \in (0, 1)$  and  $q \succ p$ . We have two cases: either  $q = \delta_b$  or  $q \neq \delta_b$ . In the first case, note that  $1 \geq K(p, t) \geq K(q, t) = 1$  for all  $t \in (0, 1)$ , that is,  $K(p, t) = 1$  for all  $t \in (0, 1)$ . Since each  $k(\cdot, t)$  is strictly increasing and normalized, we have that  $p = \delta_b = q$ , a contradiction with  $q \succ p$ . In the second case, we have that  $\hat{V}(q) \in (0, 1)$ . On the one hand, since  $\succsim$  admits a representation a la Dekel, note that

$$\hat{V}(q) = K\left(q, \hat{V}(q)\right) \leq K\left(p, \hat{V}(q)\right). \quad (25)$$

On the other hand, by working hypothesis, we have  $q \succ p$  which implies that  $\hat{V}(q) > \hat{V}(p)$ . It follows that

$$\begin{aligned} \hat{V}(q) > \hat{V}(p) &= K\left(\hat{V}(p)\delta_b + \left(1 - \hat{V}(p)\right)\delta_w, \hat{V}(q)\right) \\ &= K\left(\hat{V}(p)\delta_b + \left(1 - \hat{V}(p)\right)\delta_w, \hat{V}(p)\right) = \hat{V}(p) = K\left(p, \hat{V}(p)\right). \end{aligned}$$

In particular, we have that

$$K\left(\hat{V}(p)\delta_b + \left(1 - \hat{V}(p)\right)\delta_w, \hat{V}(p)\right) = \hat{V}(p) = K\left(p, \hat{V}(p)\right) \quad (26)$$

and

$$\hat{V}(q) > K\left(\hat{V}(p)\delta_b + (1 - \hat{V}(p))\delta_w, \hat{V}(q)\right). \quad (27)$$

Define  $r = \hat{V}(p)\delta_b + (1 - \hat{V}(p))\delta_w$ . By (25) and (27) and since  $K$  is affine wrt the first component, it follows that there exists  $\lambda \in (0, 1]$  such that

$$K\left(\lambda p + (1 - \lambda)r, \hat{V}(q)\right) = \hat{V}(q),$$

proving that  $\lambda p + (1 - \lambda)r \sim q$ . By (26) and since  $\succcurlyeq$  is a Betweenness preference, we have that  $r \sim p$ , which yields that  $p \sim \lambda p + (1 - \lambda)r \sim r$ . We can conclude that  $q \succ p \sim \lambda p + (1 - \lambda)r \sim q$ , a contradiction.  $\blacksquare$

**Proposition 8** *Let  $\succcurlyeq$  be a Betweenness preference. If  $p \succ' q$ , then  $K(p, t) \geq K(q, t)$  for all  $t \in (0, 1)$ .*

**Proof.** Consider  $p, q \in \Delta$ . By contradiction, assume that  $p \succ' q$  and that there exists  $\bar{t} \in (0, 1)$  such that  $K(p, \bar{t}) < K(q, \bar{t})$ . Then, there exist  $\lambda \in (0, 1]$  and  $y \in [w, b]$  such that  $\hat{V}(\lambda p + (1 - \lambda)\delta_y) = \bar{t}$ .<sup>17</sup> It follows that

$$\begin{aligned} \bar{t} &= K(\lambda p + (1 - \lambda)\delta_y, \bar{t}) = \lambda K(p, \bar{t}) + (1 - \lambda)K(\delta_y, \bar{t}) \\ &< \lambda K(q, \bar{t}) + (1 - \lambda)K(\delta_y, \bar{t}) = K(\lambda q + (1 - \lambda)\delta_y, \bar{t}). \end{aligned}$$

Define  $r_1 = \lambda p + (1 - \lambda)\delta_y$  and  $r_2 = \lambda q + (1 - \lambda)\delta_y$  so that  $\bar{t} = \hat{V}(r_1)$ . In particular, we have that

$$\hat{V}(r_1) < K\left(r_2, \hat{V}(r_1)\right). \quad (28)$$

Since  $p \succ' q$  and  $\succ'$  satisfies Independence, it follows that  $r_1 \succ' r_2$ . Since  $\succ'$  is a subrelation of  $\succcurlyeq$ , this implies that  $r_1 \succcurlyeq r_2$ , that is,  $\hat{V}(r_1) \geq \hat{V}(r_2)$ . Define  $r_3 = \hat{V}(r_2)\delta_b + (1 - \hat{V}(r_2))\delta_w$ . On the one hand, it is immediate to see that  $r_2 \sim r_3$ . On the other hand, by (28), we have that

$$K\left(r_3, \hat{V}(r_1)\right) = \hat{V}(r_2) \leq \hat{V}(r_1) < K\left(r_2, \hat{V}(r_1)\right).$$

Since  $K$  is affine wrt the first component, there exists  $\gamma \in [0, 1)$  such that

$$K\left(\gamma r_2 + (1 - \gamma)r_3, \hat{V}(r_1)\right) = \hat{V}(r_1),$$

yielding that  $\gamma r_2 + (1 - \gamma)r_3 \sim r_1$ . Since  $\succcurlyeq$  satisfies Betweenness and  $r_2 \sim r_3$ , this yields that

$$r_2 \sim \gamma r_2 + (1 - \gamma)r_3 \sim r_1.$$

We can then conclude that  $\hat{V}(r_2) = \hat{V}(r_1)$ , that is,  $\hat{V}(r_1) = \hat{V}(r_2) = K\left(r_2, \hat{V}(r_2)\right) = K\left(r_2, \hat{V}(r_1)\right)$ , a contradiction with (28).  $\blacksquare$

---

<sup>17</sup>If  $\hat{V}(p) \geq \bar{t} > 0 = \hat{V}(\delta_w)$ , then  $y = w$  and if  $\hat{V}(p) < \bar{t} < 1 = \hat{V}(\delta_b)$ , then  $y = b$ . The existence of  $\lambda$  is then granted by the continuity of  $\hat{V}$ .



**Proposition 9** *If  $\succsim$  is a Betweenness preference, then*

$$p \succsim' q \iff \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \quad \forall v \in \mathcal{W}_{\text{bet}}.$$

*Moreover, the set  $\mathcal{W}_{\text{bet}}$  is either a singleton or infinite.*

**Proof.** Define  $\succsim''$  by

$$p \succsim'' q \iff \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \quad \forall v \in \mathcal{W}_{\text{bet}}.$$

By Proposition 8, we have that if  $p \succsim' q$ , then  $K(p, t) \geq K(q, t)$  for all  $t \in (0, 1)$ , that is,  $p \succsim'' q$ . By Proposition 7, if  $p \succsim'' q$ , that is  $K(p, t) \geq K(q, t)$  for all  $t \in (0, 1)$ , then  $p \succsim q$ . By (Cerreia-Vioglio et al., 2017, Lemma 1 and Footnote 10), we can conclude that  $p \succsim'' q$  implies  $p \succsim' q$ , proving that  $\succsim''$  coincides with  $\succsim'$ . Finally, assume that  $\mathcal{W}_{\text{bet}}$  is not a singleton. It follows that there exist  $t_1, t_2 \in (0, 1)$  and  $\bar{x} \in (w, b)$  such that  $k(\bar{x}, t_1) \neq k(\bar{x}, t_2)$ . Wlog, assume that  $k(\bar{x}, t_1) < k(\bar{x}, t_2)$ . By contradiction, assume that  $|\mathcal{W}_{\text{bet}}| \in \mathbb{N}$ . By the intermediate value theorem and since  $k(\bar{x}, \cdot)$  is continuous on  $(0, 1)$ , it follows that

$$\{k(\bar{x}, t)\}_{t \in (0, 1)} \supseteq [k(\bar{x}, t_1), k(\bar{x}, t_2)].$$

Since  $k(\bar{x}, t_1) < k(\bar{x}, t_2)$ , it follows that  $|\{k(\bar{x}, t)\}_{t \in (0, 1)}| = \infty$ , a contradiction with  $|\{k(\bar{x}, t)\}_{t \in (0, 1)}| \leq |\mathcal{W}_{\text{bet}}| \in \mathbb{N}$ .  $\blacksquare$

We now prove Theorem 4 and Proposition 2.

**Proof of Theorem 4.** (ii) implies (i). By (Cerreia-Vioglio et al., 2015, Theorem 1), the statement trivially follows.

(i) implies (ii). Since  $\succsim$  is a Betweenness preference, it satisfies Weak Order, Continuity, and Strict First Order Stochastic Dominance. By Theorem 2 and Remark 2, and since  $\mathcal{W}_{\text{bet}} = \{k(\cdot, t)\}_{t \in (0, 1)}$  represents  $\succsim'$ , it follows that  $\mathcal{W}$  in (6) can be chosen to be  $\mathcal{W}_{\text{bet}}$ . This yields (6) and, in particular, (9) with inf in place of min. Note that for each  $v \in \mathcal{W}_{\text{bet}}$  we have that  $V(\delta_w) = w = c(\delta_w, v)$  and  $V(\delta_b) = b = c(\delta_b, v)$ . Thus the inf is attained for  $\delta_w$  and  $\delta_b$ . The proof below yields that the inf is attained at each  $p \in \Delta$ , proving (9).

