An Explicit Representation for Disappointment Aversion and Other Betweenness Preferences*

Simone Cerreia-Vioglio\textsuperscript{a}, David Dillenberger\textsuperscript{b}, Pietro Ortoleva\textsuperscript{c}
\textsuperscript{a}Universit`a Bocconi and Igier, \textsuperscript{b}University of Pennsylvania, \textsuperscript{c}Princeton University

May 2019

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) or its counterpart. Our results help studying the consequences of disappointment aversion in applications: we show that introducing disappointment aversion in typical problems has similar effects as increasing risk aversion under expected utility. We apply these results to the case of Bayesian Persuasion with disappointment averse senders.

JEL: D80, D81

KEYWORDS: Betweenness, Cautious Expected Utility, Disappointment Aversion, Non-Expected Utility, Utility Representation.

---

*We thank three anonymous referees and the editor, Rani Spiegler, for very useful suggestions. We thank Soo Hong Chew, Tommaso Denti, 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-1763326). 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 because it generalizes expected utility by adding only one parameter. Despite its appeal, there is one limitation to its applicability: 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, for example, to solve typical optimization problems. 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.\(^1\)

The goal of this paper is to address these issues. First, we provide an explicit representation for DA preferences, showing that it can be easily obtained using solely the components of its implicit one. Second, we generalize this result: we provide an explicit representation for Betweenness preferences that satisfy either Negative Certainty Independence (Dillenberger, 2010; Cerreia-Vioglio et al., 2015), or its positive version, Positive Certainty Independence. Third, we show how these explicit characterizations may be useful, for example, in comparative statics exercises performed in applications.

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

\[
v = \mathbb{E}_p(k_v)
\]

where \( k_v \) is given by

\[
k_v(x) = \begin{cases} 
  u(x) & \text{if } u(x) \leq v \\
  \frac{u(x) + \beta u}{1 + \beta} & \text{if } u(x) > v
\end{cases}
\]

Here \( u \) is a utility function over money, and \( \beta \in (-1, \infty) \) is 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 additional weight \( \beta \) to disappointing outcomes – those with a utility lower than the value of the lottery itself.\(^2\) The DA model is thus a model of endogenous reference dependence:

---

\(^1\)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.

\(^2\)To see this, note that the value of a simple lottery \( p \) can be equivalently defined as the unique \( v \) that solves

\[
v = \frac{\sum_{\{x:u(x)>v\}} u(x)p(x) + (1 + \beta) \sum_{\{x:u(x)\leq v\}} u(x)p(x)}{1 + \beta \sum_{\{x:u(x)\leq v\}} p(x)}.
\]
possible payoffs generate disappointment (or elation) depending on how their utilities compare to an endogenously-determined value – the utility of the lottery.\footnote{We should stress that this is conceptually and behaviorally distinct from other models of endogenous reference dependence under risk, e.g., Köszegi and Rabin (2006, 2007). For example, the DA model satisfies Betweenness, while both models above do not. See Masatlioglu and Raymond (2016) for further discussion on the implications of these alternative models.} 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. Using our results, we also show additional properties of this model: for example, that it exhibits prudence only if it is expected utility.

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), or its counterpart (Positive Certainty Independence). The previous results are corollaries of this more general theorem. Again, the explicit representation is the minimum (resp., maximum) of the certainty equivalents using the functions (called local utilities) present in Dekel (1986)’s implicit representation.

There are at least two benefits of having an explicit representation. The first one is conceptual: it may help capturing the mental process adopted by the agent. While highly idealized, one can easily imagine a cautious 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.\footnote{A related argument appears in Chapter 17 of Gilboa (2009). Dekel and Lipman (2010) argued that “Although 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 model’s predictions more. Perhaps more importantly, the story affects our intuitions about the model and hence whether and how we use and extend it.”}
The second and possibly main advantage of an explicit representation is practical: it might facilitate the application of these models by simplifying optimization problems with these preferences. This is particularly relevant because DA preferences, while continuous, are not even Gateaux differentiable (Safra and Segal, 2009). Therefore, one cannot apply standard differential methods or Machina (1982)'s local utility approach and its extensions to the Gateaux case. We conclude the paper by showing how our explicit representations may help. Using Sion’s minmax theorem in conjunction with our characterizations, we show that the solution to a maximization problem over a convex and compact set of options with DA preferences coincides with the solution of the original problem under expected utility, but with the agent being more risk averse. As the latter is typically easier to solve, this may greatly simplify the problem when it comes to comparative statics. The key point is not that disappointment aversion increases risk aversion, but rather that the solution under disappointment aversion must also be a solution under expected utility with a more concave utility. We conclude by illustrating the usefulness of this result with an application to the case of Bayesian Persuasion. Focusing on a simple case studied in Dworczak and Martini (2018) under expected utility, we show that making the sender of information disappointment averse induces her to commit to reveal less information.

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. We denote by $C([w, b])$ the space of continuous functions on $[w, b]$ and endow it with the topology induced by the supnorm. The set $U_{\text{nor}} \subseteq C([w, b])$ is the collection of all strictly increasing and continuous functions $v : [w, b] \to \mathbb{R}$ 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, $\succ_{\text{FSD}}$ denotes the First Order Stochastic Dominance relation, that is, $p \succ_{\text{FSD}} q$ means $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for all $v \in U_{\text{nor}}$.

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

A 1 (Weak Order) The relation $\succ$ is complete and transitive.
A 2 (Continuity) For each \( q \in \Delta \), the sets \( \{ p \in \Delta : p \succ q \} \) and \( \{ p \in \Delta : q \succ p \} \) are closed.

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

\[ p \succ_{\text{FSD}} q \implies p \succ q. \]

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 mixing: if satisfied, then the agent has no preference for, or aversion to, mixing 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:

Theorem 1 (Dekel, 1986) If \( \succ \) is a Betweenness preference, then there exists a function \( k : [w, b] \times [0, 1] \to \mathbb{R} \) 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 \( t \in [0, 1] \),
4. \( \succ \) can be represented by a continuous utility function which strictly preserves first order stochastic dominance, \( \hat{V} : \Delta \to [0, 1] \), where for each \( p \in \Delta \), \( \hat{V}(p) \) is the unique number in \([0, 1]\) such that

\[ \int_{[w,b]} k(x, \hat{V}(p)) dp = \hat{V}(p). \]  \hspace{1cm} (2)

\[ ^5 \text{Dekel's original result deals with a generic set of consequences and considers a weaker form of monotonicity. At the same time, it uses a stronger form of Betweenness. Given these differences, we prove Theorem 1 in Appendix B. For convenience, we focus on the normalized representation (that is, } k(\cdot, 0) \text{ and } k(\cdot, 1) \text{ are not assumed to be continuous, they are implicitly assumed to be integrable, given Equation (2) in the representation.} \]
Fixing $t$, the function $k(\cdot,t)$ is called the \textit{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 provides a representation for Betweenness preferences, it does not provide an explicit one: indeed, $\tilde{V}$ is the solution to Equation (2), thus a fixed point of a functional equation.

An important class of Betweenness preferences is the one generated by 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
\[
\int_{[w,b]} \tilde{k}(x,\tilde{V}(p)) \, dp = \tilde{V}(p)
\]
where $\tilde{k} : [w,b] \times \text{Im} u \rightarrow \mathbb{R}$ is defined by
\[
\tilde{k}(x,s) = \begin{cases} 
 u(x) & \text{if } u(x) \leq s \\
 \frac{u(x)+\beta s}{1+\beta} & \text{if } u(x) > s 
\end{cases} \forall x \in [w,b], \forall s \in \text{Im} u;
\]
here $u$ is a strictly increasing continuous utility function from $[w,b]$ to $\mathbb{R}$ and $\beta \in (-1,\infty)$.\footnote{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 (see Remark 4).} We discussed its interpretation in the Introduction. We say that a binary relation is a \textit{DA preference} if and only if it admits a utility function $\tilde{V}$ which satisfies (3) for some pair $(u,\beta)$.

\textbf{Negative Certainty Independence.} As noted by Dillenberger (2010), a DA preference with $\beta > 0$ satisfies the following axiom.

\textbf{A 5 (Negative Certainty Independence)} For each $p,q \in \Delta$, $x \in [w,b]$, and $\lambda \in [0,1]$
\[
p \succeq \delta_x \implies \lambda p + (1-\lambda)q \succneq \lambda \delta_x + (1-\lambda)q.
\]

Negative Certainty Independence, initially suggested by 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 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:7

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

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

(6)

is a continuous utility representation of $\succeq$.

### 3 Explicit Representations

We start by providing an explicit representation of DA preferences.

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

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

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

(7)

is a continuous utility representation of $\succeq$.

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

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

(8)

is a continuous utility representation of $\succeq$.

3. If $\beta < 0$, then $V : \Delta \to \mathbb{R}$, 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 $\succeq$.

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

---

7More precisely, Cerreia-Vioglio et al. (2015) state the result below as an equivalence using a weaker form of monotonicity. However, for ease of comparison with Theorem 1, we provide it using Strict First Order Stochastic Dominance.
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.

As discussed above, when \( \beta > 0 \) Gul’s model satisfies Negative Certainty Independence. We thus know that it must admit a Cautious Expected Utility representation. However, from previous results we do not know which utilities are used. The content of Theorem 3 is to show that this involves precisely the local utilities used in the implicit representation. Thus, the explicit representation can be derived directly from the implicit one. When \( \beta < 0 \), the model does not satisfy Negative Certainty Independence, but its counterpart 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 to derive further properties of DA preferences. Recall the notion of prudence (also known as downside risk aversion): a preference for additional risk on the upside rather than the downside of a gamble (Eeckhoudt and Schlesinger, 2006).\(^8\) 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 \( p \succ_{\text{pru}} q \Rightarrow p \succ q \). 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.

