Cautious Expected Utility and the Certainty Effect*

Simone Cerreia-Vioglio† David Dillenberger‡ Pietro Ortoleva§

August 2014

Abstract

Many violations of the Independence axiom of Expected Utility can be traced to subjects' attraction to risk-free prospects. The key axiom in this paper, Negative Certainty Independence (Dillenberger, 2010), formalizes this tendency. Our main result is a utility representation of all preferences over monetary lotteries that satisfy Negative Certainty Independence together with basic rationality postulates. Such preferences can be represented as if the agent were unsure of how to evaluate a given lottery p; instead, she has in mind a set of possible utility functions over outcomes and displays a cautious behavior: she computes the certainty equivalent of p with respect to each possible function in the set and picks the smallest one. The set of utilities is unique in a well-defined sense. We show that our representation can also be derived from a 'cautious' completion of an incomplete preference relation.

JEL: *D80, D81*

Keywords: *Preferences under risk, Allais paradox, Negative Certainty Independence, Incomplete preferences, Cautious Completion, Multi-Utility representation.*

*We thank Faruk Gul, Yoram Halevy, Fabio Maccheroni, Massimo Marinacci, Efe Ok, Wolfgang Pesendorfer, Gil Riella, and Todd Sarver for very useful comments and suggestions. The co-editor and the referees provided valuable comments that improved the paper significantly. We greatly thank Selman Erol for his help in the early stages of the project. Cerreia-Vioglio gratefully acknowledges the financial support of ERC (advanced grant, BRSCDP-TEA). Ortoleva gratefully acknowledges the financial support of NSF grant SES-1156091.

†Department of Decision Sciences and IGIER, Università Bocconi. Email: simone.cerreia@unibocconi.it.

‡Department of Economics, University of Pennsylvania. Email: ddill@sas.upenn.edu.

§Department of Economics, Columbia University. Email: pietro.ortoleva@columbia.edu.

1 Introduction

Despite its ubiquitous presence in economic analysis, the paradigm of Expected Utility is often violated in choices between risky prospects. While such violations have been documented in many different experiments, a specific preference pattern emerges as one of the most prominent: the tendency of people to favor risk-free options – the so-called Certainty Effect (Kahneman and Tversky, 1979). This is shown in Allais' Common Ratio and Common Consequence paradoxes, as well as in many more recent experimental studies. (Section 6 discusses this evidence at length.)

Dillenberger (2010) suggests a way to define the Certainty Effect behaviorally, by introducing an axiom called *Negative Certainty Independence* (NCI). NCI states that for any two lotteries p and q, any number λ in $[0,1]$, and any lottery δ_x that yields the prize x for sure, if p is preferred to δ_x then $\lambda p + (1-\lambda) q$ is preferred to $\lambda \delta_x + (1-\lambda) q$. That is, if the sure outcome x is not enough to compensate the decision maker (henceforth DM) for the risky prospect p, then mixing it with any other lottery, thus eliminating its certainty appeal, will not result in the mixture of δ_x being preferred to the corresponding mixture of p. NCI is weaker than the Independence axiom, and in particular it permits Independence to fail when the Certainty Effect is present – allowing the DM to favor certainty, but ruling out the converse behavior.

In this paper we characterize the class of continuous, monotone, and complete preference relations, defined on lotteries over some interval of monetary prizes, that satisfy NCI. That is, we characterize a new class of preferences that accommodate the Certainty Effect, together with very basic rationality postulates. We show that any such preference relation admits the following representation: there exists a set $\mathcal{W}$ of strictly increasing (Bernoulli) utility functions over monetary outcomes, such that the value of any lottery p is given by

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

where $c(p,v)$ is the certainty equivalent of lottery p calculated using the utility function v. That is, if we denote by $\mathbb{E}_p(v)$ the expected utility of p with respect to v, then $c(p,v) = v^{-1}(\mathbb{E}_p(v))$.

We call this representation a Cautious Expected Utility representation and interpret it as follows. The DM acts as if she were unsure how to evaluate each given lottery: she does not have one, but a set of possible utility functions over monetary outcomes. She then reacts to this multiplicity using a form of *caution*: she evaluates each lottery according to the lowest possible certainty equivalent corresponding to some function in the set. Note that $c(\delta_x, v) = x$ for all $v \in \mathcal{W}$, which means that while the DM acts with caution when evaluating general lotteries, such caution does not play a role when evaluating degenerate ones. This captures the Certainty Effect.

The Cautious Expected Utility model is also linked with the notion of a completion of incomplete preferences. Consider a DM who has an incomplete preference relation over lotteries, which is well-behaved (that is, a reflexive, transitive, monotone, and continuous binary relation that satisfies Independence). The DM is asked to choose between two options that the original preference relation is unable to compare; she then needs to choose a rule to complete her ranking. Suppose that she follows what we call a Cautious Completion: if the original relation is unable to compare a lottery p with a degenerate lottery δ_x, then in the completion the DM opts for the latter – "when in doubt, go with certainty." We show that there always exists a unique Cautious Completion and that it admits a Cautious Expected Utility representation.[1]

We then discuss the relation of our model with the theoretical and empirical literature. We argue that our model suggests a useful way of interpreting existing empirical evidence and derives new theoretical predictions, especially emphasizing the following three points.

First, similarly to the most popular alternatives to Expected Utility (the Rank Dependent Utility (RDU) model of Quiggin, 1982 and the Betweenness class of Dekel, 1986 and Chew, 1989), our model accommodates the Certainty Effect by weakening the Independence axiom. In terms of generality, preferences in our class neither nest, nor are nested in, those that satisfy Betweenness, and are distinct from RDU, in the sense that RDU satisfies NCI only in the limiting case of Expected Utility. In terms of the key axiom delineating the departure from Expected Utility, while NCI is yet to be tested in its fullness, it is built on the Certainty Effect, which has solid empirical support. On the other hand, the analogous axioms (Comonotonic/Ordinal Independence and Betweenness) are frequently violated in experiments.

Second, our model can accommodate some evidence on the Certainty Effect (e.g., the presence of Allais-type behavior with large stakes but not with small ones), which

[1] In fact, we show that it admits a Cautious Expected Utility representation with the same set of utilities that can be used to represent the original incomplete preference relation (in the sense of the Expected Multi-Utility representation of Dubra et al., 2004).

poses difficulties to many popular alternative models.

Third, our model is consistent with the main stylized facts of preferences under risk as surveyed in Camerer (1995) and Starmer (2000). Moreover, some of these facts are predicted by our model without requiring any additional assumptions on the functional form.

The remainder of the paper is organized as follows. Section 2 presents the axiomatic structure, the main representation theorem, and the uniqueness properties of the representation. Section 3 characterizes risk attitudes and comparative risk aversion. Section 4 presents the result on the completion of incomplete preference relations. Section 5 surveys related theoretical models. Section 6 discusses the experimental evidence. All proofs appear in the Appendices.

2 The Model

2.1 Framework

Consider a compact interval $[w, b] \subset \mathbb{R}$ of monetary prizes. Let Δ 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 $[w, b]$ and by p, q, r generic elements of Δ. We denote by $\delta_x \in \Delta$ the degenerate lottery (Dirac measure at x) that gives the prize $x \in [w, b]$ with certainty. The primitive of our analysis is a binary relation $\succcurlyeq$ over Δ. The symmetric and asymmetric parts of $\succcurlyeq$ are denoted by $\sim$ and $\succ$, respectively. The certainty equivalent of a lottery $p \in \Delta$ is a prize $x_p \in [w, b]$ such that $\delta_{x_p} \sim p$.

We start by imposing the following basic axioms on $\succcurlyeq$.

Axiom 1 (Weak Order)**.** *The relation $\succcurlyeq$ is complete and transitive.*

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

Axiom 3 (Weak Monotonicity)**.** *For each $x, y \in [w, b]$, $x \geq y$ if and only if $\delta_x \succcurlyeq \delta_y$.*

The three axioms above are standard postulates. Weak Order is a common assumption of rationality. Continuity is needed to represent $\succcurlyeq$ through a continuous utility function. Finally, under the interpretation of Δ as monetary lotteries, Weak Monotonicity simply implies that more money is better than less.