We next prove (10). Consider  $p \in \Delta \setminus \{\delta_w, \delta_b\}$ . Since  $\succsim$  satisfies Strict First Order Stochastic Dominance, we have that  $\hat{V}(p) \in (0, 1)$  and it is the unique number in  $[0, 1]$  such that

$$\int_{[w, b]} k(x, \hat{V}(p)) dp = \hat{V}(p). \quad (29)$$

Define  $v_p = k(\cdot, \hat{V}(p)) \in \mathcal{W}_{\text{bet}}$ . Define  $\bar{x} \in [w, b]$  to be such that  $\bar{x} = c(p, v_p)$ . Note that

$$v_p(\bar{x}) = v_p(c(p, v_p)) = v_p\left(v_p^{-1}\left(\int_{[w, b]} k(x, \hat{V}(p)) dp\right)\right) = \int_{[w, b]} k(x, \hat{V}(p)) dp.$$

By (29), it follows that

$$\int_{[w,b]} k(x, \hat{V}(p)) d\delta_{\bar{x}} = v_p(\bar{x}) = \hat{V}(p).$$

Since  $\succsim$  is a Betweenness preference, we have that  $\hat{V}(\delta_{\bar{x}}) = \hat{V}(p)$ , that is,  $\delta_{\bar{x}} \sim p$  and so  $\bar{x} = x_p$ . This yields that

$$V(p) = x_p = \bar{x} = c(p, v_p),$$

proving that the inf is attained at  $v_p$ . ■

**Proof of Proposition 2.** Before starting, define  $V : \Delta \rightarrow \mathbb{R}$  by

$$V(p) = \inf_{v \in \mathcal{W}_{\text{bet}}} c(p, v) \quad \forall p \in \Delta.$$

Define  $v_t = k(\cdot, t)$  for all  $t \in [0, 1]$ . Recall that  $\Delta_0$  is the subset of all simple lotteries, that is,  $\Delta_0 = \text{co}(\{\delta_x\}_{x \in [w,b]})$ .

*Claim:* If  $s, t \in (0, 1)$  and  $f_{s,t}$  is convex at  $t$ , then for each  $p \in \Delta_0$

$$\mathbb{E}_p(v_t) = t \implies c(p, v_t) \leq c(p, v_s).$$

*Proof of the Claim.* Let  $p \in \Delta_0$  and  $\mathbb{E}_p(v_t) = t$ . If  $p = \delta_x$ , then the statement is trivially true, since  $c(p, v_s) = x = c(p, v_t)$ . Otherwise, we have that there exist  $n \in \mathbb{N} \setminus \{1\}$ ,  $\{x_i\}_{i=1}^n \subseteq [w, b]$ , and  $\{\lambda_i\}_{i=1}^n \subseteq [0, 1]$  such that  $\sum_{i=1}^n \lambda_i = 1$  and  $\sum_{i=1}^n \lambda_i \delta_{x_i} = p$ . Define  $t_i = v_t(x_i) \in [0, 1]$  for all  $i \in \{1, \dots, n\}$ . Since  $\mathbb{E}_p(v_t) = t$ , this implies that  $\sum_{i=1}^n \lambda_i t_i = \sum_{i=1}^n \lambda_i v_t(x_i) = \mathbb{E}_p(v_t) = t$ . Since  $f_{s,t}$  is convex at  $t$ , we have that

$$f_{s,t}(\mathbb{E}_p(v_t)) = f_{s,t}(t) \leq \sum_{i=1}^n \lambda_i f_{s,t}(t_i) = \sum_{i=1}^n \lambda_i f_{s,t}(v_t(x_i)) = \sum_{i=1}^n \lambda_i v_s(x_i) = \mathbb{E}_p(v_s).$$

Since  $v_s = f_{s,t} \circ v_t$ , we have that  $f_{s,t} = v_s \circ v_t^{-1}$ . This implies that  $c(p, v_t) \leq c(p, v_s)$ .  $\square$

(i) implies (ii). Let  $p \in \Delta \setminus \{\delta_w, \delta_b\}$ . Since  $\succsim$  satisfies Strict First Order Stochastic Dominance, we have that  $\hat{V}(p) \in (0, 1)$  and it is the unique number in  $[0, 1]$  such that

$$\int_{[w,b]} k(x, \hat{V}(p)) dp = \hat{V}(p). \quad (30)$$

Define  $t = \hat{V}(p)$  and consider  $v_t$ . Let also  $s$  be an element of  $(0, 1)$  and consider  $v_s$  as well as  $f_{s,t}$ . Since  $\Delta_0$  is dense in  $\Delta$ , we have that there exists a sequence  $\{q_n\}_{n \in \mathbb{N}} \subseteq \Delta_0$  such that  $q_n \rightarrow p$ . Since  $\succsim$  satisfies Weak Order, we have that either  $\{n \in \mathbb{N} : q_n \succsim p\}$  is infinite or  $\{n \in \mathbb{N} : p \succsim q_n\}$  is infinite or both. We have two cases:

1.  $|\{n \in \mathbb{N} : q_n \succcurlyeq p\}| = \infty$ . It follows that there exists a subsequence  $\{q_{n_k}\}_{k \in \mathbb{N}}$  such that  $q_{n_k} \rightarrow p$  and  $q_{n_k} \succcurlyeq p$  for all  $k \in \mathbb{N}$ . Since  $\succcurlyeq$  satisfies Weak Order, Continuity, and Strict First Order Stochastic Dominance, it follows that for each  $k \in \mathbb{N}$  there exists  $\lambda_{n_k} \in [0, 1]$  such that  $p_{n_k} = \lambda_{n_k} q_{n_k} + (1 - \lambda_{n_k}) \delta_w \sim p$ . By (30) and since  $q_{n_k} \rightarrow p$ , we have that  $\mathbb{E}_{q_{n_k}}(v_t) \rightarrow \mathbb{E}_p(v_t) = t$ . By (30) and since  $p_{n_k} \sim p$ , we have that  $\mathbb{E}_{p_{n_k}}(v_t) = t$  for all  $k \in \mathbb{N}$ . This implies that

$$\begin{aligned} 0 &= \lim_k [\mathbb{E}_{p_{n_k}}(v_t) - \mathbb{E}_{q_{n_k}}(v_t)] = \lim_k (1 - \lambda_{n_k}) (v_t(w) - \mathbb{E}_{q_{n_k}}(v_t)) \\ &= (v_t(w) - t) \lim_k (1 - \lambda_{n_k}), \end{aligned}$$

proving that  $\lambda_{n_k} \rightarrow 1$ , since  $v_t(w) - t = 0 - t \neq 0$ . It follows that  $\{p_{n_k}\}_{k \in \mathbb{N}} \subseteq \Delta_0$ ,  $p_{n_k} \sim p$  for all  $k \in \mathbb{N}$ , and  $p_{n_k} \rightarrow p$ .

2.  $|\{n \in \mathbb{N} : p \succcurlyeq q_n\}| = \infty$ . It follows that there exists a subsequence  $\{q_{n_k}\}_{k \in \mathbb{N}}$  such that  $q_{n_k} \rightarrow p$  and  $p \succcurlyeq q_{n_k}$  for all  $k \in \mathbb{N}$ . Since  $\succcurlyeq$  satisfies Weak Order, Continuity, and Strict First Order Stochastic Dominance, it follows that for each  $k \in \mathbb{N}$  there exists  $\lambda_{n_k} \in [0, 1]$  such that  $p_{n_k} = \lambda_{n_k} q_{n_k} + (1 - \lambda_{n_k}) \delta_b \sim p$ . By (30) and since  $q_{n_k} \rightarrow p$ , we have that  $\mathbb{E}_{q_{n_k}}(v_t) \rightarrow \mathbb{E}_p(v_t) = t$ . By (30) and since  $p_{n_k} \sim p$ , we have that  $\mathbb{E}_{p_{n_k}}(v_t) = t$  for all  $k \in \mathbb{N}$ . This implies that

$$\begin{aligned} 0 &= \lim_k [\mathbb{E}_{p_{n_k}}(v_t) - \mathbb{E}_{q_{n_k}}(v_t)] = \lim_k (1 - \lambda_{n_k}) (v_t(b) - \mathbb{E}_{q_{n_k}}(v_t)) \\ &= (v_t(b) - t) \lim_k (1 - \lambda_{n_k}), \end{aligned}$$

proving that  $\lambda_{n_k} \rightarrow 1$ , since  $v_t(b) - t = 1 - t \neq 0$ . It follows that  $\{p_{n_k}\}_{k \in \mathbb{N}} \subseteq \Delta_0$ ,  $p_{n_k} \sim p$  for all  $k \in \mathbb{N}$ , and  $p_{n_k} \rightarrow p$ .