**A General Result.** We now generalize Theorem 3 showing 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:

(i) \( \succ \) satisfies Negative Certainty Independence;

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

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

\(^8\)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.
is a continuous utility representation of $\succeq$. In particular, for each $p \in \Delta \setminus \{\delta_w, \delta_b\}$ the function $v_p = k\left(\cdot, \hat{V}(p)\right)$ is such that

$$v_p \in \arg\min_{v \in W_{\text{bet}}} c(p, v).$$

(10)

Like in the case of the previous result, the contribution of Theorem 4 does not lie in showing that these preferences admit an explicit representation of the Cautious Expected Utility class – this was already known (it follows from Theorem 2). As before, the contribution lies in showing that the utilities involved are exactly the local utilities identified in Theorem 1 and included in $W_{\text{bet}}$. Here too, the explicit representation can be derived directly from 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).

**Remark 1** A specular version of this theorem also holds for Positive Certainty Independence. In particular, by keeping the same premises, Theorem 4 takes a similar form with (i) and (ii) replaced by:

(i)' $\succeq$ satisfies Positive Certainty Independence;

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

$$V(p) = \max_{v \in W_{\text{bet}}} c(p, v) \quad \forall p \in \Delta,$$

(11)

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

$$v_p \in \arg\max_{v \in W_{\text{bet}}} c(p, v).$$

(12)

\[ \nabla \]

While Theorem 4 provides an explicit characterization for Betweenness preferences that satisfy Negative Certainty Independence, a natural question is how to check if a given Betweenness preference satisfies Negative Certainty Independence. In Appendix A we show how this could be easily done using properties of the local utilities. This result allows us to derive an example of Betweenness preferences that satisfy Negative Certainty Independence but do not belong to the class of DA preferences.

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

$$k(x, t) = \begin{cases} x & \text{if } x \leq t \\ x(x - t) + t & \text{if } x > t \end{cases} \quad \forall x \in [0, 1], \forall t \in [0, 1].$$
This retains the idea of disappointment aversion, but allows the weight, $x - t$, to depend on $x$. In Appendix B we show that this preference relation satisfies Negative Certainty Independence and therefore admits an explicit representation as in Theorem 4 with the utilities above.

\[ \nabla \]

4 Applications: A General Result and Bayesian Persuasion

We now turn to illustrate how our explicit representations can be used to provide simple comparative statics results in applications. Typically in economic models, agents need to pick the best action from a convex and compact set of alternatives. Solving such problems with DA preferences, or more general Betweenness preferences, can be however not trivial: as pointed out in the introduction, standard differential methods cannot be used as DA preferences are not even Gateaux differentiable. We now show how our explicit representation results might help.

Consider a Betweenness preference $\succeq$ that satisfies Negative Certainty Independence — for DA preferences, this corresponds to the typical case of $\beta \geq 0$. By Theorem 4, $\succeq$ is represented by

\[
V(p) = \min_{v \in W_{\text{bet}}} c(p, v) \quad \forall p \in \Delta \tag{13}
\]

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

a) $k(\cdot, t)$ is strictly increasing on $[w, b]$ for all $t \in [0, 1]$; and

b) $k$ 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 2** Let $\succeq$ be a Betweenness preference that satisfies Negative Certainty Independence and such that $k$ satisfies a and b. If $A \subseteq \Delta$ is convex and compact, then

\[
\max_{p \in A} \min_{v \in c(\hat{c}(W_{\text{bet}}))} c(p, v) = \min_{v \in c(\hat{c}(W_{\text{bet}}))} \max_{p \in A} c(p, v).
\]

In particular, if $p^* \in A$ is such that $V(p^*) \geq V(p)$ for all $p \in A$, then there exists $\hat{v} \in c(\hat{c}(W_{\text{bet}}))$ such that $E_{p^*}(\hat{v}) \geq E_p(\hat{v})$ for all $p \in A$.

Proposition 2 shows that any alternative that maximizes the original preference in $A$ is also a maximizer of an expected utility preference with Bernoulli utility $\hat{v}$ which is

\[ 9 \text{Example 1 is a special case of Chew (1985)’s model of semi-implicit weighted utility, where } [w, b] \text{ is set to be equal to } [0, 1]. \]
a convex combination of the utilities in $W_{\text{bet}}$. This result could be useful in performing standard comparative statics exercises. We first provide a general discussion; in the next subsection, we apply this to the case of Bayesian Persuation.

Consider an economic model which boils down to specify a maximization problem of the form

$$\max V(p) \text{ subject to } p \in A$$

(14)

where $V : \Delta \rightarrow \mathbb{R}$ is continuous and represents the preferences of an agent. For simplicity, assume also it admits a unique maximizer. As it is often the case, assume that $V$ has been first considered to be an expected utility functional with a strictly increasing and continuous von Neumann-Morgenstern function $u$, and call $p_{\text{EU}}^*$ the solution of (14). Now suppose that we are interested in knowing what happens when agents are instead disappointment averse ($\beta > 0$), and call $p_{\text{DA}}^*$ the solution for this case. How do the predictions of the model change? Proposition 2, derived thanks to our explicit representations, simplifies solving this question. It shows that $p_{\text{DA}}^*$ must also be the solution of the same optimization problem, but for an expected utility agent with von Neumann-Morgenstern function $\hat{v} \in \text{co}(W_{\text{bet}})$. It follows immediately from the shape of the functions in $W_{\text{bet}}$ that each $v \in \text{co}(W_{\text{bet}})$ is more concave than $u$, which is easy to see because our characterization theorems give a precise functional form to these functions. Therefore, the solution to the problem with DA preferences must coincide with the solution of the original problem under expected utility, but with the agent being more risk averse. This means that comparative statics exercises in terms of introducing disappointment aversion are equivalent to ones under expected utility in terms of concavity of $u$. As the latter are typically easy to solve, this may greatly simplify the problem. The conceptual contribution here does not lie in showing that adding disappointment aversion increases risk aversion – this is immediate and well-known (Gul, 1991). Rather, it lies in showing that the solution under disappointment aversion must also be a solution under expected utility – thus, often easy to solve – with a more risk averse agent. We illustrate a practical application of this next.

### 4.1 Bayesian Persuasion without Expected Utility

A large literature has been devoted to analyzing the problem of ‘Bayesian Persuasion’ – models of strategic communication with commitment power. Different papers have analyzed possible variations of the original model of Kamenica and Gentzkow (2011). We now use our results to study how the solution changes when the sender of information has DA preferences rather than expected utility, focusing on a simple case that extends the one studied in Dworczak and Martini (2018, Section 6.1).\(^{10}\) Two agents,

\(^{10}\)We gratefully acknowledge many useful discussions with Tommaso Denti in writing this section.
Sender and Receiver, share a full-support prior, with continuous distribution, \( r \) on a state of the world (the value of a project) \( x \in [0, 1] \).\(^\text{11}\) Sender observes the state before Receiver has to make a choice, and can commit to a disclosure policy to Receiver, chosen before observing the realization of the state.

Receiver chooses the amount of costly effort \( e \in [0, \bar{e}] \) to exert, where \( \bar{e} \leq 1 \). If \( y \) is the mean of Receiver’s posterior at the time of choice, her expected utility is

\[
\hat{u}(y, e) = \sigma y e - e^\alpha \quad \text{where } 0 < \sigma < 1 \text{ and } \alpha > 1. \tag{15}
\]

Standard calculations show that she exerts effort

\[
e^\ast(y) = \min \left\{ \left( \frac{\sigma y}{\alpha} \right)^{\frac{1}{\alpha-1}}, \bar{e} \right\}.
\]

Sender instead chooses a disclosure policy to commit to. Her utility is increasing in both the effort and the receiver’s posterior mean \( y \), and corresponds to:\(^\text{12}\)

\[
u(y) = (1 - \sigma) e^\ast(y) y \quad \forall y \in [0, 1]. \tag{16}
\]

As standard in this literature, Sender’s problem can be rephrased as choosing a distribution over Receiver’s posteriors; but since in this case only their means matter, this simplifies the problem to choosing a distribution over posterior means. Note that fully revealing the state corresponds to generating a distribution of posterior means equal to \( r \); while revealing less information implies a disclosure policy \( p \) of which \( r \) is a mean preserving spread – denote this by \( p \geq_{\text{MPS}} r \).\(^\text{13}\) Thus, Sender’s problem corresponds to choosing a distribution from the set \( A = \{ p \in \Delta : p \geq_{\text{MPS}} r \} \).