2.2 Negative Certainty Independence (NCI)

Consider a version of Allais' paradox (called the Common Ratio Effect), in which subjects choose between $A = \delta_{3000}$ and $B = 0.8\delta_{4000} + 0.2\delta_0$, and also between $C = 0.25\delta_{3000} + 0.75\delta_0$ and $D = 0.2\delta_{4000} + 0.8\delta_0$. The typical finding is that the majority of subjects tend to *systematically* violate the Independence axiom of Expected Utility by choosing the pair A and D.[2,3] Kahneman and Tversky (1979) called this pattern of behavior the Certainty Effect. The next axiom, introduced in Dillenberger (2010), captures the Certainty Effect with the following relaxation of Independence.

Axiom 4 (Negative Certainty Independence)**.** *For each $p, q \in \Delta$, $x \in [w, b]$, and $\lambda \in [0, 1]$,*

$$p \succcurlyeq \delta_x \;\Rightarrow\; \lambda p + (1-\lambda)q \succcurlyeq \lambda\delta_x + (1-\lambda)q. \qquad \text{(NCI)}$$

NCI states that if the sure outcome x is not enough to compensate the DM for the risky prospect p, then mixing it with any other lottery, thus eliminating its certainty appeal, will not result in the mixture of x being more attractive than the corresponding mixture of p. In particular, x_p, the certainty equivalent of p, might not be enough to compensate for p when part of a mixture.[4] In this sense NCI captures the Certainty Effect. When applied to the Common Ratio experiment, NCI only posits that if B is chosen in the first problem, then D must be chosen in the second one. Specifically, it allows the DM to choose the pair A and D, in line with the typical pattern of choice. Coherently with this interpretation, NCI captures the Certainty Effect as defined by Kahneman and Tversky (1979) – except that, as opposed to the latter, NCI applies also to lotteries with three or more possible outcomes and can thus apply to broader evidence.[5] For example, the *exact* same theoretical restrictions can be used to address also Allais' Common Consequence Effect.[6]

[2]Recall that a binary relation $\succcurlyeq$ satisfies *Independence* if and only if for each $p, q, r \in \Delta$ and for each $\lambda \in (0, 1]$, we have $p \succcurlyeq q$ if and only if $\lambda p + (1-\lambda)\, r \succcurlyeq \lambda q + (1-\lambda)\, r$.

[3]This example is taken from Kahneman and Tversky (1979). Of 95 subjects, 80 percent chose A over B, 65 percent chose C over D, and more than half chose the pair (A, D). Note that prospects C and D are the 0.25:0.75 mixture of prospects A and B, respectively, with δ_0. This means that the only pairs of choices consistent with Expected Utility are (A, C) and (B, D).

[4]We show in Appendix A (Proposition 6) that our axioms imply that $\succcurlyeq$ preserves First Order Stochastic Dominance and thus for each lottery $p \in \Delta$ there exists a unique certainty equivalent x_p.

[5]Kahneman and Tversky (1979) define the Certainty Effect as the requirement that for $x, y \in [w, b]$ and $\alpha, \beta \in (0, 1)$, if $\alpha\delta_y + (1-\alpha)\,\delta_0$ is indifferent to δ_x, then $\alpha\beta\delta_y + (1-\alpha\beta)\,\delta_0$ is preferred to $\beta\delta_x + (1-\beta)\,\delta_0$. Note that this immediately follows from NCI.

[6]In Allais' Common Consequence Effect, subjects choose between $A = \delta_{1M}$ and $B = 0.1\delta_{5M} +$

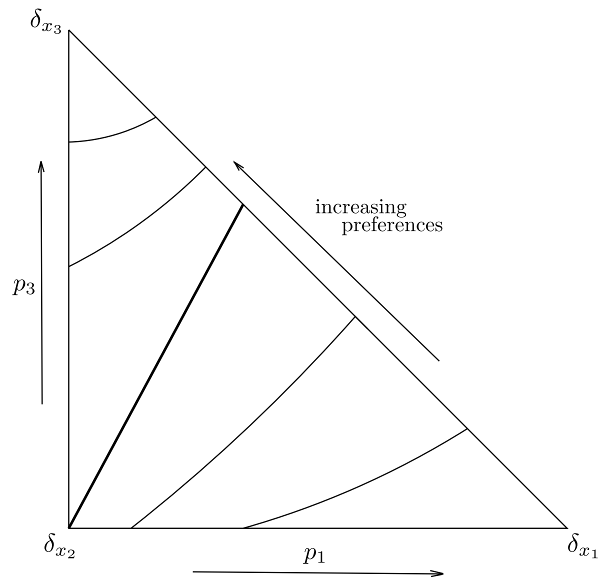

Figure 1: NCI and the Marschak-Machina triangle

Besides capturing the Certainty Effect, NCI puts additional structure on preferences over Δ. For example, NCI (in addition to the other basic axioms) implies Convexity: for each $p, q \in \Delta$, if $p \sim q$ then $\lambda p + (1-\lambda) q \succcurlyeq q$ for all $\lambda \in [0,1]$.[7] NCI thus suggests weak preference for randomization between indifferent lotteries. Similarly, NCI implies that if $p \sim \delta_x$ then $\lambda p + (1-\lambda)\delta_x \sim p$ for all $\lambda \in [0,1]$, which means neutrality towards mixing a lottery with its certainty equivalent.

To illustrate these restrictions, we now discuss the pattern of indifference curves in any Marschak-Machina triangle, which represents all lotteries over three fixed outcomes $x_3 > x_2 > x_1$ (see Figure 1). NCI implies three restrictions on these curves: (i) by Convexity, all curves must be *convex*; (ii) the bold indifference curve through the origin (which represents the lottery δ_{x_2}) is *linear*, due to neutrality towards mixing a lottery with its certainty equivalent; and (iii) this bold indifference curve is also the *steepest*, that is, its slope relative to the (p_1, p_3) coordinates exceeds that of any

$0.89\delta_{1M} + 0.01\delta_0$, and also between $C = 0.11\delta_{1M} + 0.89\delta_0$ and $D = 0.1\delta_{5M} + 0.9\delta_0$. The typical pattern of choice is the pair (A, D). It can be shown that when combined with Axioms 1–3, the only way to violate NCI is to choose the pair (B, C), and that any such violation of NCI has a corresponding violation in the Common Ratio Effect (and vice versa). Our focus on the Common Ratio Effect in motivating NCI is made mostly for explanatory purposes.

[7] To see this, assume $p \sim q$ and apply NCI twice to obtain that $\lambda p + (1-\lambda) q \succcurlyeq \lambda\delta_{x_p} + (1-\lambda) q \succcurlyeq \lambda\delta_{x_p} + (1-\lambda)\delta_{x_q} = \delta_{x_q} \sim q$.

other indifference curve in the triangle.[8] Since, as explained by Machina (1982), the slope of an indifference curve expresses local attitude towards risk (with greater slope corresponds to higher local risk aversion), property (iii) captures the Certainty Effect by, loosely speaking, requiring that local risk aversion is at its peak when it involves a degenerate lottery. In Section 6 we show that this pattern of indifference curves is consistent with a variety of experimental evidence on decision making under risk.

2.3 Representation Theorem

Before stating our representation theorem, we introduce some notation. We say that a function $V : \Delta \to \mathbb{R}$ represents $\succcurlyeq$ when $p \succcurlyeq q$ if and only if $V(p) \geq V(q)$. Denote by $\mathcal{U}$ the set of continuous and strictly increasing functions v from $[w, b]$ to $\mathbb{R}$. We endow $\mathcal{U}$ with the topology induced by the supnorm. For each lottery p and function $v \in \mathcal{U}$, $\mathbb{E}_p(v)$ denotes the expected utility of p with respect to v. The certainty equivalent of lottery p calculated using the utility function v is thus $c(p, v) = v^{-1}(\mathbb{E}_p(v)) \in [w, b]$.