In both cases, it follows that there exists a sequence  $\{p_{n_k}\}_{k \in \mathbb{N}} \subseteq \Delta_0$  such that  $p_{n_k} \sim p$  for all  $k \in \mathbb{N}$  and  $p_{n_k} \rightarrow p$ . The condition  $p_{n_k} \sim p$  yields that  $\mathbb{E}_{p_{n_k}}(v_t) = t$  for all  $k \in \mathbb{N}$ . By the previous claim and since  $\{p_{n_k}\}_{k \in \mathbb{N}} \subseteq \Delta_0$  and  $f_{s,t}$  is convex at  $t$ , this implies that  $c(p_{n_k}, v_t) \leq c(p_{n_k}, v_s)$  for all  $k \in \mathbb{N}$ . By passing to the limit and since  $s \in (0, 1)$  was arbitrarily chosen, we obtain that

$$c(p, v_t) \leq c(p, v_s) \quad \forall s \in (0, 1). \quad (31)$$

We can conclude that

$$V(p) = \min_{s \in (0, 1)} c(p, v_s) = \min_{v \in \mathcal{W}_{\text{bet}}} c(p, v) = c(p, v_t).$$

By using the same technique in the proof of (i) implies (ii) in Theorem 4, we have that  $\bar{x} = c(p, v_t)$  is such that  $p \sim \delta_{\bar{x}}$ , that is,  $\bar{x} = x_p$ . Since  $p \in \Delta \setminus \{\delta_w, \delta_b\}$  was arbitrarily

chosen, we have that  $V(p) = x_p$  for all  $p \in \Delta$ .<sup>18</sup> This implies that  $V$  is a utility representation of  $\succsim$ . Since  $\succsim$  satisfies Continuity and  $V(\delta_x) = x$  for all  $x \in [w, b]$ , it is immediate to see that  $V$  is continuous. By Theorem 4, this implies that  $\succsim$  satisfies Negative Certainty Independence.

(ii) implies (i). By Theorem 4, we have that  $V : \Delta \rightarrow \mathbb{R}$ , defined by

$$V(p) = \min_{v \in \mathcal{W}_{\text{bet}}} c(p, v) = \min_{s \in (0,1)} c(p, v_s) \quad \forall p \in \Delta,$$

is a continuous utility representation of  $\succsim$ . By contradiction, assume that there exist  $t \in (0, 1)$  and  $s' \in (0, 1)$  such that  $f_{s',t}$  is not convex at  $t$ . It follows that there exist  $n \in \mathbb{N} \setminus \{1\}$ ,  $\{t_i\}_{i=1}^n \subseteq [0, 1]$ , and  $\{\lambda_i\}_{i=1}^n \subseteq [0, 1]$  such that  $\sum_{i=1}^n \lambda_i = 1$  and  $t = \sum_{i=1}^n \lambda_i t_i$  as well as  $f_{s',t}(t) > \sum_{i=1}^n \lambda_i f_{s',t}(t_i)$ . Consider  $\{x_i\}_{i=1}^n$  such that  $v_t(x_i) = t_i$ . Define  $p \in \Delta_0$  to be such that  $p = \sum_{i=1}^n \lambda_i \delta_{x_i}$ . It follows that  $\mathbb{E}_p(v_t) = \sum_{i=1}^n \lambda_i v_t(x_i) = \sum_{i=1}^n \lambda_i t_i = t$ . Since  $\succsim$  is a Betweenness preference, this implies that  $p \sim \delta_{\bar{x}}$  where  $\bar{x} = c(p, v_t)$ . In particular, we have that  $x_p = \bar{x}$ . At the same time, we also have that

$$f_{s',t}(\mathbb{E}_p(v_t)) = f_{s',t}(t) > \sum_{i=1}^n \lambda_i f_{s',t}(t_i) = \sum_{i=1}^n \lambda_i f_{s',t}(v_t(x_i)) = \sum_{i=1}^n \lambda_i v_{s'}(x_i) = \mathbb{E}_p(v_{s'}).$$

Since  $f_{s',t} = v_{s'} \circ v_t^{-1}$ , we can conclude that

$$\min_{s \in (0,1)} c(p, v_s) = V(p) = x_p = c(p, v_t) > c(p, v_{s'}) \geq \min_{s \in (0,1)} c(p, v_s),$$

a contradiction. ■

**Proof of Remark 1.** Denote  $f_{s,t}$  simply by  $f$ . Let  $t \in (0, 1)$ . Assume that  $f : [0, 1] \rightarrow [0, 1]$  is such that  $\partial f(t) \neq \emptyset$ . By assumption, we have that there exists  $m \in \mathbb{R}$  such that

$$f(t') - f(t) \geq m(t' - t) \quad \forall t' \in [0, 1].$$

Define  $g : [0, 1] \rightarrow \mathbb{R}$  by  $g(t') = mt' + l$  for all  $t' \in [0, 1]$  where  $l = f(t) - mt$ . Note that

$$f(t) = g(t) \text{ and } g(t') \leq f(t') \quad \forall t' \in [0, 1].$$

Next consider  $n \in \mathbb{N}$ ,  $\{t_i\}_{i=1}^n \subseteq [0, 1]$ , and  $\{\lambda_i\}_{i=1}^n \subseteq [0, 1]$  such that  $\sum_{i=1}^n \lambda_i = 1$  and  $\sum_{i=1}^n \lambda_i t_i = t$ . It follows that

$$f(t) = g(t) = g\left(\sum_{i=1}^n \lambda_i t_i\right) = \sum_{i=1}^n \lambda_i g(t_i) \leq \sum_{i=1}^n \lambda_i f(t_i),$$

proving convexity at  $t$ . ■

---

<sup>18</sup>Clearly,  $V(\delta_x) = x$  if either  $x = w$  or  $x = b$ .

**Proof of Theorem 3.** Before starting, we need to define few objects:  $\alpha : \text{Im } u \rightarrow \mathbb{R}$ ,  $\gamma : \text{Im } u \rightarrow \mathbb{R}$ ,  $g : [0, 1] \rightarrow \mathbb{R}$ ,  $\hat{V} : \Delta \rightarrow \mathbb{R}$  and  $k : [w, b] \times [0, 1] \rightarrow \mathbb{R}$ . We set

$$\alpha(s) = \frac{1}{\tilde{k}(b, s) - \tilde{k}(w, s)} \text{ and } \gamma(s) = \frac{-\tilde{k}(w, s)}{\tilde{k}(b, s) - \tilde{k}(w, s)} \quad \forall s \in \text{Im } u.$$

We also set

$$g(\lambda) = \tilde{V}(\lambda\delta_b + (1 - \lambda)\delta_w) \quad \forall \lambda \in [0, 1]$$

and, since  $g$  is strictly increasing, continuous, and  $\text{Im } g = \text{Im } u = \text{Im } \tilde{V}$ ,<sup>19</sup>

$$\hat{V}(p) = g^{-1}(\tilde{V}(p)) \quad \forall p \in \Delta.$$

Finally, we set

$$k(x, t) = \alpha(g(t))\tilde{k}(x, g(t)) + \gamma(g(t)) \quad \forall x \in [w, b], \forall t \in [0, 1].$$

It is easy to check that  $k$  and  $\hat{V}$  satisfy all the assumptions of Theorem 1.<sup>20</sup> Since

$$k(\cdot, t) = \alpha(g(t))\tilde{k}(\cdot, g(t)) + \gamma(g(t)) \quad \forall t \in [0, 1]$$

and  $g : [0, 1] \rightarrow \text{Im } u$  is strictly increasing, continuous, and onto, we have that for each  $t \in [0, 1]$  there exists an element  $z \in \text{Im } u$  such that  $k(\cdot, t)$  is a positive affine transformation of  $\tilde{k}(\cdot, z)$ . Similarly, for each  $z \in \text{Im } u$  there exists an element  $t \in [0, 1]$  such that  $\tilde{k}(\cdot, z)$  is a positive affine transformation of  $k(\cdot, t)$ . Recall that  $\mathcal{W}_{\text{da}} = \left\{ \tilde{k}(\cdot, z) \right\}_{z \in \text{Im } u}$ . It follows that  $\inf_{v \in \mathcal{W}_{\text{bet}}} c(p, v) = \min_{v \in \mathcal{W}_{\text{da}}} c(p, v)$  for all  $p \in \Delta$ .