Let \( \bar{y} = \left( \frac{\alpha}{\sigma} \right) \bar{e}^{\alpha-1} \),\(^\text{14}\) i.e., the smallest posterior mean for which the agent will exert maximum effort \( \bar{e} \). Dworczak and Martini (2018) assume Sender has an expected utility preference. Denote by \( p_{\text{EU}}^* \) an optimal disclosure policy. They show that if \( \mathbb{E}_r(X) \geq \bar{y} \), then it is optimal to reveal nothing, i.e., \( p_{\text{EU}}^* = \delta_{\mathbb{E}_r(X)} \). Otherwise, let \( x^* \) be such that

\[\text{11}\text{Here, the distribution of } r \text{ is the function } F_r : [0, 1] \to [0, 1] \text{ defined by } F_r(t) = r([0, t]) \quad \forall t \in [0, 1].\]

Moreover, \( x \) is the realization of a measurable random variable \( X \) which has distribution \( F_r \).

\[\text{12}\text{Note that in line with the analysis of Dworczak and Martini (2018), Sender’s utility depends only on Receiver’s posterior mean. These utilities would emerge, for example, in the case in which Sender and Receiver agree that the latter pays a price } (1 - \sigma) ye \text{ to acquire the entire payoff from a project, the expectation of which is } ye. \text{ (Indeed, } e \text{ in this case could be understood as either the probability of success or an increase in value.)}\]

\[\text{13}\text{Formally, for each } p, q \in \Delta, \text{ we write } p \geq_{\text{MPS}} q \text{ if and only if } \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \text{ for all real-valued, continuous, and concave functions } v \text{ on } [0, 1]. \text{ In other words, } \geq_{\text{MPS}} \text{ is the concave order over } \Delta.\]

\[\text{14}\text{Like Dworczak and Martini (2018), to make the analysis interesting, we assume that } 0 < \bar{y} < 1.\]
\( \bar{y} = \mathbb{E}_r (X | x \geq x^*) \). Then, it is optimal to disclose \( x \) if \( x < x^* \), and reveal only that \( x \geq x^* \) if \( x \geq x^* \). Formally, the distribution of \( p^*_{\text{EU}} \) is

\[
F_{\text{EU}} (x) = \begin{cases} 
F_r (x) & x < x^* \\
F_r (x^*) & x^* \leq x < \bar{y} \\
1 & x \geq \bar{y} 
\end{cases} \quad \forall x \in [0, 1].
\]

This is an intuitive result. When the prior is such that it is already optimal to exert maximal effort, Sender should reveal nothing. When this is not the case, Sender may wish to reveal the exact state for low values – to avoid Receiver exerting too low effort – and only reveal that the state is above a threshold for all values such that the posterior is \( \bar{y} \). It is easy to see how this leads to a higher expected utility than revealing nothing. In the latter case, the effort would be constant but suboptimal. By contrast, this scheme guarantees that effort is maximal for the largest set of high values of \( x \), where effort is most valuable.

We now extend this result by modifying the preferences of Sender, making them disappointment averse instead of expected utility: we assume they have a representation as in (7), where \( u \) is as in (16) and \( \beta > 0 \). How does this affect the optimal disclosure policy? As discussed in the previous section, any solution of the problem with DA preferences coincides with the solution with an expected utility decision maker who is more risk averse than the original one. Because the latter is easy to study, as we show below, this greatly simplifies obtaining these comparative statics. Denote by \( p^*_{\text{DA}} \) a solution of the problem under disappointment aversion.

**Proposition 3** Let Sender have DA preferences. If \( \mathbb{E}_r (X) \geq \bar{y} \), then it is still optimal to reveal no information, that is, \( p^*_{\text{DA}} = p^*_{\text{EU}} = \delta_{\mathbb{E}_r (X)} \). Otherwise, it is optimal to reveal less information than the case of expected utility, that is, \( p^*_{\text{DA}} \geq_{\text{MPS}} p^*_{\text{EU}} \).

This result is intuitive once we see it through the lense of an increase in risk aversion under expected utility. Clearly, when \( \mathbb{E}_r (X) \geq \bar{y} \), it remains optimal to reveal no information as it guarantees maximal effort in any instance. But now consider the case in which \( \mathbb{E}_r (X) < \bar{y} \). Under expected utility, we have seen that it was optimal to reveal the state exactly until a value \( x^* \); above it, the Sender should only reveal that the state is above that threshold. This was optimal because it was maximizing the probability that the highest effort was exerted when most valuable. But note that this disclosure policy, while it has a higher expected utility, also has higher spread: for example, for low values of \( x \), the Sender receives very little as both the state and the effort are now low. This means that if the Sender is more risk averse, this disclosure policy may no longer be optimal: in this case she may prefer a policy with less dispersion, which in turn must be a policy that discloses less information.
Appendix

A When do Betweenness Preferences Satisfy NCI?

While Theorem 4 provides an explicit characterization for Betweenness preferences that satisfy Negative Certainty Independence, a natural question is how to actually check if a given Betweenness preference does or does not satisfy Negative Certainty Independence. The next result, paired with the next remark, provides a rather simple tool to answer this question, solely in terms of the functional form of $k$.

In stating these results, the following notation will be useful: given $f : [0, 1] \to [0, 1]$, we say that $f$ is convex (resp., concave) at $t \in (0, 1)$ if and only if for each $n \in \mathbb{N}$, \( \left\{ t_i \right\}_{i=1}^{n} \subseteq [0, 1] \), and \( \left\{ \lambda_i \right\}_{i=1}^{n} \subseteq [0, 1] \) such that \( \sum_{i=1}^{n} \lambda_i = 1 \)

\[
 t = \sum_{i=1}^{n} \lambda_i t_i \implies f(t) \leq \sum_{i=1}^{n} \lambda_i f(t_i) \quad \text{(resp.,} \geq \text{)}.
\]

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, $f_{s,t} : [0, 1] \to [0, 1]$ is such that $k(x, s) = f_{s,t}(k(x, t))$ for all $x \in [w, b]$. 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 4** 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 (resp., concave) at $t$;

(ii) $\succsim$ satisfies Negative (resp., Positive) Certainty Independence.

Proposition 4 characterizes Negative Certainty Independence within the class of Betweenness preferences, just in terms of the parameters of their representation. In fact, testing Negative (resp., Positive) Certainty Independence amounts to checking if for each $t \in (0, 1)$ the transformations $f_{s,t}$ are convex (resp., concave) at $t$ for all $t \in (0, 1)$. This is a handy tool since $f_{s,t} = k(\cdot, s) \circ k^{-1}(\cdot, t)$ and is thus computable. Moreover, checking convexity and concavity at $t$ is rather simple in light of next remark.

**Remark 2** Convexity at $t$ is implied by the following sufficient condition: the subdifferential of $f_{s,t}$ is nonempty at $t$, $\partial f_{s,t}(t) \neq \emptyset$. 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 $g : [0, 1] \to \mathbb{R}$ such that $g(t') = mt' + l$
for all $t' \in [0, 1]$, where $l, m \in \mathbb{R}$ and $f_{s,t}(t) = g(t)$ as well as $g(t') \leq f_{s,t}(t')$ for all $t' \in [0, 1]$.\(^{15}\)

We illustrate the usefulness of these results in Appendix B, while proving that the preferences in Example 1 satisfy Negative Certainty Independence.

## B Proofs

The appendix is structured as follows. We begin by proving Theorem 1, a mildly modified version of Dekel’s representation result. We proceed by recalling the notion of expected utility core and its role for Cautious Expected Utility preferences (Formula (19) and Remark 3). Using it, we characterize the expected utility core of Betweenness preferences (Propositions 5-7). These allow us to prove our most general representation result, Theorem 4 (resp., Remark 1). We then move to proofs for the special case of DA preferences. First, we discuss the issue of renormalization of Gul’s locals (Remark 4); with this in mind, we prove Theorem 3 as a special case of Theorem 4. This allows us to prove the incompatibility of prudence with DA preferences (Proposition 1). Next, we prove the results of our application section: our maxmin result (Proposition 2) and the Bayesian Persuasion application (Proposition 3). Finally, we prove Proposition 4 and Remark 2 in Appendix A, and use them to show that the preferences in Example 1 indeed satisfy Negative Certainty Independence.

### Proof of Theorem 1

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

\[
p \succeq q \implies p \succeq \lambda p + (1 - \lambda) q \quad \forall \lambda \in (0, 1)
\]

and

\[
p > q \implies p > \lambda p + (1 - \lambda) q \quad \forall \lambda \in (0, 1).
\]

The proof of the first implication is routine.\(^{16}\) As for the second, suppose $p \succ q$. By the first implication, we have that $p \succeq \lambda p + (1 - \lambda) q \succeq q$ 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:

1. $p = \delta_b$. Since $p \succ q$, we obtain 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.

\(^{15}\)Clearly, a specular sufficient condition can be provided for concavity at $t$. One needs to focus on the superdifferential at $t$, rather than the subdifferential, and the above inequality will be reversed.

\(^{16}\)For example, it can be proved by using the techniques of (Cerreia-Vioglio et al., 2011, Lemma 56).
2. $p \neq \delta_b$. Since $\succ$ satisfies Betweenness, it follows that

$$1 \geq \lambda \geq \bar{\lambda} \Rightarrow \lambda p + (1 - \lambda) q \sim p.$$  \hspace{1cm} (17)

Since $\succ$ satisfies Strict First Order Stochastic Dominance, we have that $\gamma p + (1 - \gamma) \delta_b > p$ for all $\gamma \in (0, 1)$. By (17) and since $\succ$ satisfies Strict First Order Stochastic Dominance, we can conclude that

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

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

$$\lambda p + (1 - \lambda) r_{n, \gamma} =$$

$$= \lambda p + (1 - \lambda) [\eta (\gamma p + (1 - \gamma) \delta_b) + (1 - \eta) q]$$

$$= (\lambda + (1 - \lambda) \eta \gamma) p$$

$$+ (1 - \lambda - (1 - \lambda) \eta \gamma) \left[ \frac{(1 - \lambda) \eta (1 - \gamma)}{1 - \lambda - (1 - \lambda) \eta \gamma} \delta_b + \frac{(1 - \lambda) (1 - \eta)}{1 - \lambda - (1 - \lambda) \eta \gamma} q \right].$$

Since $\gamma p + (1 - \gamma) \delta_b > p > q$ 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

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

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

$$(\lambda + (1 - \lambda) \bar{\eta}_\gamma \gamma) p + (1 - \lambda - (1 - \lambda) \bar{\eta}_\gamma \gamma) \hat{r} \sim p.$$  \hspace{1cm} (18)