Definition 1. *Let $\succcurlyeq$ be a binary relation on Δ and $\mathcal{W}$ a subset of $\mathcal{U}$. The set $\mathcal{W}$ is a Cautious Expected Utility representation of $\succcurlyeq$ if and only if the function $V : \Delta \to \mathbb{R}$, defined by*

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

represents $\succcurlyeq$. We say that $\mathcal{W}$ is a Continuous Cautious Expected Utility representation if and only if V is also continuous.

Theorem 1. *Let $\succcurlyeq$ be a binary relation on Δ. The following statements are equivalent:*

(i) The relation $\succcurlyeq$ satisfies Weak Order, Continuity, Weak Monotonicity, and Negative Certainty Independence;

(ii) There exists a Continuous Cautious Expected Utility representation of $\succcurlyeq$.

According to a Cautious Expected Utility representation, the DM has a set $\mathcal{W}$ of possible utility functions over monetary outcomes. While each of these functions is strictly increasing, i.e., agrees that "more money is better", these functions may have different curvatures: it is as if the DM is unsure how to evaluate each lottery. The

[8] The *steepest middle slope* property is formally derived in Lemma 3 of Dillenberger (2010).

DM reacts to this multiplicity with *caution*: she evaluates each lottery p by using the utility function that returns the lowest *certainty equivalent*.

As an example, suppose the DM needs to evaluate the lottery p that pays either \$0 or \$10,000, both equally likely. The DM may find it difficult to give a precise answer, but, instead, finds it conceivable that her certainty equivalent of p falls in the range $[\$3,500, \$4,500]$, and that this interval is tight (that is, the end points are plausible evaluations). This is the first component of the representation: the DM has a set of plausible valuations that she considers. Nevertheless, when asked how much she would be willing to pay in order to obtain p, she is cautious and answers *at most* \$3,500. This is the second component. Note that if $\mathcal{W}$ contains only one element then the model reduces to Expected Utility. Moreover, since each $v \in \mathcal{W}$ is strictly increasing, it preserves monotonicity with respect to First Order Stochastic Dominance.

An important feature of the representation in Theorem 1 is that the DM uses the utility function that minimizes the certainty equivalent of a lottery, instead of its expected utility. This leads to the Certainty Effect: while the DM acts with caution when evaluating general lotteries, caution does not play a role when evaluating degenerate ones – independently of the utility function being used, the certainty equivalent of receiving the prize x for sure is x.

Example 1. *Let* $[w, b] \subseteq [0, \infty)$ *and* $\mathcal{W} = \{u_1, u_2\}$*, where*

$$u_1(x) = -\exp(-\beta x), \ \beta > 0; \text{ and } u_2(x) = x^{\alpha}, \ \alpha \in (0, 1).$$

Consider the Common Ratio example from Section 2.2. Let $\alpha = 0.8$ *and* $\beta = 0.0002$*. We have* $V(B) = c(B, u_1) \simeq 2904 < 3000 = V(A)$*, but* $V(D) = c(D, u_2) \simeq 535 > 530 \simeq c(C, u_2) = V(C)$*. Thus* $A \succ B$ *but* $D \succ C$*.*

Example 1 shows that one could address experimental evidence related to the Certainty Effect using a set $\mathcal{W}$ that includes only two utility functions. The key feature that allows this is that there is no unique $v \in \mathcal{W}$ which minimizes the certainty equivalents for all lotteries; otherwise, only v would matter and behavior would coincide with Expected Utility.[9]

[9] This implies, for example, that if all utilities in $\mathcal{W}$ have constant relative risk aversion (that is, $v_i \in \mathcal{W}$ only if $v_i(x) = x^{\alpha_i}$ for some $\alpha_i \in (0, 1)$), then preferences will be indistinguishable from Expected Utility with coefficient of relative risk aversion equals $1 - \min_j \alpha_j$. In Section 2.6 we suggest some convenient parametric class of utility functions that can be used in applications.

The interpretation of the Cautious Expected Utility representation is different from some of the most prominent existing models of non-Expected Utility. For example, the common interpretation of Rank Dependent Utility is that the DM knows her utility function but she distorts probabilities. By contrast, in a Cautious Expected Utility representation the DM takes probabilities at face value, but she is unsure of which utility function to use and applies caution. Rather, the interpretation of our model is reminiscent of the Maxmin Expected Utility of Gilboa and Schmeidler (1989) under ambiguity, in which the DM has a set of probabilities and evaluates acts using the worst probability in the set. Our model can be seen as a corresponding model under risk. This analogy with Maxmin Expected Utility will be strengthened by our analysis in Section 4, where we argue that both models can be derived from extending incomplete preferences using a cautious rule.

Before proceeding, we introduce a notion that will play a major role in the subsequent analysis. Let $\succcurlyeq'$ be the largest subrelation of $\succcurlyeq$ that satisfies the Independence axiom. Formally, define $\succcurlyeq'$ on Δ by

$$p \succcurlyeq' q \iff \lambda p + (1-\lambda) r \succcurlyeq \lambda q + (1-\lambda) r \qquad \forall \lambda \in (0,1], \forall r \in \Delta. \tag{1}$$

In the context of choice under risk, this derived relation was proposed and characterized by Cerreia-Vioglio (2009). It parallels a notion introduced in the context of ambiguity by Ghirardato et al. (2004) (see also Cerreia-Vioglio et al., 2011). Since it satisfies Independence, $\succcurlyeq'$ is often interpreted as including the comparisons that the DM is confident in making. We refer to $\succcurlyeq'$ as the Linear Core of $\succcurlyeq$.

2.4 Proof Sketch of Theorem 1

In what follows we sketch the proof of Theorem 1; a complete proof appears in Appendix B. We focus here only on the sufficiency of the axioms for the representation.

Step 1. Define the Linear Core of $\succcurlyeq$ *.* As we have discussed above, we introduce the binary relation $\succcurlyeq'$ on Δ defined in (1).

Step 2. Find the set $\mathcal{W} \subseteq \mathcal{U}$ *that represents* $\succcurlyeq'$*.* By Cerreia-Vioglio (2009), $\succcurlyeq'$ is reflexive, transitive, continuous, and satisfies Independence. In particular, there exists a set $\mathcal{W}$ of continuous functions on $[w, b]$ that constitutes an Expected Multi-Utility representation of $\succcurlyeq'$, that is, $p \succcurlyeq' q$ if and only if $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for all $v \in \mathcal{W}$ (see Dubra et al., 2004). Since $\succcurlyeq$ satisfies Weak Monotonicity and NCI, $\succcurlyeq'$ also satisfies Weak Monotonicity. Thus, $\mathcal{W}$ can be chosen to be composed only of strictly increasing functions.

Step 3. Representation of $\succcurlyeq$. We show that $\succcurlyeq$ admits a certainty equivalent representation, i.e., there exists $V : \Delta \to \mathbb{R}$ such that V represents $\succcurlyeq$ and $V(\delta_x) = x$ for all $x \in [w, b]$.

Step 4. Relation between $\succcurlyeq$ and $\succcurlyeq'$. We note that *(i)* $\succcurlyeq$ is a completion of $\succcurlyeq'$, i.e., $p \succcurlyeq' q$ implies $p \succcurlyeq q$; and *(ii)* for each $p \in \Delta$ and for each $x \in [w, b]$, $p \not\succcurlyeq' \delta_x$ implies $\delta_x \succ p$. The latter is an immediate implication of NCI.