1 and 2. We first show that if  $\beta \geq 0$ , then  $f_{s,t}$  is convex at  $t$  for all  $s, t \in (0, 1)$ . We do so by proving that  $\partial f_{s,t}(t) \neq \emptyset$  (see Remark 1). We thus need to compute  $v_s \circ v_t^{-1}$  where  $v_t(x) = k(x, t)$  and  $v_s(x) = k(x, s)$  for all  $x \in [w, b]$ . Fix  $s, t \in (0, 1)$ . Note that

$$v_t^{-1}(z) = \tilde{k}^{-1}\left(\frac{z - \gamma(g(t))}{\alpha(g(t))}, g(t)\right) \quad \forall z \in [0, 1].$$

Consider  $f : [0, 1] \rightarrow [0, 1]$  defined by  $f(z) = f_{s,t}(z) = \alpha(g(s))\tilde{k}\left(\tilde{k}^{-1}\left(\frac{z - \gamma(g(t))}{\alpha(g(t))}, g(t)\right), g(s)\right) + \gamma(g(s))$  for all  $z \in [0, 1]$ . Since  $s$  and  $t$  are fixed, to ease notation, define  $\alpha(g(s)) = \alpha$  and  $\gamma(g(s)) = \gamma$  as well as  $\alpha(g(t)) = \alpha'$  and  $\gamma(g(t)) = \gamma'$ . Note that  $\alpha, \alpha' > 0$ . Finally, set  $\tilde{v}_s = \tilde{k}(\cdot, g(s))$  and  $\tilde{v}_t = \tilde{k}(\cdot, g(t))$ . Therefore, we have that

$$f(z) = \alpha\tilde{v}_s\left(\tilde{v}_t^{-1}\left(\frac{z - \gamma'}{\alpha'}\right)\right) + \gamma \quad \forall z \in [0, 1].$$

<sup>19</sup>Indeed, one has that for each  $\lambda \in [0, 1]$

$$g(\lambda) = \tilde{V}(\lambda\delta_b + (1 - \lambda)\delta_w) = \frac{\lambda u(b) + (1 + \beta)(1 - \lambda)u(w)}{1 + \beta(1 - \lambda)}.$$

<sup>20</sup>Indeed, points 1 and 2 are satisfied on  $[0, 1]$  and not just  $(0, 1)$ .

Note that  $\partial f(t) \neq \emptyset$  if and only if there exists  $m \in \partial f(t)$ , that is,

$$f(t') - f(t) \geq m(t' - t) \quad \forall t' \in [0, 1].$$

Such an  $m$  exists if and only if there exists  $m \in \mathbb{R}$  such that

$$\begin{aligned} & \alpha \tilde{v}_s \left( \tilde{v}_t^{-1} \left( \frac{t' - \gamma'}{\alpha'} \right) \right) - \alpha \tilde{v}_s \left( \tilde{v}_t^{-1} \left( \frac{t - \gamma'}{\alpha'} \right) \right) \\ &= \alpha \tilde{v}_s \left( \tilde{v}_t^{-1} \left( \frac{t' - \gamma'}{\alpha'} \right) \right) + \gamma - \alpha \tilde{v}_s \left( \tilde{v}_t^{-1} \left( \frac{t - \gamma'}{\alpha'} \right) \right) - \gamma \\ &= f(t') - f(t) \geq m(t' - t) \quad \forall t' \in [0, 1], \end{aligned}$$

if and only if there exists  $m \in \mathbb{R}$  such that for each  $t' \in [0, 1]$

$$\tilde{v}_s \left( \tilde{v}_t^{-1} \left( \frac{t' - \gamma'}{\alpha'} \right) \right) - \tilde{v}_s \left( \tilde{v}_t^{-1} \left( \frac{t - \gamma'}{\alpha'} \right) \right) \geq \frac{m}{\alpha} (t' - t) = \left( \frac{m}{\alpha} \alpha' \right) \left( \frac{t' - \gamma'}{\alpha'} - \frac{t - \gamma'}{\alpha'} \right).$$

This latter condition holds if there exists  $\tilde{m} \in \partial \tilde{f} \left( \frac{t - \gamma'}{\alpha'} \right)$  where  $\tilde{f} = \tilde{v}_s \circ \tilde{v}_t^{-1}$ . Observe that: a)  $g(t) = \frac{t - \gamma'}{\alpha'}$  and b) for each  $z \in [0, 1]$

$$\tilde{k}(x, g(z)) = \begin{cases} u(x) & u(x) \leq g(z) \\ \frac{u(x) + \beta g(z)}{1 + \beta} & u(x) > g(z) \end{cases}.$$

This yields that

$$\tilde{k}^{-1}(x, g(z)) = \begin{cases} u^{-1}(x) & x \leq g(z) \\ u^{-1}(x(1 + \beta) - \beta g(z)) & x > g(z) \end{cases}$$

We now compute  $\tilde{f}(\cdot) = \tilde{v}_s \circ \tilde{v}_t^{-1}(\cdot) = \tilde{k}(\cdot, g(s)) \circ \tilde{k}^{-1}(\cdot, g(t))$ . If  $t = s$ , clearly  $\tilde{f}$  is the identity and therefore  $\partial \tilde{f}(g(t)) \neq \emptyset$ . Otherwise, there are two cases to consider:

1.  $t > s$ : In this case, we have that

$$\tilde{k} \left( \tilde{k}^{-1}(x, g(t)), g(s) \right) = \begin{cases} x & x \leq g(s) \\ \frac{x + \beta g(s)}{1 + \beta} & g(s) < x \leq g(t) \\ \frac{x(1 + \beta) + \beta[g(s) - g(t)]}{1 + \beta} & g(t) < x \end{cases}.$$

Note that

$$\frac{d}{dx} \left[ \tilde{k} \left( \tilde{k}^{-1}(x, g(t)), g(s) \right) \right] = \begin{cases} 1 & x < g(s) \\ \frac{1}{1 + \beta} & g(s) < x < g(t) \\ 1 & g(t) < x \end{cases}.$$

which clearly yields  $1 \in \partial \tilde{f}(g(t))$ , since  $\beta \geq 0$ .

2.  $t < s$ . In this case, we have that

$$\tilde{k}\left(\tilde{k}^{-1}(x, g(t)), g(s)\right) = \begin{cases} x & x \leq g(t) \\ x(1+\beta) - \beta g(t) & g(t) < x \leq \frac{g(s)+\beta g(t)}{1+\beta} \\ \frac{x(1+\beta)+\beta[g(s)-g(t)]}{1+\beta} & \frac{g(s)+\beta g(t)}{1+\beta} < x \end{cases}.$$

Note that

$$\frac{d}{dx} \left[ \tilde{k}\left(\tilde{k}^{-1}(x, g(t)), g(s)\right) \right] = \begin{cases} 1 & x < g(t) \\ 1 + \beta & g(t) < x < \frac{g(s)+\beta g(t)}{1+\beta} \\ 1 & \frac{g(s)+\beta g(t)}{1+\beta} < x \end{cases}.$$

which clearly yields  $1 \in \partial \tilde{f}(g(t))$ , since  $\beta \geq 0$ .

To sum up, we showed that if  $\beta \geq 0$ , then  $\partial \tilde{f}(g(t)) \neq \emptyset$  which yields that  $\partial f_{s,t}(t) = \partial f(t) \neq \emptyset$ . By Remark 1, it follows that  $f_{s,t}$  is convex at  $t$ . By Proposition 2 and since  $s, t \in (0, 1)$  were arbitrarily chosen, we have that  $\beta \geq 0$  implies that  $\succsim$  satisfies Negative Certainty Independence. By Theorem 4, it follows that if  $\beta \geq 0$ , then  $V : \Delta \rightarrow \mathbb{R}$ , defined by

$$V(p) = \min_{v \in \mathcal{W}_{\text{bet}}} c(p, v) = \min_{v \in \mathcal{W}_{\text{da}}} c(p, v) \quad \forall p \in \Delta$$

is a continuous utility representation of  $\succsim$  where  $\mathcal{W}_{\text{bet}} = \{k(\cdot, t)\}_{t \in (0,1)}$ . Thus, if  $\beta > 0$ , then (7) follows. If  $\beta = 0$ , then  $\mathcal{W}_{\text{da}} = \{u\}$  and (8) follows.

3. The statement follows from a specular argument.<sup>21</sup> ■

**Proof of Proposition 1.** “Only if”. By (Cerreia-Vioglio et al., 2017, Fact 2 and Lemma 1), if  $\succsim$  exhibits prudence, then any set representing  $\succsim'$  via an expected multi-utility representation must be made of functions which are differentiable on  $(w, b)$  and have convex derivatives there. By Proposition 9, we can conclude that this set can be chosen to be  $\mathcal{W}_{\text{bet}} = \{k(\cdot, t)\}_{t \in (0,1)}$ . By the discussion at the beginning of the proof of Theorem 3,  $\mathcal{W}_{\text{bet}}$  can be replaced by  $\mathcal{W}_{\text{da}} = \{\tilde{k}(\cdot, z)\}_{z \in \text{Im } u}$ . Since  $u$  is strictly increasing, the condition of differentiability of each of these local utilities forces  $u'$  to exist on  $(w, b)$  and be convex, as well as  $\beta$  to be equal to 0 and preferences to be expected utility. The “if” part is trivial. ■

---

<sup>21</sup>A sketch of a possible proof is as follows. Since  $\succsim$  satisfies Weak Order, Continuity, and Strict First Order Stochastic Dominance, there exists a continuous utility function  $V : \Delta \rightarrow \mathbb{R}$  such that  $V(\delta_x) = x$  for all  $x \in [w, b]$ . By Proposition 9 and the identification at the beginning of the proof of Theorem 3, the expected utility core  $\succsim'$  of  $\succsim$  admits an expected multi-utility representation with set  $\mathcal{W}_{\text{bet}}$ . After some tedious algebra, one can show that if  $\beta < 0$ , then  $\succsim$  satisfies Positive Certainty Independence. By mimicking Steps 4 and 5 in the proof of Theorem 1 of Cerreia-Vioglio et al. (2015), we obtain that  $V(p) = \sup_{v \in \mathcal{W}_{\text{bet}}} c(p, v) = \sup_{v \in \mathcal{W}_{\text{da}}} c(p, v)$  for all  $p \in \Delta$ . Since  $\mathcal{W}_{\text{da}}$  is compact, we have that the sup is actually attained.