By (17) and (18), we can conclude that $\hat{\lambda} \in (\bar{\lambda}, 1)$,

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

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

We recall the definition of expected utility core of $\succ$, i.e., the subrelation $\succ'$ defined as:\footnote{Under Axioms A 1-2, one can show that $\succ'$ satisfies all the assumptions of expected utility with possibly the exception of completeness; and that it is the largest subrelation of $\succ$ satisfying these properties. See Cerreia-Vioglio (2009); Cerreia-Vioglio et al. (2015, 2017).}

$$p \succ' q \iff \lambda p + (1 - \lambda) r \succ \lambda q + (1 - \lambda) r \quad \forall \lambda \in (0, 1], \forall r \in \Delta.$$  \hspace{1cm} (19)
This notion is useful for two reasons. First, as Remark 3 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 $\succeq'$. This is instrumental in proving Theorem 4 (cf. Proposition 7). Second, as shown by Cerreia-Vioglio et al. (2017), in general $\succeq'$ summarizes the risk attitudes of the decision maker irrespective of whether or not $\succeq$ satisfies Negative Certainty Independence. In particular, $\succeq$ is averse to Mean Preserving Spreads if and only if $\succeq'$ 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 3** In addition to what is stated in Theorem 2, it can also be shown that the following are true:

1. There exists a set $\mathcal{W} \subseteq U_{\text{nor}}$ such that
   \[ p \succeq' q \iff \mathbb{E}_p (v) \geq \mathbb{E}_q (v) \quad \forall v \in \mathcal{W} \tag{20} \]
   and $V : \Delta \rightarrow \mathbb{R}$ defined as in Equation (6) is a continuous utility representation of $\succeq$.

2. If $\mathcal{W} \subseteq U_{\text{nor}}$ satisfies (20), then it satisfies (6).

3. The set $\mathcal{W} \subseteq U_{\text{nor}}$ can be chosen to be
   \[ \mathcal{W}_{\text{max-nor}} = \{ v \in U_{\text{nor}} : p \succeq' q \implies \mathbb{E}_p (v) \geq \mathbb{E}_q (v) \} . \]

4. If $\mathcal{W} \subseteq U_{\text{nor}}$ satisfies (20), then
   \[ \mathcal{W} \subseteq \mathcal{W}_{\text{max-nor}} \text{ as well as } \overline{\text{co}} (\mathcal{W}) = \overline{\text{cl}} (\mathcal{W}_{\text{max-nor}}) . \]

\[ \nabla \]

We next prove a few results pertaining to the expected utility core of a Betweenness preference. These results rely on some of the techniques developed in Cerreia-Vioglio et al. (2017). 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 $t \in [0, 1]$

$$K\left(r, \hat{V}(r)\right) = \int_{[w, b]} k\left(x, \hat{V}(r)\right) dr = \hat{V}(r) = \hat{V}(r) k(b, t) + \left(1 - \hat{V}(r)\right) k(w, t)$$

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

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 5** Let $\succeq$ be a Betweenness preference. If $K(p, t) \succeq K(q, t)$ for all $t \in (0, 1)$, then $p \succeq q$.

**Proof.** Consider $p, q \in \Delta$. By contradiction, assume that $K(p, t) \succeq K(q, t)$ for all $t \in (0, 1)$ and $q \succeq p$. 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 can conclude that $p = \delta_b = q$, a contradiction with $q \succeq p$. In the second case, we have that $\hat{V}(q) \in (0, 1)$.

On the one hand, since $\succeq$ admits a representation a la Dekel, note that

$$\hat{V}(q) = K\left(q, \hat{V}(q)\right) \leq K\left(p, \hat{V}(p)\right). \quad (21)$$

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

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

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

In particular, it follows that

$$K\left(\hat{V}(p) \delta_b + \left(1 - \hat{V}(p)\right) \delta_w, \hat{V}(p)\right) = \hat{V}(p) = K\left(p, \hat{V}(p)\right) \quad (22)$$

and

$$\hat{V}(q) > K\left(\hat{V}(p) \delta_b + \left(1 - \hat{V}(p)\right) \delta_w, \hat{V}(q)\right). \quad (23)$$

Define $r = \hat{V}(p) \delta_b + \left(1 - \hat{V}(p)\right) \delta_w$. By (21) and (23) 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 (22), we have that $r \sim p$. Since $\succ$ is a Betweenness preference, this 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. 

**Proposition 6** Let $\succ$ 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 $\hat{V}((\lambda p + (1 - \lambda) \delta_y) = \bar{t}$. It follows that

$$\bar{t} = K((\lambda p + (1 - \lambda) \delta_y) = \lambda K(p, \bar{t}) + (1 - \lambda) K(\delta_y, \bar{t})$$

$$< \lambda K(q, \bar{t}) + (1 - \lambda) K(\delta_y, \bar{t}) = K(\lambda q + (1 - \lambda) \delta_y, \bar{t}).$$

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 obtain that

$$\hat{V}(r_1) < K(r_2, \hat{V}(r_1)).$$

Since $p \succ q$ and $\succ$ satisfies Independence, it follows that $r_1 \succ r_2$. Since $\succ$ is a subrelation of $\succ$, this implies that $r_1 \succ r_2$, 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 (24), we obtain that

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

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

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

yielding that $\gamma r_2 + (1 - \gamma) r_3 \sim r_1$. Since $\succ$ 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, $\hat{V}(r_1) = \hat{V}(r_2) = K(r_2, \hat{V}(r_2)) = K(r_2, \hat{V}(r_1))$, a contradiction with (24).

**Proposition 7** If $\succ$ is a Betweenness preference, then

$$p \succ q \iff E_p(v) \geq E_q(v), \quad \forall v \in \mathcal{W}_{\text{bet}}.$$ 

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

---

18If $\hat{V}(p) \geq \bar{t} > 0 = \hat{V}(\delta_w)$, 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 $\hat{V}$. 

18
Continuity, and Strict First Order Stochastic Dominance. By Theorem 2 and Remark 1), the statement trivially follows.

Proof of Theorem 4. (ii) implies (i). We now prove Theorem 4. We next prove (10). Consider \( p \in \Delta \setminus \{\delta_w, \delta_b\} \). Since \( \succcurlyeq \) 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

\[
\int_{[w, b]} k \left( x, \hat{V} (p) \right) dp = \hat{V} (p). \tag{25}
\]

Define \( v_p = k \left( \cdot, \hat{V} (p) \right) \in \mathcal{W}_{\text{bet}} \). 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 (25), it follows that

\[
\int_{[w, b]} k (x, \hat{V} (p)) d\delta_{\bar{x}} = v_p (\bar{x}) = \hat{V} (p).
\]
Since $\succeq$ is a Betweenness preference, we can conclude that $\hat{V}(\delta_x) = \hat{V}(p)$, that is, $\delta_x \sim p$ and so $\bar{x} = x_p$.\(^{19}\) This yields that

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

proving that the inf is attained at $v_p$.

\[\blacksquare\]

**Proof of Remark 1. (i)' implies (ii)'**. Since $\succeq$ satisfies Weak Order, Continuity, and Strict First Order Stochastic Dominance, there exists a continuous utility function $V : \Delta \to \mathbb{R}$ such that $V(\delta_x) = x$ for all $x \in [w, b]$. By Proposition 7, the expected utility core $\succeq'$ of $\succeq$ admits an expected multi-utility representation with set $\mathcal{W}_{\text{bet}}$. Define $\bar{V} : \Delta \to \mathbb{R}$ by $\bar{V}(p) = \sup_{v \in \mathcal{W}_{\text{bet}}} c(p, v)$ for all $p \in \Delta$. We next show that $V = \bar{V}$. Fix $p \in \Delta$. Note that $\bar{V}(p), V(p) \in [w, b]$. By construction, it is immediate to see that

$$c(\delta_{\bar{V}(p)}, v') = \bar{V}(p) = \sup_{v \in \mathcal{W}_{\text{bet}}} c(p, v) \geq c(p, v') \quad \forall v' \in \mathcal{W}_{\text{bet}}.$$

By Proposition 7, this yields that $\delta_{\bar{V}(p)} \succeq' p$. Since $\succeq'$ is a subrelation of $\succeq$, we can conclude that $\delta_{\bar{V}(p)} \succeq p$, that is, $\bar{V}(p) = V(\delta_{\bar{V}(p)}) \geq V(p)$, proving that $\bar{V} \geq V$. By contradiction, assume that $\bar{V} \not\geq V$. It would follow that there exists $p \in \Delta$ such that $\bar{V}(p) > V(p)$. This would imply that there exists $y \in [w, b]$ such that $\bar{V}(p) > y > V(p)$. We could conclude that $\delta_y \not\succeq p$. By Proposition 7 and since $\succeq$ satisfies Positive Certainty Independence, we could conclude that $\delta_y \succeq' p$, that is, $y = c(\delta_y, v) \geq c(p, v)$ for all $v \in \mathcal{W}_{\text{bet}}$, yielding that $\bar{V}(p) > y \geq \sup_{v \in \mathcal{W}_{\text{bet}}} c(p, v) = \bar{V}(p)$, a contradiction. This proves that

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

In order to prove that (12) and (11) hold, the same arguments used in proving (10) and (9) apply.