Step 5. Final step. We conclude the proof by showing that we must have $V(p) = \inf_{v\in\mathcal{W}} \; c(p, v)$ for all $p \in \Delta$. For each p, find $x \in [w, b]$ such that $p \sim \delta_x$, which means $V(p) = V(\delta_x) = x$. First note that we must have $V(p) = x \leq \inf_{v\in\mathcal{W}} c(p, v)$. If not, then we would have that $x > c(p, v)$ for some $v \in \mathcal{W}$, which means, by *Step 2*, $p \not\succcurlyeq' \delta_x$. But by *Step 4(ii)* we would obtain $\delta_x \succ p$, contradicting $\delta_x \sim p$. Second, we must have $V(p) = x \geq \inf_{v\in\mathcal{W}} c(p, v)$: if not, then we would have $x < \inf_{v\in\mathcal{W}} c(p, v)$. We could then find y such that $x < y < \inf_{v\in\mathcal{W}} c(p, v)$, which, by *Step 2*, would yield $p \succcurlyeq' \delta_y$. By *Step 4 (i),* we could conclude that $p \succcurlyeq \delta_y \succ \delta_x$, contradicting $p \sim \delta_x$.

2.5 Uniqueness and Properties of the Set of Utilities

We now discuss the uniqueness properties of the set $\mathcal{W}$ in a Cautious Expected Utility representation. We define the set of *normalized* utility functions $\mathcal{U}_{\text{nor}} = \{v \in \mathcal{U} : v(w) = 0,\ v(b) = 1\}$ and, without loss of generality, confine our attention to a normalized representation (that is, when $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$). We first show that even with this normalization, we are bound to find uniqueness properties only 'up to' the closed convex hull. Denote by $\overline{\text{co}}(\mathcal{W})$ the closed convex hull of a set $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$.

Proposition 1. *If $\mathcal{W}, \mathcal{W}' \subseteq \mathcal{U}_{\text{nor}}$ are such that $\overline{\text{co}}(\mathcal{W}) = \overline{\text{co}}(\mathcal{W}')$ then*

$$\inf_{v\in\mathcal{W}} c(p, v) = \inf_{v\in\mathcal{W}'} c(p, v) \qquad \forall p \in \Delta.$$

Moreover, $\mathcal{W}$ in general will not be unique, even up to the closed convex hull, as we can always add redundant utility functions that will never achieve the infimum. To see this, consider any set $\mathcal{W}$ in a Cautious Expected Utility representation and add to it a function $\bar{v}$ which is a continuous, strictly increasing, and strictly convex transformation of some other function $u \in \mathcal{W}$. The set $\mathcal{W} \cup \{\bar{v}\}$ will give a Cautious

Expected Utility representation of the same preference relation, as the function $\bar{v}$ will never be used in the representation.[10]

Once we remove these redundant utilities, we can identify a unique (up to the closed convex hull) set of utilities. In particular, for each preference relation that admits a Continuous Cautious Expected Utility representation, there exists a set $\widehat{\mathcal{W}}$ such that any other Cautious Expected Utility representation $\mathcal{W}$ of these preferences is such that $\overline{\text{co}}(\widehat{\mathcal{W}}) \subseteq \overline{\text{co}}(\mathcal{W})$. In this sense $\widehat{\mathcal{W}}$ is a 'minimal' set of utilities. Moreover, the set $\widehat{\mathcal{W}}$ will have a natural interpretation in our setup: it constitutes a unique (up to the closed convex hull) Expected Multi-Utility representation of the Linear Core $\succcurlyeq'$, the derived relation defined in (1).

Theorem 2. *Let $\succcurlyeq$ be a binary relation on Δ that satisfies Weak Order, Continuity, Weak Monotonicity, and Negative Certainty Independence. Then there exists $\widehat{\mathcal{W}} \subseteq \mathcal{U}_{\text{nor}}$ such that*

(i) The set $\widehat{\mathcal{W}}$ is a Continuous Cautious Expected Utility representation of $\succcurlyeq$;

(ii) If $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$ is a Cautious Expected Utility representation of $\succcurlyeq$ then $\overline{\text{co}}(\widehat{\mathcal{W}}) \subseteq \overline{\text{co}}(\mathcal{W})$;

(iii) The set $\widehat{\mathcal{W}}$ is an Expected Multi-Utility representation of $\succcurlyeq'$, that is,

$$p \succcurlyeq' q \Longleftrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \widehat{\mathcal{W}}.$$

Moreover, $\widehat{\mathcal{W}}$ is unique up to the closed convex hull.[11]

2.6 Parametric Sets of Utilities and Elicitation

In applied work, it is common to focus on parametric classes of utility functions. In this subsection we suggest two examples of parsimonious families of utility functions that are compatible with Cautious Expected Utility representation. We then remark on the issue of how to elicit the set of utilities from a finite data set.

[10] Since $u \in \mathcal{W}$ and $c(p,u) \leq c(p,\bar{v})$ for all $p \in \Delta$, there will not be a lottery p such that $\inf_{v \in \mathcal{W} \cup \{\bar{v}\}} c(p,v) = c(p,\bar{v}) < \inf_{v \in \mathcal{W}} c(p,v)$.

[11] Evren (2014) characterizes (possibly incomplete) preference relations that satisfy Independence and a form of continuity which is stronger than ours. He obtains a normalized Expected Multi-Utility representation with a compact set $\mathcal{W}$. To our knowledge, he was the first to study the notion of uniqueness of this representation up to the closed convex hull (see his Theorem 2).

The first example is the family of Expo-Power utility functions (Saha, 1993), which generalizes both constant absolute and constant relative risk aversion, given by

$$u(x) = 1 - \exp\left(-\lambda x^{\theta}\right), \text{ with } \lambda \neq 0,\ \theta \neq 0, \text{ and } \lambda\theta > 0.$$

This functional form has been applied in a variety of fields, such as finance, intertemporal choices, and agriculture economics. Holt and Laury (2002) argue that this functional form fits well experimental data that involve both low and high stakes. The second example is the set of Pareto utility functions, given by

$$u(x) = 1 - \left(1 + \frac{x}{\gamma}\right)^{-\kappa}, \text{ with } \gamma > 0 \text{ and } \kappa > 0.$$

Ikefuji et al. (2012) show that a Pareto utility function has some desirable properties.[12]

We conclude with a remark on the issue of elicitation. If one could observe the certainty equivalents for all lotteries, then the whole preference relation would be recovered and the set $\mathcal{W}$ identified (up to its uniqueness properties) – but this requires an infinite number of observations. With a finite data set, one can approximate, or partially recover, the set $\mathcal{W}$ as follows. Note that if a function v assigns to some lottery p a certainty equivalent *smaller* than the one observed in the data, then v cannot belong to $\mathcal{W}$. Therefore, by observing the certainty equivalents of a finite number of lotteries, one could exclude a set of possible utility functions and approximate the set $\mathcal{W}$ 'from above.' The set thus obtained would necessarily contain the 'true' one, and as the number of observations increases, it will shrink to coincide with $\mathcal{W}$ (or, more precisely, with a version of $\mathcal{W}$ up to uniqueness). Such elicitation would be significantly faster if, as is often the case in empirical work, one focuses on utility functions that come from a specific parametric class.

3 Cautious Expected Utility and Risk Attitudes

In this section we explore the connection between Theorem 1 and standard definitions of risk attitude, and characterize the comparative notion of "more risk averse than".

[12] If u is Pareto, then the coefficient of absolute risk aversion is $-\frac{u''(x)}{u'(x)} = \frac{\kappa+1}{x+\gamma}$, which is increasing in κ and decreasing in γ. Therefore, for a large enough interval $[w, b]$, if $\kappa_u > \kappa_v$ and $\gamma_u > \gamma_v$ then u and v are not ranked in terms of risk aversion.

Throughout this section, we mainly focus on a 'minimal' representation $\widehat{\mathcal{W}}$ as in Theorem 2.

Remark. *If $\mathcal{W}$ is a Continuous Cautious Expected Utility representation of a preference relation $\succcurlyeq$, we denote by $\widehat{\mathcal{W}}$ a set of utilities as identified in Theorem 2 (which is unique up to the closed convex hull). More formally, we can define a correspondence T that maps each set $\mathcal{W}$ that is a Continuous Cautious Expected Utility representation of some $\succcurlyeq$ to a class of subsets of $\mathcal{U}_{\text{nor}}$, $T(\mathcal{W})$, each element of which satisfies the properties of points (i)-(iii) of Theorem 2 and is denoted by $\widehat{\mathcal{W}}$.*

3.1 Characterization of Risk Attitudes

We adopt the following standard definition of risk aversion/risk seeking.