**Proof of Example 1.** For each  $t \in [0, 1]$  define  $v_t(x) = k(x, t)$  for all  $x \in [0, 1]$ . Given  $s, t \in (0, 1)$ , we need to show that  $f = v_s \circ v_t^{-1}$  is convex at  $t$ . Before starting, observe that  $v_t^{-1} : [0, 1] \rightarrow \mathbb{R}$

$$v_t^{-1}(x) = \begin{cases} x & \text{if } x \leq t \\ \frac{t + \sqrt{t^2 + 4(x-t)}}{2} & \text{if } x > t \end{cases} \quad \forall x \in [0, 1].$$

Clearly, if  $s = t$ , then  $f = v_s \circ v_t^{-1}$  is the identity on  $[0, 1]$  and it is convex at  $t$ . We then have two cases:

1.  $t > s$ . In this case, we have that for each  $x \in [0, 1]$

$$f(x) = v_s(v_t^{-1}(x)) = \begin{cases} x & \text{if } x \leq s \\ x^2 - sx + s & \text{if } s < x \leq t \\ \left(\frac{t + \sqrt{t^2 + 4(x-t)}}{2}\right)^2 - s \left(\frac{t + \sqrt{t^2 + 4(x-t)}}{2}\right) + s & \text{if } x > t \end{cases}.$$

Consider  $g : [0, 1] \rightarrow \mathbb{R}$  to be such that  $g(x) = m(x-t) + f(t)$  and  $m = \max\{2t - s, 1\}$ . We have three cases:

(a)  $0 \leq t' \leq s$ . Note that

$$g(0) = f(t) - mt \leq f(t) - t = t^2 - st + s - t = (t-1)(t-s) < 0.$$

We can conclude that

$$\begin{aligned} g(t') &= m(t' - t) + f(t) \leq f(t) + t' - t \\ &= t' + f(t) - t \leq t' = f(t'). \end{aligned}$$

(b)  $s < t' \leq t$ . Define  $h : [0, 1] \rightarrow \mathbb{R}$  by  $h(x) = x^2 - sx + s$  for all  $x \in [0, 1]$ . Note that  $h(t) = f(t)$  and  $h'(t) = 2t - s \leq m$ , yielding  $h'(t)(t' - t) \geq m(t' - t)$  for all  $t' \leq t$ . Since  $h$  is convex, we have that

$$f(t') = h(t') \geq h'(t)(t' - t) + h(t) \geq m(t' - t) + f(t) = g(t') \quad \forall t' \in (s, t].$$

(c)  $t' > t$ . Define  $\tilde{h} : [t, 1] \rightarrow \mathbb{R}$  by  $\tilde{h}(x) = \left(\frac{t + \sqrt{t^2 + 4(x-t)}}{2}\right)^2 - s \left(\frac{t + \sqrt{t^2 + 4(x-t)}}{2}\right) + s$  for all  $x \in [t, 1]$ . It follows that  $\tilde{h}$  is concave. Note that  $\tilde{h}(t) = f(t) = g(t)$ . Since  $\tilde{h}$  is concave and  $g$  is affine, it is enough to verify that  $\tilde{h}(1) \geq g(1)$  to prove that  $f(t') = \tilde{h}(t') \geq g(t')$  for all  $t' \in [t, 1]$ . Since  $t \in (0, 1)$  and  $\tilde{h}(1) = 1$ , observe that if  $m = 2t - s$ , then

$$\begin{aligned} g(1) &= m(1-t) + f(t) = (2t-s)(1-t) + t^2 - st + s \\ &= 2t - 2t^2 - s + st + t^2 - st + s \\ &= 2t - t^2 = t + t(1-t) \leq 1 = \tilde{h}(1). \end{aligned}$$



Since  $0 < s < t < 1$ , if  $m = 1$ , then

$$\begin{aligned} g(1) - \tilde{h}(1) &= g(1) - 1 = 1 - t + f(t) - 1 = 1 - t + t^2 - st + s - 1 \\ &= -t + t^2 - st + s = t(t-1) + s(1-t) \\ &= (t-s)(t-1) < 0, \end{aligned}$$

proving that  $g(1) < \tilde{h}(1)$ .

Subpoints a–c just showed that the subdifferential of  $f$  is nonempty at  $t$  and, in particular,  $f$  is convex at  $t$ .

2.  $t < s$ . In this case, we have that for each  $x \in [0, 1]$

$$f(x) = v_s(v_t^{-1}(x)) = \begin{cases} x & \text{if } x \leq t \\ \frac{t + \sqrt{t^2 + 4(x-t)}}{2} & \text{if } t < x \leq \bar{s} \\ \left(\frac{t + \sqrt{t^2 + 4(x-t)}}{2}\right)^2 - s \left(\frac{t + \sqrt{t^2 + 4(x-t)}}{2}\right) + s & \text{if } x > \bar{s} \end{cases}$$

where  $\bar{s}$  is such that  $\frac{t + \sqrt{t^2 + 4(\bar{s}-t)}}{2} = s$ .<sup>22</sup> Consider  $g : [0, 1] \rightarrow \mathbb{R}$  to be such that  $g(x) = x$ . We have three cases:

(a)  $0 \leq t' \leq t$ . Clearly, we have that  $f(t') \geq g(t')$ .

(b)  $t < t' \leq \bar{s}$ . Define  $h : [t, \bar{s}] \rightarrow \mathbb{R}$  by  $h(x) = \frac{t + \sqrt{t^2 + 4(x-t)}}{2}$  for all  $x \in [t, \bar{s}]$ . Since  $h$  is concave and  $g$  is affine, if we verify that  $h(t) \geq g(t)$  and  $h(\bar{s}) \geq g(\bar{s})$ , then  $f(t') = h(t') \geq g(t')$  for all  $t' \in [t, \bar{s}]$ . Note that  $h(t) = t = g(t)$ . On the other hand, we have that

$$\begin{aligned} h(\bar{s}) &= \frac{t + \sqrt{t^2 + 4(\bar{s}-t)}}{2} \geq \frac{t + \sqrt{t^2 + 4\bar{s}(\bar{s}-t)}}{2} \\ &= \frac{t + \sqrt{t^2 + 4\bar{s}^2 - 4\bar{s}t}}{2} = \frac{t + \sqrt{(2\bar{s}-t)^2}}{2} = \bar{s} = g(\bar{s}). \end{aligned}$$

(c)  $t' > \bar{s}$ . Define  $\tilde{h} : [\bar{s}, 1] \rightarrow \mathbb{R}$  by  $\tilde{h}(x) = \left(\frac{t + \sqrt{t^2 + 4(x-t)}}{2}\right)^2 - s \left(\frac{t + \sqrt{t^2 + 4(x-t)}}{2}\right) + s$  for all  $x \in [\bar{s}, 1]$ . Since  $\tilde{h}$  is convex,  $\tilde{h}(1) = 1$ , and  $\tilde{h}'(1) = \frac{2-s}{2-t} \in (0, 1)$ , we have that

$$\begin{aligned} \tilde{h}(t') &\geq \tilde{h}'(1)(t'-1) + \tilde{h}(1) \geq 1(t'-1) + \tilde{h}(1) \\ &= t' - 1 + 1 = t' = g(t') \quad \forall t' \in [\bar{s}, 1]. \end{aligned}$$

---

<sup>22</sup>Since  $\frac{t + \sqrt{t^2 + 4(t-t)}}{2} = t < s < 1 = \frac{t + \sqrt{t^2 + 4(1-t)}}{2}$  and the map  $x \mapsto \frac{t + \sqrt{t^2 + 4(x-t)}}{2}$  is strictly increasing and continuous on  $[t, 1]$ , we have that  $\bar{s}$  exists and  $\bar{s} > t$ .