(ii)' implies (i)'\). Consider $x \in [w, b]$ and $p, q \in \Delta$ as well as $\lambda \in [0, 1]$. Note that

$$\delta_x \succeq p \implies \sup_{v \in \mathcal{W}_{\text{bet}}} c(\delta_x, v) \geq \sup_{v \in \mathcal{W}_{\text{bet}}} c(p, v) \implies x \geq \sup_{v \in \mathcal{W}_{\text{bet}}} c(p, v)$$

$$\implies c(\delta_x, v) = x \geq c(p, v) \quad \forall v \in \mathcal{W}_{\text{bet}} \implies \delta_x \not\succeq' p$$

$$\implies \lambda \delta_x + (1 - \lambda) q \succeq \lambda p + (1 - \lambda) q,$$

proving that $\succeq$ satisfies Positive Certainty Independence. \[\blacksquare\]

\(^{19}\)Recall that a Betweenness preference is a binary relation on $\Delta$ which satisfies Weak Order, Continuity, Strict First Order Stochastic Dominance, and Betweenness. In this case, given $p \in \Delta$, $x_p$ is the unique number such that $\delta_{x_p} \sim p$. 

20
Remark 4 As already observed, a careful inspection of Gul’s model formulation (see (4)) suggests that two types of normalizations are due to link the implicit Gul’s representation to Dekel’s one. In order to do so, we define few objects: $\alpha : \text{Im} u \to \mathbb{R}$, $\gamma : \text{Im} u \to \mathbb{R}$, $\hat{V} : \Delta \to \mathbb{R}$ and $k : [w, b] \times [0, 1] \to \mathbb{R}$. We set

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

We also set

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

and, since $g$ is strictly increasing, continuous, and $\text{Im} g = \text{Im} u = \text{Im} \hat{V}$, $^{20}$

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

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.$^{21}$ Since

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

and $g : [0, 1] \to \text{Im} u$ 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 $W_{da} = \{ \tilde{k} (\cdot, z) \}_{z \in \text{Im} u}$. In particular, this implies that $\inf_{v \in W_{bet}} c (p, v) = \min_{v \in W_{da}} c (p, v)$ for all $p \in \Delta$ as well as $\sup_{v \in W_{bet}} c (p, v) = \max_{v \in W_{da}} c (p, v)$.

Proof of Theorem 3. 1 and 2. By Dillenberger (2010) and Artstein-Avidan and Dillenberger (2015) and since $\beta \geq 0$, it follows that $\succeq$ satisfies Negative Certainty Independence. By Theorem 4 and Remark 4, it follows that if $\beta \geq 0$, then $V : \Delta \to \mathbb{R}$, defined by

$$V (p) = \min_{v \in W_{bet}} c (p, v) = \min_{v \in W_{da}} c (p, v) \quad \forall p \in \Delta$$

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

---

$^{20}$Indeed, one has that for each $\lambda \in [0, 1]$

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

$^{21}$Indeed, points 1 and 2 are satisfied on $[0, 1]$ and not just $(0, 1)$.  

21
3. By Dillenberger (2010) and Artstein-Avidan and Dillenberger (2015) and since 
\( \beta \in (-1, 0) \), it follows that \( \gg \) satisfies Positive Certainty Independence. By Remarks
1 and 4, it follows that if \( \beta \in (-1, 0) \), then \( V : \Delta \rightarrow \mathbb{R} \), defined by

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

is a continuous utility representation of \( \gg \) where \( \mathcal{W}_{\text{bet}} = \{ k(\cdot, t) \}_{t \in (0, 1)} \).

**Proof of Proposition 1.** Only if. By (Cerreia-Vioglio et al., 2017, Fact 2 and
Lemma 1), if \( \gg \) exhibits prudence, then any set representing \( \gg' \) via an expected multi-
utility representation must be made of functions which are differentiable on \((w, b)\)
and have convex derivatives there. By Proposition 7, we can conclude that this set
can be chosen to be \( \mathcal{W}_{\text{bet}} = \{ k(\cdot, t) \}_{t \in (0, 1)} \). By Remark 4, \( \mathcal{W}_{\text{bet}} \) 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.

**Proof of Proposition 2.** Define \( \mathcal{W} = c_0(\mathcal{W}_{\text{bet}}) \). Given the assumptions, \( \mathcal{W} \)
convex and compact and (13) holds with \( \mathcal{W} \) in place of \( \mathcal{W}_{\text{bet}} \).\(^{22}\) 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

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

\(^{22}\)Since \( k : [w, b] \times [0, 1] \rightarrow \mathbb{R} \) is jointly continuous and \( k(x, t) \in [0, 1] \) for all \( x \in [w, b] \)
and all \( t \in [0, 1] \), it follows that \( \mathcal{W}_{\text{bet}} = \{ k(\cdot, t) \}_{t \in (0, 1)} \) is a bounded and equicontinuous family of functions in \( C([w, b]) \).
This implies that the convex hull of \( \mathcal{W}_{\text{bet}} \), \( c_0(\mathcal{W}_{\text{bet}}) \), is bounded and equicontinuous. Since closure in
supnorm preserves boundedness and equicontinuity, we can conclude that \( c_0(\mathcal{W}_{\text{bet}}) = \text{cl}(c_0(\mathcal{W}_{\text{bet}})) \) is
a bounded and equicontinuous family of functions of \( C([w, b]) \). By Arzela-Ascoli Theorem, \( c_0(\mathcal{W}_{\text{bet}}) \)
is compact. Finally, we are left to show that \( c_0(\mathcal{W}_{\text{bet}}) \) is a subset of \( \mathcal{U}_{\text{nor}} \). Clearly, each \( v \in c_0(\mathcal{W}_{\text{bet}}) \)
is continuous and such that \( v(w) = 0 \) and \( v(b) = 1 \). Thus, we only need to prove that \( v \) is strictly
increasing. Let \( x, y \in [w, b] \) be such that \( x > y \). Since \( k \) is strictly increasing in the first argument
and continuous in the second, we have that there exists \( \bar{t} \in [0, 1] \) and \( \varepsilon > 0 \) such that

\[
\inf_{v \in \mathcal{W}_{\text{bet}}} [v(x) - v(y)] \geq \min_{t \in [0, 1]} [k(x, t) - k(y, t)] = k(x, \bar{t}) - k(y, \bar{t}) \geq \varepsilon > 0.
\]

It is immediate to show that \( \inf_{v \in c_0(\mathcal{W}_{\text{bet}})} [v(x) - v(y)] \geq \varepsilon \), yielding that \( \inf_{v \in c_0(\mathcal{W}_{\text{bet}})} [v(x) - v(y)] \geq \varepsilon \). This implies that \( v(x) > v(y) \) for all \( v \in c_0(\mathcal{W}_{\text{bet}}) \). Since \( x \) and \( y \) were arbitrary, this shows that
each element of \( c_0(\mathcal{W}_{\text{bet}}) \) is strictly increasing.
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

\[
\max_{p \in A} \min_{v \in \mathcal{W}} c(p, v) = \max_{p \in A} V(p) = V(p^*) = \min_{v \in \mathcal{W}} c(p^*, v) \leq c(p^*, \hat{v}) \leq \max_{p \in A} \min_{v \in \mathcal{W}} c(p, v).
\]

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

\[
c(p^*, \hat{v}) = \max_{p \in A} c(p, \hat{v}),
\]

proving the statement. \( \blacksquare \)

**Proof of Proposition 3.** Let \([w, b] = [0, 1]\). Denote by \( \succeq \) the DA preference of Sender and assume \( \succeq \) is represented by \( V: \Delta \rightarrow \mathbb{R} \) as in Theorem 3 where \( u \) is as in (16) and \( \beta > 0 \). The problem of Sender is

\[
\max_{p \in A} V(p) \quad \text{subject to } p \in A \quad (26)
\]

where \( A = \{p \in \Delta : p \geq_{\text{MPS}} r\} \). Before starting, we need some notation. Given an element \( p \in \Delta \), we denote by \( F_p : [0, 1] \rightarrow [0, 1] \) the distribution of \( p \), that is, \( F_p(t) = p([0, t]) \) for all \( t \in [0, 1] \). We denote by \( e(p) \) the expectation of \( p \). Recall that, provided \( \mathbb{E}_r(X) < \bar{y} \), \( \mathbb{E}_{FU} : [0, 1] \rightarrow [0, 1] \) is defined by

\[
\mathbb{E}_{FU}(x) = \begin{cases} 
F_r(x) & x < x^* \\
F_r(x^*) & x^* \leq x < \bar{y} \\
1 & x \geq \bar{y}
\end{cases} \quad \forall x \in [0, 1]
\]

and define \( \mathbb{F}_{DA} = \mathbb{F}_{DA}^* \) where \( \mathbb{F}_{DA}^* \) is a solution to (26). Define \( \mathcal{U}_{\text{conc}} \) to be the set of all functions \( v : [0, 1] \rightarrow \mathbb{R} \) which are continuous and concave. Finally, denote by \( \mathcal{B} \) the Borel \( \sigma \)-algebra of \([0, 1]\). We have two cases:

1. \( \mathbb{E}_r(X) \geq \bar{y} \). By (Dworczak and Martini, 2018, Proposition 3), \( \mathbb{E}_r(X) = c(\delta_{\mathbb{E}_r}(X), u) \geq c(p, u) \) for all \( p \in A \) and \( \delta_{\mathbb{E}_r}(X) \in \mathcal{A} \). By Remark 4 and since any \( v \in \text{co} (\mathcal{W}_{\text{bet}}) \) is more concave than \( u \), it follows that \( c(\delta_{\mathbb{E}_r}(X), v) = \mathbb{E}_r(X) \geq c(p, v) \) for all \( p \in A \) and for all \( v \in \text{co} (\mathcal{W}_{\text{bet}}) \). By Proposition 7, this yields that \( \delta_{\mathbb{E}_r}(X) \succeq' p \) for all \( p \in A \). Since \( \succeq' \) is a subrelation of \( \succeq \), it follows that \( \delta_{\mathbb{E}_r}(X) \succeq p \) for all \( p \in A \), proving the statement.