Definition 2. *We say that $\succcurlyeq$ is risk averse (resp., risk seeking) if $p \succcurlyeq q$ (resp., $q \succcurlyeq p$) whenever q is a mean preserving spread of p.*

Theorem 3. *Let $\succcurlyeq$ be a binary relation that satisfies Weak Order, Continuity, Weak Monotonicity, and Negative Certainty Independence. The following statements are true:*

(i) The relation $\succcurlyeq$ is risk averse if and only if each $v \in \widehat{\mathcal{W}}$ is concave.

(ii) The relation $\succcurlyeq$ is risk seeking if and only if each $v \in \widehat{\mathcal{W}}$ is convex.

Theorem 3 shows that the relationship found under Expected Utility between the concavity/convexity of the utility function and risk attitude holds also for the Continuous Cautious Expected Utility model – although it now involves *all* utilities in the set $\widehat{\mathcal{W}}$. In turn, this shows that our model is compatible with many types of risk attitudes. For example, despite the presence of the Certainty Effect, when all utilities are convex the DM would be risk seeking.

3.2 Comparative Risk Aversion

We now proceed to compare the risk attitudes of two individuals.

Definition 3. *Let $\succcurlyeq_1$ and $\succcurlyeq_2$ be two binary relations on Δ. We say that $\succcurlyeq_1$ is more risk averse than $\succcurlyeq_2$ if and only if for each $p \in \Delta$ and for each $x \in [w, b]$,*

$$p \succcurlyeq_1 \delta_x \quad \Longrightarrow \quad p \succcurlyeq_2 \delta_x.$$

Theorem 4. *Let $\succcurlyeq_1$ and $\succcurlyeq_2$ be two binary relations with Continuous Cautious Expected Utility representations, $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively. The following statements are equivalent:*

(i) $\succcurlyeq_1$ *is more risk averse than* $\succcurlyeq_2$*;*

(ii) Both $\mathcal{W}_1 \cup \mathcal{W}_2$ *and* $\mathcal{W}_1$ *are Continuous Cautious Expected Utility representations of* $\succcurlyeq_1$*;*

(iii) $\overline{\text{co}}\left(\widehat{\mathcal{W}_1 \cup \mathcal{W}_2}\right) = \overline{\text{co}}(\widehat{\mathcal{W}_1})$.

Theorem 4 states that DM1 is more risk averse than DM2 if and only if all the utilities in $\mathcal{W}_2$ are redundant when added to $\mathcal{W}_1$.[13,14] This result compounds two different channels that lead one DM to be more risk averse than another. The first is related to the *curvatures* of the functions in each set of utilities. For example, if each $v \in \mathcal{W}_2$ is a strictly increasing and strictly convex transformation of some $\hat{v} \in \mathcal{W}_1$, then DM2 assigns a strictly higher certainty equivalent than DM1 to any nondegenerate lottery $p \in \Delta$ (while the certain outcomes are, by construction, treated similarly in both). In particular, no member of $\mathcal{W}_2$ will be used in the representation corresponding to $\mathcal{W}_1 \cup \mathcal{W}_2$. The second channel corresponds to comparing the *size* of the two sets: if $\mathcal{W}_2 \subseteq \mathcal{W}_1$, then for each $p \in \Delta$ the certainty equivalent under $\mathcal{W}_2$ is weakly greater than that under $\mathcal{W}_1$, implying that $\succcurlyeq_1$ is more risk averse than $\succcurlyeq_2$.

To distinguish between these two different channels and characterize the behavioral underpinning of the second one, we focus on the notion of Linear Core and its representation as in Theorem 2.

Definition 4. *Let $\succcurlyeq_1$ and $\succcurlyeq_2$ be two binary relations on Δ with corresponding Linear Cores $\succcurlyeq_1'$ and $\succcurlyeq_2'$. We say that $\succcurlyeq_1$ is more indecisive than $\succcurlyeq_2$ if and only if for each $p, q \in \Delta$*

$$p \succcurlyeq_1' q \quad \Longrightarrow \quad p \succcurlyeq_2' q.$$

Since we interpret the derived binary relation $\succcurlyeq'$ as capturing the comparisons that the DM is confident in making, Definition 4 implies that DM1 is more indecisive than DM2 if whenever DM1 can confidently declare p weakly better than q, so does

[13] We thank Todd Sarver for suggesting point (iii) in Theorem 4.

[14] Note that if both $\succcurlyeq_1$ and $\succcurlyeq_2$ are Expected Utility preferences, then there are v_1 and v_2 such that $\{v_1\} = \mathcal{W}_1 = \widehat{\mathcal{W}}_1$, $\{v_2\} = \mathcal{W}_2 = \widehat{\mathcal{W}}_2$, and points (ii) and (iii) in Theorem 4 are equivalent to v_1 being a strictly increasing concave transformation of v_2.

DM2. The following result characterizes this comparative relation and links it to the comparative notion of risk aversion.

Proposition 2. *Let $\succcurlyeq_1$ and $\succcurlyeq_2$ be two binary relations that satisfy Weak Order, Continuity, Weak Monotonicity, and Negative Certainty Independence. The following statements are true:*

(i) $\succcurlyeq_1$ is more indecisive than $\succcurlyeq_2$ if and only if $\overline{\text{co}}(\widehat{\mathcal{W}_2}) \subseteq \overline{\text{co}}(\widehat{\mathcal{W}_1})$;

(ii) If $\succcurlyeq_1$ is more indecisive than $\succcurlyeq_2$ then $\succcurlyeq_1$ is more risk averse than $\succcurlyeq_2$.

4 Cautious Completions of Incomplete Preferences

Our analysis thus far has focused on the characterization of a complete preference relation that satisfies NCI (in addition to the other basic axioms). We now show that it is also related to that of a 'cautious' completion of incomplete preferences.

Consider a DM who has an *incomplete* preference relation over Δ. There might be occasions in which the DM is asked to choose among lotteries she cannot compare, and to do this she has to complete her preferences. Suppose that the DM wants to do so by applying *caution*: when in doubt between a sure outcome and a lottery, she opts for the sure outcome. Which preferences will she obtain as a completion? This analysis parallels the one of Gilboa et al. (2010), who consider an environment with ambiguity instead of risk.[15,16]

Since we study an incomplete preference relation, we require a slightly stronger notion of continuity, called Sequential Continuity (which coincides with our Continuity axiom if the binary relation is complete and transitive).

Axiom 5 (Sequential Continuity)**.** *Let $\{p_n\}_{n\in\mathbb{N}}$ and $\{q_n\}_{n\in\mathbb{N}}$ be two sequences in Δ. If $p_n \to p$, $q_n \to q$, and $p_n \succcurlyeq q_n$ for all $n \in \mathbb{N}$ then $p \succcurlyeq q$.*

[15] There is one minor formal difference: while in Gilboa et al. (2010) both the incomplete relation and its completion are primitives of the analysis, in our case the primitive is simply the incomplete preference relation over lotteries, and we study the properties of all possible completions of this kind.

[16] Riella (2013) develops a more general treatment that encompasses the result in this section and the one in Gilboa et al. (2010); he shows that a combined model could be obtained starting from a preference relation over acts that admits a Multi-Prior Expected Multi-Utility representation, as in Ok et al. (2012) and Galaabaatar and Karni (2013), and constructing a Cautious Completion.

In the rest of the section, we assume that $\succcurlyeq'$ is a reflexive and transitive binary relation over Δ, which satisfies Sequential Continuity, Weak Monotonicity, and Independence. We look for a Cautious Completion of $\succcurlyeq'$, which is defined as follows.