Subpoints a–c just showed that the subdifferential of  $f$  is nonempty at  $t$  and, in particular,  $f$  is convex at  $t$ . ■

**Proof of Proposition 4.** Define  $\mathcal{W} = \bar{c}\bar{o}(\mathcal{W}_{\text{bet}})$ . Given the assumptions,  $\mathcal{W}$  is convex and compact and (12) holds with  $\mathcal{W}$  in place of  $\mathcal{W}_{\text{bet}}$ .<sup>23</sup> First, note that the map  $c : \Delta \times \mathcal{W} \rightarrow [w, b]$ , defined by

$$c(p, v) = v^{-1}(\mathbb{E}_p(v)) \quad \forall (p, v) \in \Delta \times \mathcal{W},$$

is quasiconcave and upper semicontinuous in the first argument and quasiconvex and lower semicontinuous in the second argument. By Sion's minimax theorem and since  $A$  is a convex and compact set of  $\Delta$ , this implies that

$$\max_{p \in A} \min_{v \in \mathcal{W}} c(p, v) = \min_{v \in \mathcal{W}} \max_{p \in A} c(p, v).$$

Let  $\hat{v} \in \mathcal{W}$  be such that  $\max_{p \in A} c(p, \hat{v}) \leq \max_{p \in A} c(p, v)$  for all  $v \in \mathcal{W}$ . Note that

$$\begin{aligned} \max_{p \in A} \min_{v \in \mathcal{W}} c(p, v) &= \max_{p \in A} V(p) = V(\bar{p}) = \min_{v \in \mathcal{W}} c(\bar{p}, v) \leq c(\bar{p}, \hat{v}) \\ &\leq \max_{p \in A} c(p, \hat{v}) \leq \min_{v \in \mathcal{W}} \max_{p \in A} c(p, v). \end{aligned}$$

Since  $\max_{p \in A} \min_{v \in \mathcal{W}} c(p, v) = \min_{v \in \mathcal{W}} \max_{p \in A} c(p, v)$ , this yields that

$$c(\bar{p}, \hat{v}) = \max_{p \in A} c(p, \hat{v}),$$

proving the statement. ■

## C.1 Betweenness Preferences and Finite Representations

In this section, we further explore the intersection of Betweenness preferences and Cautious Expected Utility preferences which admit a finite representation.

We start by showing that *any canonical* representation of a Betweenness preference that satisfies Negative Certainty Independence must contain either one element or infinitely many. We prove this under an additional assumption: namely,  $\succsim$  is not infinitely risk loving, which we called NIRL.

**Proposition 10** *Let  $\succsim$  be a Betweenness preference that satisfies Negative Certainty Independence and NIRL. If  $\mathcal{W}' \subseteq \mathcal{U}_{\text{nor}}$  satisfies (6) and (13), then either  $|\mathcal{W}'| = 1$  or  $|\mathcal{W}'| = \infty$ .*

---

<sup>23</sup>Note that  $u : [w, b] \rightarrow [0, 1]$ , defined by  $u(x) = \min_{t \in [0, 1]} k(x, t)$  for all  $x \in [w, b]$ , belongs to  $\mathcal{U}_{\text{nor}}$ . In particular, we have that  $v \geq u$  for all  $v \in \mathcal{W}_{\text{bet}}$ . By Theorem 4 and Propositions 6 and 9, we have that  $\succsim$  satisfies NIRL. By Theorem 6, this implies that  $\mathcal{W}_{\text{max-nor}}$  is compact and satisfies (12). Finally, by Remark 2, we have that  $\bar{c}\bar{o}(\mathcal{W}_{\text{bet}}) = \text{cl}(\mathcal{W}_{\text{max-nor}}) = \mathcal{W}_{\text{max-nor}}$ .

**Proof.** Recall that  $\mathcal{W}_{\text{bet}} = \{k(\cdot, t)\}_{t \in (0,1)}$ . Define by  $\mathcal{E}$  the set of extreme points of  $\mathcal{W}_{\text{max-nor}}$ . By contradiction, assume that  $|\mathcal{W}'| = n \in \mathbb{N} \setminus \{1\}$ .

*Step 1: The set  $\mathcal{E}$  is a nonempty finite subset of  $\text{cl}(\mathcal{W}_{\text{bet}})$ , contains more than one element, and satisfies (6) and (13).*

*Proof of the Step.* By Theorems 2, 6, and 4, Proposition 9, and Remark 2, it follows that  $\mathcal{W}_{\text{max-nor}}$  is compact and therefore  $\bar{\text{co}}(\mathcal{W}') = \bar{\text{co}}(\mathcal{W}_{\text{bet}}) = \text{cl}(\mathcal{W}_{\text{max-nor}}) = \mathcal{W}_{\text{max-nor}}$ . By (Megginson, 1998, Theorem 2.10.6 and Corollary 2.10.16) and Proposition 9, it follows that  $\emptyset \neq \mathcal{E} \subseteq \text{cl}(\mathcal{W}_{\text{bet}})$  and  $\mathcal{E} \subseteq \text{cl}(\mathcal{W}') = \mathcal{W}'$ . It follows that  $\mathcal{E}$  is a nonempty finite subset of  $\text{cl}(\mathcal{W}_{\text{bet}})$ . Clearly, it contains more than one element. Otherwise, by Krein-Milman's theorem (see (Megginson, 1998, Theorem 2.10.6)),  $1 = |\bar{\text{co}}(\mathcal{E})| = |\bar{\text{co}}(\mathcal{W}')| \geq |\mathcal{W}'| > 1$ , a contradiction. Moreover, by Krein-Milman's theorem again, it is immediate to see that  $\mathcal{E}$  satisfies (13). By Remark 2, it follows that  $\mathcal{E}$  satisfies (6) as well.  $\square$

By Step 1, we have that  $\mathcal{E} \subseteq \text{cl}(\mathcal{W}_{\text{bet}})$ . In particular,  $\text{cl}(\mathcal{W}_{\text{bet}})$  contains more than one element. By Proposition 9, it follows that  $\mathcal{W}_{\text{bet}}$  cannot be a singleton and, therefore, it contains infinitely many elements. Moreover, by Theorem 4, we have that

$$V(p) = \min_{v \in \mathcal{E}} c(p, v) = \min_{v \in \mathcal{W}_{\text{bet}}} c(p, v) \quad \forall p \in \Delta. \quad (32)$$

By the Krein-Milman's Theorem and (Aliprantis and Border, 2006, Corollary 5.30) and since  $\mathcal{E}$  is finite, it follows that

$$\text{co}(\mathcal{E}) = \bar{\text{co}}(\mathcal{E}) = \mathcal{W}_{\text{max-nor}} \supseteq \mathcal{W}_{\text{bet}}.$$

Since  $\mathcal{E}$  is finite and not a singleton, there exists a finite collection  $\mathcal{E} = \{v_m\}_{m=1}^n \subseteq \mathcal{U}_{\text{nor}}$  with  $n \geq 2$ . Consider  $\bar{t} \in (0, 1)$  such that  $k(\cdot, \bar{t}) \neq v_m$  for all  $m \in \{1, \dots, n\}$ . Since  $\mathcal{W}_{\text{bet}}$  is infinite, such a  $\bar{t}$  exists. It follows that  $\bar{v} = k(\cdot, \bar{t}) = \sum_{m=1}^n \lambda_m v_m$  where  $\lambda_m \geq 0$  for all  $m \in \{1, \dots, n\}$  and  $\sum_{m=1}^n \lambda_m = 1$ . Moreover,  $\bar{v} \notin \mathcal{E}$ . Define  $I(\bar{t}) = \{p \in \Delta : \hat{V}(p) = \bar{t}\}$ . By Theorem 4, we have that

$$c(p, \bar{v}) = V(p) \quad \forall p \in I(\bar{t}). \quad (33)$$

*Step 2: There exists  $i \in \{1, \dots, n\}$  for each  $p \in I(\bar{t})$  such that  $V(p) = c(p, v_i)$ .*

*Proof of the Step.* By contradiction, assume that for each  $i \in \{1, \dots, n\}$  there exists  $p_i \in I(\bar{t})$  such that  $V(p_i) \neq c(p_i, v_i)$ . Since (32) holds, this implies that  $V(p_i) < c(p_i, v_i)$ . Fix  $i \in \{1, \dots, n\}$ . By (33) and since  $p_i \in I(\bar{t})$ , it follows that if  $\lambda_i > 0$ , then

$$V(p_i) = c(p_i, \bar{v}) > V(p_i),$$

a contradiction,<sup>24</sup> that is,  $\lambda_i = 0$ . Since  $i$  was arbitrarily chosen, then  $\lambda_i = 0$  for all  $i \in \{1, \dots, n\}$ , a contradiction with  $\sum_{m=1}^n \lambda_m = 1$ .  $\square$

---

<sup>24</sup>Note that

$$c(p_i, v_m) \geq V(p_i) \quad \forall m \in \{1, \dots, n\} \text{ and } c(p_i, v_i) > V(p_i),$$

By Step 2 and (33), it follows that there exists  $i \in \{1, \dots, n\}$  such that  $c(p, v_i) = V(p)$  and  $c(p, \bar{v}) = V(p)$  for all  $p \in I(\bar{t})$ , that is,

$$\mathbb{E}_p(v_i) = v_i(V(p)) \text{ and } \mathbb{E}_p(\bar{v}) = \bar{v}(V(p)) \quad \forall p \in I(\bar{t}). \quad (34)$$

In the next two steps, we show that  $\bar{v} = v_i$ . First, define  $\bar{p} = \bar{t}\delta_b + (1 - \bar{t})\delta_w$ . It is immediate to check that  $\bar{p} \in I(\bar{t})$ . Let then  $\bar{x}$  be such that  $\delta_{\bar{x}} \sim \bar{p}$ . Since  $\bar{t} \in (0, 1)$ , we have that  $\bar{x} \in (w, b)$ .