2. \( \mathbb{E}_r(X) < \bar{y} \). Since \( A \) is compact and \( V \) is continuous, let \( \mathbb{F}_{DA}^* \) be a solution of (26). By Proposition 2, there exists \( \hat{v} \in \text{co} (\mathcal{W}_{\text{bet}}) \) such that

\[
\mathbb{E}_{\mathbb{F}_{DA}^*}(\hat{v}) \geq \mathbb{E}_p(\hat{v}) \quad \forall p \in A. \quad (27)
\]

By contradiction, assume that \( \mathbb{F}_{DA}^* \not\succeq_{\text{MPS}} \mathbb{F}_{EU}^* \). Define \( \Phi_{DA}, \Phi_{EU} : [0, 1] \rightarrow \mathbb{R} \) by

\[
\Phi_{DA}(t) = \int_0^t F_{DA}(x) \, dx \quad \text{and} \quad \Phi_{EU}(t) = \int_0^t F_{EU}(x) \, dx \quad \forall t \in [0, 1].
\]

\[23\text{Recall that, equivalently, } p \geq_{\text{MPS}} q \text{ if and only if } \int_0^1 F_p(x) \, dx \leq \int_0^1 F_q(x) \, dx \text{ for all } t \in [0, 1] \text{ and } \int_0^1 F_p(x) \, dx = \int_0^1 F_q(x) \, dx.\]
It is immediate to see that $\Phi_{DA}$ and $\Phi_{EU}$ are continuous functions. Since, by construction, $p_{EU}^*, p_{DA}^* \geq_{MPS} r$, we necessarily have that $\Phi_{DA}(0) = \Phi_{EU}(0)$ as well as

$$\Phi_{DA}(1) = \int_0^1 F_{DA}(x) \, dx = \int_0^1 F_{r}(x) \, dx = \int_0^1 F_{EU}(x) \, dx = \Phi_{EU}(1). \tag{28}$$

Since $p_{DA}^* \not\geq_{MPS} p_{EU}^*$, this yields that there exists $\bar{t} \in (0, 1)$ such that

$$\Phi_{DA}(\bar{t}) = \int_0^\bar{t} F_{DA}(x) \, dx > \int_0^\bar{t} F_{EU}(x) \, dx = \Phi_{EU}(\bar{t}). \tag{29}$$

Since $p_{DA}^* \geq_{MPS} r$ and $F_{EU}(x) = F_{r}(x)$ for all $x \in [0, x^*]$, we can conclude that

$$\Phi_{DA}(t) = \int_0^t F_{DA}(x) \, dx \leq \int_0^t F_{r}(x) \, dx = \int_0^t F_{EU}(x) \, dx = \Phi_{EU}(t) \quad \forall t \in [0, x^*].$$

This implies that $\bar{t}$ in (29) must belong to $(x^*, 1)$. Since $\Phi_{DA}$ and $\Phi_{EU}$ are continuous as well as $\Phi_{DA}(x^*) \leq \Phi_{EU}(x^*)$ and $\Phi_{DA}(\bar{t}) > \Phi_{EU}(\bar{t})$ for some $\bar{t} \in (x^*, 1)$, we have that

$$\hat{t} = \min \left\{ t \in [x^*, 1] : \Phi_{DA}(t) = \Phi_{EU}(t) \right\}$$

is well defined. Note that $\hat{t} > \hat{t} \geq x^*$ and $\hat{t} \leq \bar{y}$.\textsuperscript{24,25} By (28), we have that

$$\int_0^\hat{t} F_{DA}(x) \, dx + \int_\hat{t}^1 F_{DA}(x) \, dx = \Phi_{DA}(1) = \Phi_{EU}(1) = \int_0^\hat{t} F_{EU}(x) \, dx + \int_\hat{t}^1 F_{EU}(x) \, dx,$$

\textsuperscript{24} Otherwise, $\bar{t} \leq \hat{t}$. We have two cases:

1. $\bar{t} = \hat{t}$. In this case, we would have that

$$0 = \Phi_{DA}(\bar{t}) - \Phi_{EU}(\bar{t}) = \Phi_{DA}(\bar{t}) - \Phi_{EU}(\bar{t}) > 0,$$

a contradiction.

2. $\bar{t} < \hat{t}$. In this case, we would have that

$$\Phi_{DA}(x^*) - \Phi_{EU}(x^*) \leq 0 < \Phi_{DA}(\bar{t}) - \Phi_{EU}(\bar{t}).$$

Since $\Phi_{DA} - \Phi_{EU}$ is a continuous function on $[x^*, \bar{t})$, this would imply that there exists $\tilde{t} \in [x^*, \bar{t}] \subseteq [x^*, 1]$ such that

$$\Phi_{DA}(\tilde{t}) - \Phi_{EU}(\tilde{t}) = 0,$$

that is, $\tilde{t} \leq \bar{t} < \tilde{t} \leq \hat{t}$, a contradiction.

\textsuperscript{25} By contradiction, assume that $\hat{t} > \bar{y}$. Since $F_{EU}$ is such that $F_{EU}(x) = 1 \geq F_{DA}(x)$ for all $x \in [\bar{y}, 1]$. This would imply that

$$\Phi_{EU}(\hat{t}) = \int_0^\hat{t} F_{EU}(x) \, dx = \int_0^\hat{t} F_{EU}(x) \, dx + \int_\hat{t}^1 F_{EU}(x) \, dx = \Phi_{EU}(\hat{t}) + \int_\hat{t}^1 F_{EU}(x) \, dx$$

$$= \Phi_{DA}(\hat{t}) + \int_\hat{t}^1 F_{EU}(x) \, dx \geq \Phi_{DA}(\hat{t}) + \int_\hat{t}^1 F_{DA}(x) \, dx$$

$$= \int_0^\hat{t} F_{DA}(x) \, dx + \int_\hat{t}^\bar{t} F_{DA}(x) \, dx = \int_0^\hat{t} F_{DA}(x) \, dx = \Phi_{DA}(\hat{t}),$$

a contradiction with (29).
yielding that
\[ \int_{i}^{\bar{y}} F_{DA} (x) \, dx = \int_{i}^{\bar{y}} F_{EU} (x) \, dx. \]

Note that \( \bar{t} \) can be chosen to be \( \bar{y} \).\(^2\) This yields that
\[ 0 < \int_{i}^{\bar{y}} F_{DA} (x) \, dx - \int_{i}^{\bar{y}} F_{EU} (x) \, dx = \int_{i}^{\bar{y}} F_{EU} (x) \, dx - \int_{\bar{y}}^{1} F_{DA} (x) \, dx. \]

Since \( \int_{\bar{y}}^{1} F_{EU} (x) \, dx - \int_{i}^{\bar{y}} F_{DA} (x) \, dx > 0 \) and \( F_{EU} (x) = 1 \geq F_{DA} (x) \) for all \( x \in [\bar{y}, 1] \), we have that \( F_{DA} (\bar{y}) < 1 \) and, in particular, \( p_{DA} ((\bar{y}, 1]) > 0 \). Since \( e (p_{DA}^*) = E_r (X) < \bar{y} \), note also that \( 1 > p_{DA}^* ([\bar{y}, 1]) \geq p_{DA}^* ([\bar{y}, 1]) > 0 \). With this in mind, define \( q_{DA} (B) = p_{DA}^* (B \cap [0, \bar{y}]) / p_{DA}^* ([0, \bar{y}]) \) and \( r_{DA} (B) = p_{DA}^* (B \cap [\bar{y}, 1]) / p_{DA}^* ([\bar{y}, 1]) \) for all \( B \in \mathcal{B} \).

Since \( p_{DA}^* ([\bar{y}, 1]) \in (0, 1) \), \( q_{DA} \) and \( r_{DA} \) are well defined elements of \( \Delta \). Let also \( \lambda \) be \( p_{DA}^* ([0, \bar{y})) = 1 - p_{DA}^* ([\bar{y}, 1]) \in (0, 1) \). Note that \( p_{DA}^* = \lambda q_{DA} + (1 - \lambda) r_{DA} \) as well as \( e (q_{DA}) \in [0, \bar{y}) \) and \( e (r_{DA}) = y \in [\bar{y}, 1] \). We now have two cases:

\footnote{First, by (29) and since \( \Phi_{DA} (\bar{t}) = \Phi_{EU} (\bar{t}) \), observe that}

\[ \Phi_{DA} (\bar{t}) + \int_{i}^{\bar{t}} F_{DA} (x) \, dx = \int_{0}^{\bar{t}} F_{DA} (x) \, dx + \int_{i}^{\bar{t}} F_{DA} (x) \, dx = \int_{0}^{\bar{t}} F_{DA} (x) \, dx \]

\[ > \int_{0}^{\bar{t}} F_{EU} (x) \, dx = \int_{0}^{\bar{t}} F_{EU} (x) \, dx + \int_{\bar{t}}^{\bar{t}} F_{EU} (x) \, dx \]

\[ = \Phi_{EU} (\bar{t}) + \int_{i}^{\bar{t}} F_{EU} (x) \, dx, \]

yielding that
\[ \int_{i}^{\bar{t}} F_{DA} (x) \, dx > \int_{i}^{\bar{t}} F_{EU} (x) \, dx. \quad (30) \]

Since \( F_{EU} \) is such that \( F_{EU} (x) = 1 \geq F_{DA} (x) \) for all \( x \in [\bar{y}, 1] \), note that if \( \bar{t} \geq \bar{y} \), then
\[ 0 < \int_{i}^{\bar{t}} \left[ F_{DA} (x) - F_{EU} (x) \right] \, dx = \int_{i}^{\bar{y}} \left[ F_{DA} (x) - F_{EU} (x) \right] \, dx + \int_{\bar{t}}^{\bar{y}} \left[ F_{DA} (x) - F_{EU} (x) \right] \, dx \]
\[ \leq \int_{i}^{\bar{y}} \left[ F_{DA} (x) - F_{EU} (x) \right] \, dx. \]

Since \( F_{EU} (x) = F_{EU} (x^*) \) for all \( x \in [x^*, \bar{y}] \), vice versa, if \( \bar{t} < \bar{y} \), then it follows that \( F_{DA} (\bar{t}) > F_{EU} (\bar{t}) \).