Definition 5. *Let $\succcurlyeq'$ be a binary relation on Δ. We say that the binary relation $\hat{\succcurlyeq}$ is a Cautious Completion of $\succcurlyeq'$ if and only if the following hold:*

1. *The relation $\hat{\succcurlyeq}$ satisfies Weak Order, Weak Monotonicity, and for each $p \in \Delta$ there exists $x \in [w, b]$ such that $p \hat{\sim} \delta_x$;*
2. *For each $p, q \in \Delta$, if $p \succcurlyeq' q$ then $p \hat{\succcurlyeq} q$;*
3. *For each $p \in \Delta$ and $x \in [w, b]$, if $p \not\succcurlyeq' \delta_x$ then $\delta_x \hat{\succ} p$.*

Point 1 imposes few minimal requirements of rationality on $\hat{\succcurlyeq}$ and the existence of a certainty equivalent for each lottery p. (Weak Monotonicity will imply that this certainty equivalent is unique.) In point 2, we assume that the relation $\hat{\succcurlyeq}$ extends $\succcurlyeq'$. Finally, point 3 requires that such a completion of $\succcurlyeq'$ is done with caution.

Theorem 5. *If $\succcurlyeq'$ is a reflexive and transitive binary relation on Δ that satisfies Sequential Continuity, Weak Monotonicity, and Independence, then $\succcurlyeq'$ admits a unique Cautious Completion $\hat{\succcurlyeq}$ and there exists a set $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$ such that for all $p, q \in \Delta$*

$$p \succcurlyeq' q \iff \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{W}$$

and

$$p \hat{\succcurlyeq} q \iff \inf_{v \in \mathcal{W}} c(p, v) \geq \inf_{v \in \mathcal{W}} c(q, v).$$

Moreover, $\mathcal{W}$ is unique up to the closed convex hull.

Theorem 5 shows that, given a binary relation $\succcurlyeq'$ which satisfies all the tenets of Expected Utility except completeness, there always exists a unique Cautious Completion $\hat{\succcurlyeq}$. Most importantly, such completion admits a Cautious Expected Utility representation, using the same set of utilities as in the Expected Multi-Utility representation of the original preference $\succcurlyeq'$. This shows that the Cautious Expected Utility model could also represent the behavior of a subject who might be unable to compare some of the available options and, when asked to extend her ranking, does so by being cautious. Together with Theorem 1, Theorem 5 shows that this behavior is indistinguishable from that of a subject who starts with a complete preference relation and satisfies Axioms 1–4.

Theorem 5 strengthens the link between our model and the Maxmin Expected Utility model of Gilboa and Schmeidler (1989). Gilboa et al. (2010) show that the latter could be derived as a completion of an incomplete preference relation over Anscombe-Aumann acts that satisfies the same assumptions as $\succcurlyeq'$ (adapted to their domain), by applying a form of caution according to which, when in doubt, the DM chooses a constant act. Similarly, here we derive the Cautious Expected Utility model by extending an incomplete preference over lotteries using a form of caution according to which, when in doubt, the DM chooses a risk-free lottery.[17]

5 Related Literature

Dillenberger (2010) introduces the NCI axiom and studies its implications in dynamic settings. Under specific assumptions on preferences over two-stage lotteries, he shows that NCI is a necessary and sufficient condition to a property called "preference for one-shot resolution of uncertainty." Dillenberger, however, does not provide a utility representation for preferences that satisfy NCI, as in Theorem 1.

Cerreia-Vioglio (2009) characterizes the class of continuous and complete preference relations that satisfy Convexity. Loosely speaking, Cerreia-Vioglio shows that there exists a set $\mathcal{W}$ of normalized Bernoulli utility functions, and a real function U on $\mathbb{R} \times \mathcal{W}$, such that preferences are represented by $V(p) = \inf_{v \in \mathcal{W}} U\left(\mathbb{E}_p\left(v\right), v\right)$. Using this representation, Cerreia-Vioglio interprets Convexity as a behavioral property that captures a preference for hedging that may arise in the face of uncertainty about the value of outcomes, future tastes, and/or the degree of risk aversion. He suggests the choice of the minimal certainty equivalent as a criterion to resolve uncertainty about risk attitudes and as a completion procedure. Since NCI implies Convexity (see Section 2.2), our model is a special case of his (from the representation, this can be seen by setting $U\left(\mathbb{E}_p\left(v\right), v\right) = v^{-1}(\mathbb{E}_p\left(v\right)) = c\left(p, v\right)$).

A popular generalization of Expected Utility is the Rank Dependent Utility (RDU) model of Quiggin (1982), also used within Cumulative Prospect Theory (Tversky and

[17]The condition in Gilboa et al. (2010) is termed Default to Certainty; Point 3 of Definition 5 is the translation of this condition to the context of choice under risk.

Kahneman, 1992). In this model, individuals weight probability in a nonlinear way using a distortion function $f:[0,1]\to[0,1]$, which is strictly increasing and onto.[18] If $f(p)=p$ then RDU reduces to Expected Utility. If f is convex, then larger weight is given to inferior outcomes, leading to a pessimistic probability distortion suitable to explain the Allais paradoxes. Apart from the different interpretation of RDU compared to our Cautious Expected Utility representation, as discussed in Section 2.3, the two models are behaviorally distinct: Dillenberger (2010) shows that RDU satisfies NCI only in the limiting case of Expected Utility.[19]

Another popular class of continuous and monotone preferences is the one introduced by Dekel (1986) and Chew (1989) based on the Betweenness axiom.[20] Under this postulate, indifference curves in the Marschak-Machina triangle are linear, but they are not necessarily parallel as in Expected Utility. One prominent example of such preferences is Gul (1991)'s model of Disappointment Aversion (denoted DA in Figure 2), which adds one parameter, $\beta\in(-1,\infty)$ to Expected Utility. Artstein-Avidan and Dillenberger (2011) show that Gul's preferences satisfy NCI if and only if $\beta\geq 0$; and Dillenberger and Erol (2013) provide an example of a continuous, monotone, and complete preference relation that satisfies NCI but not Betweenness. Thus, preferences in our class neither nest, nor are nested in, those that satisfy Betweenness.

Figure 2 summarizes our discussion thus far.[21]

Maccheroni (2002) (see also Chatterjee and Krishna, 2011) studies the following model: there exists a set $\mathcal{T}$ of utilities over outcomes, such that preferences are rep-

[18]If we order the prizes in the support of a finite lottery p, with $x_1<x_2<...<x_n$, then the functional form for RDU is:

$$V(p)=u(x_n)f(p(x_n))+\sum_{i=1}^{n-1}u(x_i)[f(\sum_{j=i}^{n}p(x_j))-f(\sum_{j=i+1}^{n}p(x_j))],$$

where $f:[0,1]\to[0,1]$ is strictly increasing and onto, and $u:[w,b]\to\mathbb{R}$ is increasing.

[19]Bell and Fishburn (2003) show that Expected Utility is the only RDU with the property that for each binary lottery p and $x\in[w,b]$, $p\sim\delta_x$ implies $\alpha p+(1-\alpha)\delta_x\sim\delta_x$. This property is implied by NCI (see Section 2.2). Geometrically, it corresponds to the linear indifference curve through the origin in any Marschak-Machina triangle (Figure 1 in Section 2.2).

[20]The Betweenness axiom states that for each $p,q\in\Delta$ and $\lambda\in(0,1)$, $p\succ q$ (resp., $p\sim q$) implies $p\succ\lambda p+(1-\lambda)q\succ q$ (resp., $p\sim\lambda p+(1-\lambda)q\sim q$).

[21]Chew and Epstein (1989) show that there is no intersection between RDU and Betweenness other than Expected Utility (see also Bell and Fishburn, 2003). Whether or not RDU satisfies Convexity depends on the curvature of the distortion function f; in particular, concave f implies Convexity. In addition to Disappointment Aversion with negative β, an example of preferences that satisfy Betweenness but do not satisfy NCI is Chew (1983)'s model of Weighted Utility.