*Step 3:*  $v_i(\bar{x}) = \bar{v}(\bar{x})$ .

*Proof of the Step.* Since  $\succsim$  satisfies Negative Certainty Independence, we have that  $\delta_{\bar{x}} \sim \gamma\bar{p} + (1 - \gamma)\delta_{\bar{x}}$ , that is,  $\bar{x} = V(\delta_{\bar{x}}) = V(\gamma\bar{p} + (1 - \gamma)\delta_{\bar{x}})$  and  $\gamma\bar{p} + (1 - \gamma)\delta_{\bar{x}} \in I(\bar{t})$  for all  $\gamma \in [0, 1]$ . By (34), it follows that for each  $\gamma \in (0, 1)$

$$\mathbb{E}_{\gamma\bar{p} + (1 - \gamma)\delta_{\bar{x}}}(v_i) = v_i(V(\gamma\bar{p} + (1 - \gamma)\delta_{\bar{x}})) \text{ and } \mathbb{E}_{\gamma\bar{p} + (1 - \gamma)\delta_{\bar{x}}}(\bar{v}) = \bar{v}(V(\gamma\bar{p} + (1 - \gamma)\delta_{\bar{x}})),$$

that is,  $\gamma\bar{t} + (1 - \gamma)v_i(\bar{x}) = v_i(\bar{x})$  and  $\gamma\bar{t} + (1 - \gamma)\bar{v}(\bar{x}) = \bar{v}(\bar{x})$  for all  $\gamma \in (0, 1)$ . By choosing  $\gamma = \frac{1}{2}$  and subtracting the two equations, we can conclude that

$$\frac{1}{2}(v_i(\bar{x}) - \bar{v}(\bar{x})) = v_i(\bar{x}) - \bar{v}(\bar{x}),$$

that is,  $v_i(\bar{x}) = \bar{v}(\bar{x})$ . □

*Step 4:*  $v_i = \bar{v}$ .

*Proof of the Step.* By contradiction, assume that  $v_i \neq \bar{v}$ . Since  $v_i, \bar{v} \in \mathcal{U}_{\text{nor}}$ , this implies that there exists  $x \in (w, b)$  such that  $v_i(x) \neq \bar{v}(x)$ . By Step 3, we have two cases:

1.  $x > \bar{x}$ . There exists  $\gamma \in (0, 1)$  such that  $\gamma\delta_x + (1 - \gamma)\delta_w \sim \delta_{\bar{x}} \in I(\bar{t})$ . By (34) and Step 3, we have that

$$\gamma v_i(x) = \mathbb{E}_{\gamma\delta_x + (1 - \gamma)\delta_w}(v_i) = v_i(\bar{x}) = \bar{v}(\bar{x}) = \mathbb{E}_{\gamma\delta_x + (1 - \gamma)\delta_w}(\bar{v}) = \gamma\bar{v}(x),$$

that is,  $\gamma v_i(x) = \gamma\bar{v}(x)$ , a contradiction, since  $\gamma \neq 0$  and  $v_i(x) \neq \bar{v}(x)$ .

2.  $x < \bar{x}$ . There exists  $\gamma \in (0, 1)$  such that  $\gamma\delta_x + (1 - \gamma)\delta_b \sim \delta_{\bar{x}} \in I(\bar{t})$ . By (34) and Step 3, we have that

$$\gamma v_i(x) + 1 - \gamma = \mathbb{E}_{\gamma\delta_x + (1 - \gamma)\delta_b}(v_i) = v_i(\bar{x}) = \bar{v}(\bar{x}) = \mathbb{E}_{\gamma\delta_x + (1 - \gamma)\delta_b}(\bar{v}) = \gamma\bar{v}(x) + 1 - \gamma,$$

that is,  $\gamma v_i(x) = \gamma\bar{v}(x)$ , a contradiction, since  $\gamma \neq 0$  and  $v_i(x) \neq \bar{v}(x)$ . □

---

that is,

$$\mathbb{E}_{p_i}(v_m) \geq v_m(V(p_i)) \quad \forall m \in \{1, \dots, n\} \text{ and } \mathbb{E}_{p_i}(v_i) > v_i(V(p_i)).$$

If  $\lambda_i > 0$ , then

$$\mathbb{E}_{p_i}(\bar{v}) = \sum_{m=1}^n \lambda_m \mathbb{E}_{p_i}(v_m) > \sum_{m=1}^n \lambda_m v_m(V(p_i)) = \bar{v}(V(p_i)).$$

By Step 4, we can conclude that  $v_i = \bar{v}$ , a contradiction with  $\bar{v} \notin \mathcal{E}$ . ■

**Corollary 2** *Let  $\succsim$  be a Betweenness preference that satisfies Negative Certainty Independence and NIRL. If  $\mathcal{W}' \subseteq \mathcal{U}_{\text{nor}}$  satisfies (6) and (13), then  $\succsim$  violates Independence if and only if  $|\mathcal{W}'| = \infty$ .<sup>25</sup>*

We next turn to Cautious Expected Utility preferences that admit a finite representation, that is,  $|\mathcal{W}| < \infty$  where  $\mathcal{W}$  is only assumed to represent  $\succsim$ , but might fail to represent  $\succsim'$ . This was analyzed in Proposition 3. From an economic point of view, this is an important class. Indeed, one could speculate that in applications, finite specifications may be appealing. The proposition shows that this apparently harmless assumption comes with a behavioral counterpart: either expected utility, or violations of Betweenness. From a technical point of view, this result is the conceptual counterpositive of Proposition 10. Apart from few technical details, Proposition 10 says that, given a Cautious Expected Utility preference which is not expected utility, if it satisfies Betweenness, then any *canonical* representation must contain infinite elements. Thus, the counterpositive of this statement is, given a Cautious Expected Utility preference, if a *canonical* representation contains finitely many elements, then it violates Betweenness. Starting from a finite specification  $\mathcal{W}$  of the Cautious Expected Utility model, does not rule out that any canonical representation might contain infinitely many elements. Indeed, a priori, given a canonical representation  $\mathcal{W}'$ , we only know that  $\bar{\text{co}}(\mathcal{W}) \supseteq \bar{\text{co}}(\mathcal{W}') = \mathcal{W}_{\text{max-nor}}$ .<sup>26</sup> One could have a situation similar to the following one: Think of  $\mathcal{W}$  as a finite set of points (loosely speaking, say four vertices of a square, so that  $\text{co}(\mathcal{W}) = \bar{\text{co}}(\mathcal{W})$  is the square) and  $\mathcal{W}'$  as a circle fully contained in  $\text{co}(\mathcal{W})$ ; clearly  $\mathcal{W}$  is finite and  $\mathcal{W}'$  is not. Moreover,  $\mathcal{W}'$  cannot be fully refined to be finite, otherwise one would not be able to obtain  $\bar{\text{co}}(\mathcal{W}') = \mathcal{W}_{\text{max-nor}}$ . However, Proposition 3 shows that this cannot be the case.

**Proof of Proposition 3.** Consider  $\mathcal{W}$ . Since  $\mathcal{W}$  satisfies (6), recall that

$$V(p) = \min_{v \in \mathcal{W}} c(p, v) \quad \forall p \in \Delta.$$

Since  $\mathcal{W}$  is finite, it is compact. By Theorem 6, this implies that  $\mathcal{W}_{\text{max-nor}}$  is compact. Say that  $v \in \mathcal{W}$  is redundant in  $\mathcal{W}$  if and only if for each  $p \in \Delta \setminus \{\delta_x\}_{x \in [w, b]}$  there exists

---

<sup>25</sup>We remind the reader that Independence is the standard assumption:

$$p \succsim q \implies \lambda p + (1 - \lambda)r \succsim \lambda q + (1 - \lambda)r \quad \forall r \in \Delta, \forall \lambda \in [0, 1].$$

<sup>26</sup>For the sake of simplicity, we are assuming that  $\mathcal{W}'$  is compact as well.