It follows that \( F_{DA} (x) \geq F_{DA} (\bar{t}) > F_{EU} (\bar{t}) = F_{EU} (x^*) = F_{EU} (x) \) for all \( x \in [\bar{t}, \bar{y}] \subseteq [x^*, \bar{y}] \), yielding that
\[ \int_{i}^{\bar{y}} \left[ F_{DA} (x) - F_{EU} (x) \right] \, dx = \int_{i}^{\bar{t}} \left[ F_{DA} (x) - F_{EU} (x) \right] \, dx + \int_{\bar{t}}^{\bar{y}} \left[ F_{DA} (x) - F_{EU} (x) \right] \, dx > 0. \]

It follows that
\[ \int_{i}^{\bar{y}} F_{DA} (x) \, dx + \int_{\bar{y}}^{1} F_{DA} (x) \, dx = \int_{i}^{1} F_{DA} (x) \, dx = \int_{i}^{1} F_{EU} (x) \, dx = \int_{i}^{\bar{y}} F_{EU} (x) \, dx + \int_{\bar{y}}^{1} F_{EU} (x) \, dx. \]

25
imply that \( p^*_{DA} = \lambda q_{DA} + (1 - \lambda) r_{DA} = \lambda q_{DA} + (1 - \lambda) \delta_y \). In particular, we would have that \( F_{DA}(\bar{y}) = 1 \), a contradiction.

** \( y > \bar{y} \). Since \( \lambda \in (0, 1) \) and \( e(q_{DA}) < \bar{y} < e(r_{DA}) \), consider \( \varepsilon, \tau \in (0, \min \{\lambda, 1 - \lambda\}) \) such that \( \frac{\tau}{\tau + \varepsilon} e(q_{DA}) + \frac{\varepsilon}{\tau + \varepsilon} e(r_{DA}) = \bar{y} \). It follows that \( (\lambda - \tau), (1 - \lambda - \varepsilon) \in (0, 1) \). Consider the probability

\[
\hat{p}_{DA} = (\lambda - \tau) q_{DA} + (\tau + \varepsilon) \delta_y + (1 - \lambda - \varepsilon) r_{DA}.
\]

Since \( e\left(\frac{\tau}{\tau + \varepsilon} q_{DA} + \frac{\varepsilon}{\tau + \varepsilon} r_{DA}\right) = \bar{y} \), note that

\[
E_{\delta_y}(v) \geq E_{\frac{\tau}{\tau + \varepsilon} q_{DA} + \frac{\varepsilon}{\tau + \varepsilon} r_{DA}}(v) = \frac{\tau}{\tau + \varepsilon} E_{q_{DA}}(v) + \frac{\varepsilon}{\tau + \varepsilon} E_{r_{DA}}(v) \quad \forall v \in U_{conc}.
\]

This implies that

\[
E_{\hat{p}_{DA}}(v) = E_{(\lambda - \tau) q_{DA} + (\tau + \varepsilon) \delta_y + (1 - \lambda - \varepsilon) r_{DA}}(v) \\
= (\lambda - \tau) E_{q_{DA}}(v) + (\tau + \varepsilon) E_{\delta_y}(v) + (1 - \lambda - \varepsilon) E_{r_{DA}}(v) \\
= \lambda E_{q_{DA}}(v) + (1 - \lambda) E_{r_{DA}}(v) \\
+ (\tau + \varepsilon) \left( E_{\delta_y}(v) - \frac{\tau}{\tau + \varepsilon} E_{q_{DA}}(v) - \frac{\varepsilon}{\tau + \varepsilon} E_{r_{DA}}(v) \right) \\
\geq \lambda E_{q_{DA}}(v) + (1 - \lambda) E_{r_{DA}}(v) = E_{\hat{p}_{DA}}(v) \quad \forall v \in U_{conc}
\]

that is, \( \hat{p}_{DA} \geq_{MPS} p^*_{DA} \geq_{MPS} r \) and, in particular, \( \hat{p}_{DA} \in A \). Finally, note that the affine function \( \ell : [0, 1] \to \mathbb{R} \), defined by \( \ell(x) = (1 - \sigma) \bar{e}x \) for all \( x \in [0, 1] \), is such that

\[
\ell(x) > u(x) \quad \forall x \in (0, \bar{y}) \quad \text{and} \quad \ell(x) = u(x) \quad \forall x \in [\bar{y}, 1] \cup \{0\}.
\]
Note also that supp \( q_{DA} \neq \{0\}. \)\(^{27}\) This implies that

\[
u (\bar{y}) = \ell (\bar{y}) = \ell \left( e \left( \frac{\tau}{\tau + \varepsilon} q_{DA} + \frac{\varepsilon}{\tau + \varepsilon} r_{DA} \right) \right) = E_{\frac{\tau}{\tau + \varepsilon} q_{DA} + \frac{\varepsilon}{\tau + \varepsilon} r_{DA}} (\ell) = \frac{\tau}{\tau + \varepsilon} E_{q_{DA}} (\ell) + \frac{\varepsilon}{\tau + \varepsilon} E_{r_{DA}} (\ell) > \frac{\tau}{\tau + \varepsilon} E_{q_{DA}} (u) + \frac{\varepsilon}{\tau + \varepsilon} E_{r_{DA}} (u) = E_{\frac{\tau}{\tau + \varepsilon} q_{DA} + \frac{\varepsilon}{\tau + \varepsilon} r_{DA}} (u).
\]

Since \( \hat{v} \) is a continuous, strictly increasing, and concave transformation of \( u, \) we have that

\[
E_{\hat{u}} (\hat{v}) = \hat{u} (\bar{y}) > E_{\frac{\tau}{\tau + \varepsilon} q_{DA} + \frac{\varepsilon}{\tau + \varepsilon} r_{DA}} (\hat{v}) = \frac{\tau}{\tau + \varepsilon} E_{q_{DA}} (\hat{v}) + \frac{\varepsilon}{\tau + \varepsilon} E_{r_{DA}} (\hat{v}).
\]

We can conclude that

\[
E_{\hat{p}_{DA}} (\hat{v}) = E_{(\lambda - \tau) q_{DA} + (\tau + \varepsilon) \delta_{\bar{y}} + (1 - \lambda - \varepsilon) r_{DA}} (\hat{v}) = (\lambda - \tau) E_{q_{DA}} (\hat{v}) + (\tau + \varepsilon) E_{\delta_{\bar{y}}} (\hat{v}) + (1 - \lambda - \varepsilon) E_{r_{DA}} (\hat{v}) + \lambda E_{\hat{v}} + (1 - \lambda) E_{r_{DA}} (\hat{v}) + (\tau + \varepsilon) \left( E_{\delta_{\bar{y}}} (\hat{v}) - \frac{\tau}{\tau + \varepsilon} E_{q_{DA}} (\hat{v}) - \frac{\varepsilon}{\tau + \varepsilon} E_{r_{DA}} (\hat{v}) \right)
\]

a contradiction with (27).

Cases * and ** prove the statement. \( \blacksquare \)

We conclude by proving the results in Appendix A which are instrumental in proving the statement contained in Example 1.

\(^{27}\)If supp \( q_{DA} = \{0\}, \) there would exist \( k \in (0, 1) \) such that \( F_{DA} (x) = k \) for all \( x \in [0, \bar{y}] \). In this case, note that \( F_r (0) \geq F_{DA} (0) \). Otherwise, since \( F_r \) is continuous, if \( F_r (0) < F_{DA} (0) \), then there would exist \( \bar{x} \in [0, 1] \) such that \( F_r (x) < F_{DA} (0) \) for all \( x \in [0, \bar{x}] \). It would follow that \( F_r (x) < F_{DA} (0) \leq F_{DA} (x) \) for all \( x \in [0, \bar{x}] \), yielding that

\[
\int_0^{\bar{x}} F_r (x) \, dx < \int_0^{\bar{x}} F_{DA} (x) \, dx,
\]

a contradiction with \( p_{DA}^{*} \geq_{MPS} r \). Since \( F_{EU} (x) = F_r (x) \) for all \( x \in [0, x^*] \) and \( F_r (0) \geq F_{DA} (0) \), it follows that \( F_{EU} (x) = F_r (x) \geq F_r (0) \geq F_{DA} (0) = F_{DA} (x) \) for all \( x \in [0, x^*] \). Moreover, \( F_{EU} (x) = F_{EU} (x^*) \geq F_{DA} (0) = F_{DA} (x) \) for all \( x \in [x^*, \bar{y}] \). Finally, since \( F_{EU} (x) = 1 \geq F_{DA} (x) \) for all \( x \in [\bar{y}, 1] \), we can conclude that \( F_{EU} (x) \geq F_{DA} (x) \) for all \( x \in [0, 1] \). This would yield that

\[
\int_0^{x^*} F_{DA} (x) \, dx \leq \int_0^{x^*} F_{EU} (x) \, dx,
\]

a contradiction with (29).
Proof of Proposition 4. Before starting, define $V : \Delta \to \mathbb{R}$ by

$$V(p) = \inf_{v \in W_{bet}} c(p, v) \quad (\text{resp., } = \sup_{v \in W_{bet}} c(p, v)) \quad \forall p \in \Delta.$$ 

Define $v_t = k(\cdot, t)$ for all $t \in [0, 1]$. Denote also by $\Delta_0$ the subset of $\Delta$ of all simple lotteries (convex linear combinations of Dirac measures), that is, $\Delta_0 = \text{co} \left( \{ \delta_x \}_{x \in [w, b]} \right)$.