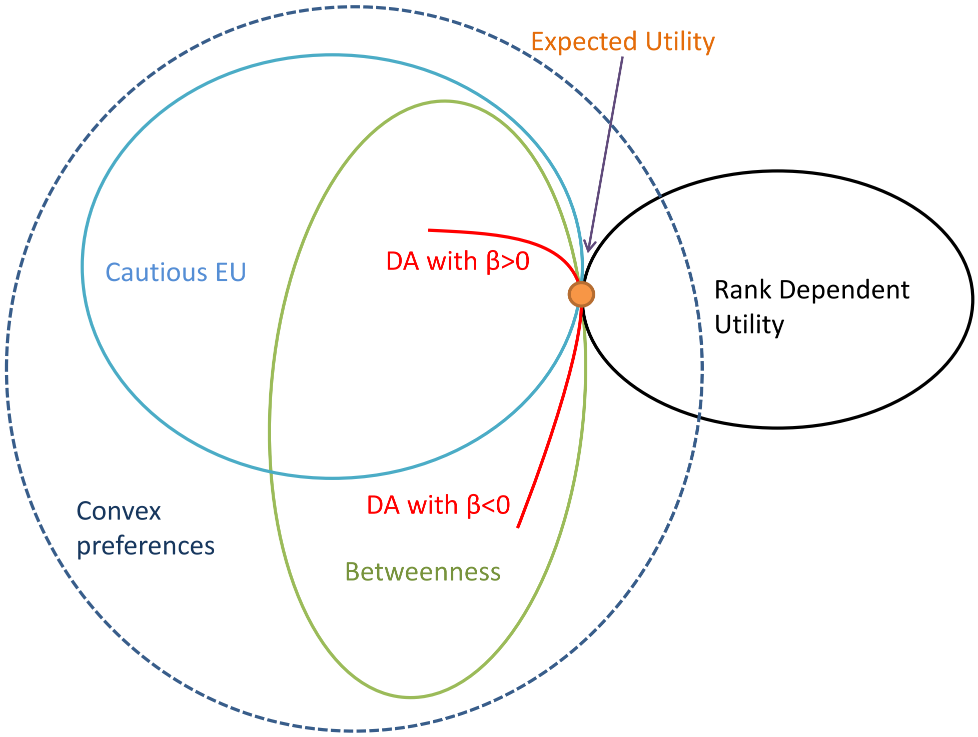

Figure 2: Cautious Expected Utility and other models

resented by $V(p) = \min_{v \in \mathcal{T}} \mathbb{E}_p(v)$. Maccheroni's interpretation that "the most pessimist of [the DM] selves gets the upper hand over the others," is closely related to the interpretation of our functional form. In addition, both models satisfy Convexity. There are, however, two main differences between the two models. First, Maccheroni's model cannot (and was not meant to) address the Certainty Effect: since certainty equivalents are not used, also degenerate lotteries have multiple evaluations. Second, besides Convexity, Maccheroni's other key axiom, which requires Independence when mixing lotteries with the best possible outcome (b), is distinct from NCI.

Schmidt (1998) develops a model in which the value of any nondegenerate lottery p is $\mathbb{E}_p(u)$, whereas the value of the degenerate lottery δ_x is $v(x)$. The Certainty Effect is captured by requiring $v(x) > u(x)$ for all x. Schmidt's model violates both Continuity and Monotonicity, while we confine our attention to preferences that satisfy both of these basic properties. In addition, his model also violates NCI.[22] Other discontinuous

[22]For example, take $u(x) = x$ and $v(x) = 2x$, and note that $V(\delta_3) = 6 > 4 = V(\delta_2)$, but $V(\delta_2) = 4 > 2.5 = V(0.5\delta_3 + 0.5\delta_2)$. (The statement in Dillenberger and Erol (2013) that Schmidt's model satisfies NCI is incorrect.)

specifications of the Certainty Effect include, for example, Gilboa (1988) and Jaffray (1988), which are models of 'expected utility with security level'.

Dean and Ortoleva (2014) present a model that, when restricted to preferences over lotteries, generalizes pessimistic RDU. In their model, the DM has one utility function and a set of convex probability distortion functions; she then evaluates each lottery using the most pessimistic of these distortions. The exact relation between their model and ours remains an open question.

Machina (1982) studies a model with minimal restrictions imposed on preferences apart from requiring them to be smooth (Fréchet differentiable). One of the main behavioral assumptions proposed by Machina is Hypothesis II, which implies that indifference curves in the Marschak-Machina triangle *Fan Out*, that is, they become steeper as one moves in the north-west direction. The steepest middle slope property (Section 2.2) implies that our model can accommodate Fanning Out in the lower-right part of the triangle (from where most evidence on Allais-type behavior had come), while global Fanning Out is ruled out.

6 Experimental Evidence

In this section we discuss the experimental evidence pertaining to the two Allais paradoxes– the Common Ratio and Common Consequence effects– and more broadly the empirical regularities of preferences under risk. In light of this evidence, we evaluate the fit of our model and that of the two most popular alternatives to Expected Utility, RDU and Betweenness, which are also consistent with Allais-type behavior.[23] Our discussion relies on the surveys in Camerer (1995) and Starmer (2000), and on more recent findings.

The presence of Allais-type behavior when one alternative in the choice set is risk-free is extensively documented (see Camerer, 1995, Section C1). A natural question is whether violations of Expected Utility in these experiments depend on the special nature of certainty, or whether they would be exhibited also when certainty is not

[23] In reviewing the evidence related to the Common Ratio and Common Consequence effects, we find it useful to discuss them together. For the purpose of our paper this is inconsequential, since our model puts the exact same restrictions on behavior in both problems. Moreover, the two phenomena are often discussed either interchangeably or jointly to emphasize specific issues; for example, even though the term Certainty Effect was first referred to as a special case of the Common Ratio Effect, in more recent literature it is typically understood as any violation of Expected Utility resulting from experiments that include one lottery that is risk-free.

involved. Some early evidence initially suggested the latter (e.g., Kahneman and Tversky, 1979 and MacCrimmon and Larsson, 1979; see Machina, 1987 for a survey).[24] A second wave of empirical studies, however, provide evidence that certainty does matter. The results in Cohen and Jaffray (1988), Conlisk (1989), Camerer (1992), Sopher and Gigliotti (1993), Harless and Camerer (1994), and Humphrey (2000), all indicate that Allais-type violations of Expected Utility are much *less* frequent when the safe option is moved away from certainty and, more generally, when only non-degenerate lotteries over common outcomes are involved.[25] This pattern is documented as one of the key stylized facts of preferences under risk. Some recent evidence even argues that non-Expected Utility behavior completely disappears unless certainty is involved (see Andreoni and Harbaugh, 2010, and Andreoni and Sprenger, 2010, 2012).[26]

The Cautious Expected Utility model is compatible with more violations of Expected Utility as we approach the boundaries of the Marschak-Machina triangle. Moreover, it *implies* that Allais-type violations cannot be more prominent when a risk-free prospect is not involved than when it is. This is formalized in the following proposition (the proof is an immediate implication of NCI and is omitted).

Proposition 3. *Let $\succcurlyeq$ be a binary relation on Δ that satisfies Weak Order, Continuity, Weak Monotonicity, and Negative Certainty Independence. Fix $x_3 > x_2 > x_1$, and for $\alpha, \beta, \gamma \in [0,1]$ consider the lotteries $A = \delta_{x_2}$; $B = \alpha\delta_{x3} + (1-\alpha)\,\delta_{x_1}$; $C = \beta A + (1-\beta)\,\delta_{x_1}$; $D = \beta B + (1-\beta)\,\delta_{x_1}$; $E = \gamma C + (1-\gamma)\,\delta_{x_1}$; and $F = \gamma D + (1-\gamma)\,\delta_{x_1}$. Then*

$$C \succ D \;\; and \;\; E \prec F \quad \Rightarrow \quad A \succ B$$

In words, if we observe a Common-Ratio violation for the pairs CD and EF, then we must also observe such violation for AB and EF, when A is a risk-free lottery.