$\bar{v} \in \mathcal{W} \setminus \{v\}$  such that  $c(p, v) \geq c(p, \bar{v})$ . Define

$$\mathcal{W}_1 = \begin{cases} \mathcal{W} \setminus \{v\} & \text{if } \exists v \in \mathcal{W} \text{ redundant in } \mathcal{W} \\ \mathcal{W} & \text{if } \nexists v \in \mathcal{W} \text{ redundant in } \mathcal{W} \end{cases}.$$

Note that in both cases

$$V(p) = \min_{v \in \mathcal{W}_1} c(p, v) \quad \forall p \in \Delta.$$

If  $\mathcal{W}_1 = \mathcal{W}$ , then we stop. Otherwise, we compute

$$\mathcal{W}_2 = \begin{cases} \mathcal{W}_1 \setminus \{v\} & \text{if } \exists v \in \mathcal{W}_1 \text{ redundant in } \mathcal{W}_1 \\ \mathcal{W}_1 & \text{if } \nexists v \in \mathcal{W}_1 \text{ redundant in } \mathcal{W}_1 \end{cases}.$$

Note that  $V(p) = \min_{v \in \mathcal{W}_2} c(p, v)$  for all  $p \in \Delta$ . By iterating this procedure and since  $\mathcal{W}$  is finite, we get to a set  $\mathcal{W}_k \subseteq \mathcal{W}$  with  $k \in \mathbb{N}$  such that  $\mathcal{W}_k = \mathcal{W}_{k+1}$  and  $V(p) = \min_{v \in \mathcal{W}_k} c(p, v)$  for all  $p \in \Delta$ . If  $|\mathcal{W}_k| = 1$ , then clearly  $\succsim$  is expected utility and satisfies Independence. If  $|\mathcal{W}_k| \in \mathbb{N} \setminus \{1\}$ , then we show that  $\mathcal{W}_k$  represents not only  $\succsim$ , but also  $\succsim'$ . Since  $\mathcal{W}_k$  is finite, we enumerate it as  $\{v_l\}_{l=1}^m$ . By (Cerrei-Vioglio et al., 2015, Theorem 2), note that  $\bar{\text{co}}(\mathcal{W}_k) \supseteq \mathcal{W}_{\max\text{-nor}}$ . By contradiction, assume that  $\bar{\text{co}}(\mathcal{W}_k) \supset \mathcal{W}_{\max\text{-nor}}$ . Since  $\mathcal{W}_{\max\text{-nor}}$  is convex and compact, we have that  $\mathcal{W}_k \not\subseteq \mathcal{W}_{\max\text{-nor}}$ .<sup>27</sup> Without loss of generality, say that  $v_1$  is the element in  $\mathcal{W}_k$  that does not belong to  $\mathcal{W}_{\max\text{-nor}}$ . Since  $\mathcal{W}_k = \mathcal{W}_{k+1}$ , we have that no element in  $\mathcal{W}_k$  is redundant in  $\mathcal{W}_k$ . In particular,  $v_1$  has this property. This implies that there exists  $\hat{p} \in \Delta$  such that

$$\hat{x} = c(\hat{p}, v_1) < c(\hat{p}, v_l) \quad \forall l \in \{2, \dots, m\}. \quad (35)$$

At the same time, by (Aliprantis and Border, 2006, Corollary 5.30) and since  $\text{co}(\mathcal{W}_k) = \bar{\text{co}}(\mathcal{W}_k) \supset \mathcal{W}_{\max\text{-nor}}$  and

$$V(p) = \min_{v \in \mathcal{W}_{\max\text{-nor}}} c(p, v) \quad \forall p \in \Delta,$$

there exists  $\bar{v} \in \mathcal{W}_{\max\text{-nor}}$  such that  $c(\hat{p}, \bar{v}) = V(\hat{p}) = c(\hat{p}, v_1)$ . Moreover, there exists  $\{\lambda_l\}_{l=1}^m \subseteq [0, 1]$  such that  $\sum_{l=1}^m \lambda_l = 1$  and  $\bar{v} = \sum_{l=1}^m \lambda_l v_l$ . Since  $v_1 \notin \mathcal{W}_{\max\text{-nor}}$ , this implies that  $\lambda_1 \neq 1$  and  $\lambda_l > 0$  for some  $l \in \{2, \dots, m\}$ . By (35), this implies that

$$\begin{aligned} \mathbb{E}_{\hat{p}}(v_l) > v_l(\hat{x}) \quad \forall l \in \{2, \dots, m\} &\implies \sum_{l=1}^m \lambda_l \mathbb{E}_{\hat{p}}(v_l) > \sum_{l=1}^m \lambda_l v_l(\hat{x}) \\ &\implies \mathbb{E}_{\hat{p}}(\bar{v}) > \bar{v}(\hat{x}) \implies c(\hat{p}, \bar{v}) > \hat{x} = c(\hat{p}, v_1), \end{aligned}$$

a contradiction.

<sup>27</sup>Otherwise,  $\bar{\text{co}}(\mathcal{W}_k) \subseteq \mathcal{W}_{\max\text{-nor}}$ , a contradiction.

We thus showed that  $\text{co}(\mathcal{W}_k) = \bar{\text{co}}(\mathcal{W}_k) = \mathcal{W}_{\text{max-nor}}$ . This implies that  $\mathcal{W}_k$  satisfies both (6) and (13). Moreover,  $\mathcal{W}_k$  is finite. By Theorem 6, this implies that  $\succsim$  satisfies Strict First Order Stochastic Dominance, Negative Certainty Independence, and NIRL. By Proposition 10 and since  $\mathcal{W}_k$  is finite, we can conclude that either  $|\mathcal{W}_k| = 1$  and  $\succsim$  satisfies Independence or  $1 < |\mathcal{W}_k| < \infty$  and  $\succsim$  violates Betweenness. ■

## References

- ALIPRANTIS, C. D. AND K. C. BORDER (2006): *Infinite Dimensional Analysis*, Springer, Berlin Heidelberg New York.
- ALIPRANTIS, C. D. AND O. BURKINSHAW (1998): *Principles of Real Analysis*, Academic Press, New York.
- ARTSTEIN-AVIDAN, S. AND D. DILLENBERGER (2015): “Dynamic Disappointment Aversion,” Mimeo University of Pennsylvania.
- CAROTHERS, N. L. (2000): *Real analysis*, Cambridge University Press.
- CERREIA-VIOGLIO, S. (2009): “Maxmin Expected Utility on a Subjective State Space: Convex Preferences under Risk,” Mimeo, Bocconi University.
- CERREIA-VIOGLIO, S., D. DILLENBERGER, AND P. ORTOLEVA (2015): “Cautious Expected Utility and the Certainty Effect,” *Econometrica*, 83, 693–728.
- CERREIA-VIOGLIO, S., F. MACCHERONI, AND M. MARINACCI (2017): “Stochastic dominance analysis without the independence axiom,” *Management Science*, 63, 1097–1109.
- CERREIA-VIOGLIO, S., F. MACCHERONI, M. MARINACCI, AND L. MONTRUCCHIO (2011): “Uncertainty Averse Preferences,” *Journal of Economic Theory*, 146, 1275–1330.
- CHEW, S. H. (1985): “Implicit Weighted and Semi-Weighted Utility Theories, M-estimators, and Non-Demand Revelation of Second-Price Auctions for an Uncertain Auction Object,” Working Paper 155, John Hopkins University.
- (1989): “Axiomatic utility theories with the betweenness property,” *Annals of Operations Research*, 19, 273–298.
- CHEW, S. H. AND K. R. MACCRIMMON (1979a): “Alpha-nu Choice Theory: a Generalization of Expected Utility Theory,” Working paper 669, University of British Columbia.
- (1979b): “Alpha Utility Theory, Lottery Composition and the Allais Paradox,” Working paper 686, University of British Columbia.
- DEKEL, E. (1986): “An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom,” *Journal of Economic Theory*, 40, 304–318.
- DEKEL, E. AND B. L. LIPMAN (2010): “How (not) to do decision theory,” *Annual Review of Economics*, 2, 257–282.

- DILLENBERGER, D. (2010): “Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior,” *Econometrica*, 78, 1973–2004.
- EECKHOUDT, L. AND H. SCHLESINGER (2006): “Putting risk in its proper place,” *American Economic Review*, 96, 280–289.
- GILBOA, I. (2009): *Theory of decision under uncertainty*, Cambridge University Press Cambridge.
- GUL, F. (1991): “A theory of disappointment aversion,” *Econometrica*, 59, 667–686.
- KIMBALL, M. S. (1990): “Precautionary Saving in the Small and in the Large,” *Econometrica*, 58, 53–73.
- KÖSZEGI, B. AND M. RABIN (2006): “A Model of Reference-Dependent Preferences,” *Quarterly Journal of Economics*, 121, 1133–1165.
- (2007): “Reference-dependent risk attitudes,” *The American Economic Review*, 97, 1047–1073.
- MACHINA, M. J. (1982): ““Expected Utility” Analysis without the Independence Axiom,” *Econometrica*, 50, 277–323.
- MASATLIOGLU, Y. AND C. RAYMOND (2016): “A behavioral analysis of stochastic reference dependence,” *American Economic Review*, 106, 2760–82.
- MEGGINSON, R. E. (1998): *An introduction to Banach space theory*, Springer.
- MENEZES, C., C. GEISS, AND J. TRESSLER (1980): “Increasing downside risk,” *The American Economic Review*, 70, 921–932.
- SAFRA, Z. AND U. SEGAL (2009): “Risk aversion in the small and in the large: Calibration results for betweenness functionals,” *Journal of risk and uncertainty*, 38, 27–37.
- SION, M. (1958): “On general minimax theorems,” *Pacific Journal of Mathematics*, 8, 171–176.