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

$$E_p(v_t) = t \implies c(p, v_t) \leq c(p, v_s) \quad (\text{resp., } \geq).$$

Proof of the Claim. Let $p \in \Delta_0$ and $E_p(v_t) = t$. 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 $i \in \{1, ..., n\}$. Since $E_p(v_t) = t$, this implies that

$$\sum_{i=1}^n \lambda_i t_i = \sum_{i=1}^n \lambda_i v_t(x_i) = E_p(v_t) = t.$$ 

Since $f_{s,t}$ is convex (resp., concave) at $t$, it follows that

$$f_{s,t}(E_p(v_t)) = f_{s,t}(t) \leq \sum_{i=1}^n \lambda_i f_{s,t}(t_i) = \sum_{i=1}^n \lambda_i f_{s,t}(v_t(x_i)) = \sum_{i=1}^n \lambda_i v_s(x_i) = E_p(v_s) \quad (\text{resp., } \geq).$$

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

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

$$\int_{[w, b]} k(x, \hat{V}(p)) \, dp = \hat{V}(p). \quad (31)$$

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 $\Delta$ and $\succ$ satisfies Weak Order, Continuity, and Strict First Order Stochastic Dominance, we have that there exists a sequence \{p_n\}_{n \in \mathbb{N}} \subseteq \Delta_0$ such that $p_n \sim p$ for all $n \in \mathbb{N}$ and $p_n \to p$. The condition $p_n \sim p$ yields that $E_{p_n}(v_t) = t$ for all $n \in \mathbb{N}$. By the previous claim and since \{p_n\}_{n \in \mathbb{N}} \subseteq \Delta_0$ and $f_{s,t}$ is convex (resp., concave) at $t$, this implies that $c(p_n, v_t) \leq c(p_n, v_s)$ (resp., $\geq$) for all $n \in \mathbb{N}$. By passing to the limit and since $s \in (0, 1)$ was arbitrarily chosen, we obtain that

$$c(p, v_t) \leq c(p, v_s) \quad (\text{resp., } \geq) \quad \forall s \in (0, 1). \quad (32)$$

We can conclude that

$$V(p) = \min_{s \in (0, 1)} c(p, v_s) = \min_{v \in W_{bet}} c(p, v) = c(p, v_t)$$

(resp., $V(p) = \max_{s \in (0, 1)} c(p, v_s) = \max_{v \in W_{bet}} c(p, v) = c(p, v_t)$).
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 \( p \sim \delta_x \),\(^2\) that is, \( \bar{x} = x_p \). Since \( p \in \Delta \setminus \{\delta_w, \delta_b\} \) was arbitrarily chosen, we obtain that \( V(p) = x_p \) for all \( p \in \Delta \).\(^3\) This implies that \( V \) is a utility representation of \( \succ \). Since \( \succ \) 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 (resp., Remark 1), this implies that \( \succ \) satisfies Negative Certainty Independence (resp., Positive Certainty Independence).

(ii) implies (i). By Theorem 4 (resp., Remark 1), we have that \( V : \Delta \rightarrow \mathbb{R} \), defined by

\[
V(p) = \min_{v \in W_{bet}} c(p, v) = \min_{s \in (0,1)} c(p, v_s) \quad \forall p \in \Delta
\]

(resp., \( V(p) = \max_{v \in W_{bet}} c(p, v) = \max_{s \in (0,1)} c(p, v_s) \quad \forall p \in \Delta \)),

is a continuous utility representation of \( \succ \). By contradiction, assume that there exist \( t \in (0,1) \) and \( s' \in (0,1) \) such that \( f_{s',t} \) is not convex (resp., concave) 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 \( f_{s',t}(t) > \sum_{i=1}^n \lambda_i f_{s',t}(t_i) \) (resp., \( < \)). 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 \( \succ \) is a Betweenness preference, this implies that \( p \sim \delta_{\bar{x}} \) where \( \bar{x} = c(p, v_t) \). In particular, this implies that \( x_p = \bar{x} \). At the same time, we also have that

\[
f_{s',t}(\mathbb{E}_p(v_t)) = f_{s',t}(t) > \sum_{i=1}^n \lambda_i f_{s',t}(t_i) = \sum_{i=1}^n \lambda_i f_{s',t}(v_t(x_i)) = \sum_{i=1}^n \lambda_i v_{s'}(x_i) = \mathbb{E}_p(v_{s'}) \quad (\text{resp., } <).
\]

\(^2\)Recall that \( p \in \Delta \setminus \{\delta_w, \delta_b\} \) and \( \hat{V}(p) \in (0,1) \) where the latter is the unique number in \([0,1]\) such that

\[
\int_{[w,b]} k\left(x, \hat{V}(p)\right) dp = \hat{V}(p).
\]

Recall also that \( v_t = k\left(x, \hat{V}(p)\right) \in W_{bet} \) and \( t = \hat{V}(p) \). Define \( \bar{x} \in [w, b] \) to be such that \( \bar{x} = c(p, v_t) \). Note that

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

It follows that

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

Since \( \succ \) is a Betweenness preference, we can conclude that \( \hat{V}(\delta_{\bar{x}}) = \hat{V}(p) \), that is, \( \delta_{\bar{x}} \sim p \) and so \( \bar{x} = x_p \).

\(^3\)Clearly, \( V(\delta_x) = x \) if either \( x = w \) or \( x = b \).
Since \( f_{s',t} = v_{s'} \circ v_t^{-1} \), we can conclude that

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

(resp., \( \max_{s \in (0,1)} c(p, v_s) = V(p) = x_p = c(p, v_t) < c(p, v_{s'}) \leq \max_{s \in (0,1)} c(p, v_s) \))

a contradiction. \( \blacksquare \)

**Proof of Remark 2.** Denote \( f_{s,t} \) simply by \( f \). Let \( t \in (0, 1) \). Assume that \( f : [0, 1] \to [0, 1] \) is such that \( \partial f(t) \neq \emptyset \). By assumption, it follows that there exists \( m \in \mathbb{R} \) such that

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

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

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

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 \( t \). \( \blacksquare \)

**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 \( v_t^{-1} : [0, 1] \to \mathbb{R} \)

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

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 \( x \in [0, 1] \)

\[
f(x) = v_s\left( v_t^{-1}(x) \right) = \begin{cases} 
\begin{array}{ll}
x & \text{if } x \leq s \\
x^2 - sx + s & \text{if } s < x \leq t
\end{array}
\end{cases} \quad \text{if } x \leq t
\]

Consider \( g : [0, 1] \to \mathbb{R} \) 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
\[ g(t') = m(t' - t) + f(t) \leq f(t) + t' - t \]
\[ = t' + f(t) - t \leq t' = f(t'). \]

(b) $s < t' \leq t$. Define $h : [0, 1] \to \mathbb{R}$ 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 $\tilde{h} : [t, 1] \to \mathbb{R}$ 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
\[ g(1) = m(1-t) + f(t) = (2t-s)(1-t) + t^2-st+s \]
\[ = 2t - 2t^2 - s + st + t^2 - st + s \]
\[ = 2t - t^2 = t + t(1-t) \leq 1 = \tilde{h}(1). \]

Since $0 < s < t < 1$, if $m = 1$, then
\[ g(1) - \tilde{h}(1) = g(1) - 1 = 1 - t + f(t) - 1 = 1 - t + t^2 - st + s - 1 \]
\[ = -t + t^2 - st + s = t(t-1) + s(1-t) \]
\[ = (t-s)(t-1) < 0, \]

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 \( \tilde{s} \) is such that \( \frac{t+\sqrt{t^2+4(\tilde{s}-t)}}{2} = s \).\(^{30}\) Consider \( g : [0,1] \to \mathbb{R} \) 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) \( t < t' \leq \tilde{s} \). Define \( h : [t,\tilde{s}] \to \mathbb{R} \) by \( h(x) = \frac{t+\sqrt{t^2+4(x-t)}}{2} \) for all \( x \in [t,\tilde{s}] \).

On the other hand, we have that \( \tilde{h}(\tilde{s}) = \frac{t+\sqrt{t^2+4(\tilde{s}-t)}}{2} = \frac{t+\sqrt{(2\tilde{s}-t)^2-4\tilde{t}^2}}{2} = \tilde{g}(\tilde{s}) \).

(c) \( t' > \tilde{s} \). Define \( \tilde{h} : [\tilde{s},1] \to \mathbb{R} \) 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 \( x \in [\tilde{s},1] \). Since \( \tilde{h} \) is convex, \( \tilde{h}(1) = 1 \), and \( \tilde{h}'(1) = \frac{2-s}{2-t} \in (0,1) \), we have that

\[
\tilde{h}(t') \geq \tilde{h}'(1)(t'-1) + \tilde{h}(1) \geq 1(t'-1) + \tilde{h}(1) = t' - 1 + 1 = t' = g(t') \quad \forall t' \in [\tilde{s},1].
\]

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

\[\blacksquare\]

References


\(^{30}\)Since \( \frac{t+\sqrt{t^2+4(\tilde{s}-t)}}{2} = t < s < 1 = \frac{t+\sqrt{t^2+4(1-t)}}{2} \) 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 \( \tilde{s} \) exists and \( \tilde{s} > t \).