Note that without further assumptions, our model does not rule out the opposite of the Common Ratio Effect away from certainty, as long as Continuity is preserved

[24]For example, Kahneman and Tversky found that 86 percent of their subjects chose $0.9\delta_{3000} + 0.1\delta_0$ over $0.45\delta_{6000} + 0.55\delta_0$, while 73 percent chose $0.01\delta_{6000} + 0.99\delta_0$ over $0.02\delta_{3000} + 0.98\delta_0$. This is an instance of the Common Ratio Effect since the ratio of the probabilities to receive 3000 and 6000 is the same in both problems (equals $\frac{0.9}{0.45} = \frac{0.02}{0.01}$).

[25]For example, in Conlisk (1989) the fraction of Expected Utility violations drops from about 50 percent in the basic version that includes one degenerate lottery to about 32 percent in the 'displaced Allais version', where each of the new prospects puts strictly positive probability on all prizes.

[26]On the other hand, some studies found a 'reverse' Certainty Effect, which is clearly incompatible with our model. See Wu et al. (2005) and Blavatskyy (2010).

(e.g., it permits $D \succ C$ and $F \prec E$, at least for small enough values of β). This is also allowed by both RDU and Betweenness; in their most general forms, these models put no restrictions on the behaviors just discussed (for example, they are compatible with fewer violations of Expected Utility away from certainty and with the absence of reverse Common Ratio Effect, but are consistent with the opposite patterns as well).

Many studies investigated whether Allais' results, and the pattern of indifference curves they suggest, persist when constructed using different kind of mixtures, e.g., when computed in the upper part of the Marschak-Machina triangle as opposed to the lower part (from where most early evidence had come). A robust finding is that indifference curves in the triangle exhibit *Mixed Fanning*: they become first steeper (Fanning Out) and then flatter (Fanning In) as we move in the north-west direction. In particular, global Fanning Out is inconsistent with the available data.[27] As we noted in Section 5, our model is compatible with Mixed Fanning and rules out global Fanning Out. Betweenness is compatible with both Mixed Fanning and Fanning Out. RDU rules out global Fanning Out (see Röell, 1987).[28]

Another robust finding is that Allais-type behavior is significantly less frequent when stakes are small rather than large.[29] Our model is compatible with this evidence: for instance, this would be the case if, as in Example 1, one of the utility functions in $\mathcal{W}$ is the most risk averse for a range of outcomes below a threshold.[30] Models based on Betweenness are also consistent with it. However, this is not the case for RDU. As its name suggests, in that model only the ranks of outcomes within the support of a lottery matter for the probability distortion, not their sizes. Thus, the presence of Allais-type behavior should be independent of the stakes.[31]

The behavioral patterns we have mentioned so far are documented for lotteries

[27] Chew and Waller (1986); Camerer (1989); Conlisk (1989); Battalio et al. (1990); Prelec (1990); Sopher and Gigliotti (1993); Wu (1994).

[28] RDU is compatible with many fanning patterns of indifference curves, but is inconsistent with Mixed Fanning in the strict sense (see Wu and Gonzalez, 1998).

[29] Conlisk (1989); Camerer (1989); Burke et al. (1996); Fan (2002); Huck and Müller (2012); Agranov and Ortoleva (2014). Some of these studies also found that Expected Utility violations are less prominent in experiments that include real rather than hypothetical payoffs.

[30] In Example 1, u_2 has a higher Arrow-Pratt coefficient of (absolute) risk aversion than u_1 for all outcomes below $\frac{1-\alpha}{\beta}$. Therefore, when restricted to lotteries with support in $\left[w, \frac{1-\alpha}{\beta}\right]$, preferences are Expected Utility with Bernoulli index u_2, but they violate Expected Utility for larger stakes.

[31] RDU implies that if we detect an Allais-type violation of Expected Utility in some range of prizes, e.g., with $x_1 < x_2 < x_3$, then similar violations of Expected Utility can be produced in any range of prizes. That is, for any $y_1 < y_3$ there exists $y_2 \in (y_1, y_3)$ and $a, b \in (0,1)$ such that $\delta_{y_2} \succ a\delta_{y_3} + (1-a)\,\delta_{y_1}$ but $b\delta_{y_2} + (1-b)\,\delta_{y_1} \prec ab\delta_{y_3} + (1-ab)\,\delta_{y_1}$.

involving only positive outcomes (gains). As it is well known, behavior may be different when losses are involved (Camerer, 1995). For example, individuals are typically risk averse with respect to gains, yet risk seeking with respect to losses – the so called Reflection Effect. Just like RDU or Betweenness, our model is not designed to distinguish between gains and losses. Yet it is consistent with risk aversion over gains and risk loving over losses: for example, if each $v \in \mathcal{W}$ is such that v is concave (resp., convex) for positive (resp., negative) outcomes. Some studies also document the opposite of Common Ratio Effect for losses (Kahneman and Tversky, 1979). As we have mentioned earlier, our model, independently of the sign of the stakes, is inconsistent with the opposite of the Certainty Effect but can accommodate local violations (away from certainty) of the Common Ratio Effect.

With respect to their main behavioral underpinnings, a stylized fact is that indifference curves are typically nonlinear, thus directly violating the Betweenness axiom.[32] In addition, there is also evidence of frequent violations of RDU's key axiom, Comonotonic/Ordinal Independence.[33] (We discuss tests of NCI in detail below.)

Overall, it appears that our model does at least as well as leading alternative models in accommodating most prominent existing evidence. And it is based on an axiom that captures the well documented Certainty Effect, instead of ancillary assumptions that are easily challenged by experiments, such as Betweenness or Comonotonic/Ordinal Independence.

We conclude with two remarks. The first one pertains to the flexibility of the Cautious Expected Utility model. It would be less appealing if our model could better fit the experimental evidence simply because it is more permissive than existing ones, as this would mean that it has less predictive power. We argue that this is not case. First, as we have seen in Section 5, the two main alternative models and ours are not nested, thus formally our model is not more permissive. Indeed, in some instances the models predictions are incomparable as, for example, with respect to the fanning properties of indifference curves. In other instances, our model has specific behavioral predictions that are *more* restrictive than those of existing ones. In the case of behavior at or away from certainty as stated in Proposition 3, for example, this additional restriction is in line with the experimental data.

On the other hand, our model is indeed permissive. For example, it generalizes

[32] Chew and Waller (1986); Bernasconi (1994); Camerer and Ho (1994); Prelec (1990).

[33] Wu (1994); Wakker et al. (1994).

Gul (1991)'s model of disappointment aversion (with $\beta > 0$). This comes from the fact that the NCI axiom captures the Certainty Effect but regulates less other aspects of behavior. Some existing models, again such as Gul (1991), may provide sharper predictions (e.g., imply 'more' Common-Ratio violations away from certainty),[34] but also rely on assumptions that have poor empirical performance, such as Betweennes. In this regard, our model is less restrictive – and thus have less predictive power – than some existing alternatives, but at the same time relies on minimal assumptions and provides a more flexible framework to applied researchers.

As a second remark, we note that the Cautious Expected Utility model has additional behavioral implications which may or not find empirical support, and that have not been subject to similar scrutiny yet. In general, while consistent with many of the findings on the Certainty Effect, to our knowledge no comprehensive tests of NCI have been conducted thus far. Due to the simplicity of the axiom, such tests should be easy to implement, and could focus either on testing the axiom itself, or on testing some of its implications when combined with the other three axioms. For example, while most of the evidence on the Certainty Effect we surveyed above is based on lotteries with at most three outcomes, NCI directly suggests that the phenomenon should be invariant to the number of prizes. In addition, NCI implies that an individual should be indifferent to any mixing between a lottery and its certainty equivalent (recall that the indifference curve through the middle outcome in any Marschak-Machina triangle must be linear). Lastly, NCI implies (weak) Convexity of preferences. Convexity has been tested experimentally, albeit possibly with smaller scrutiny. The existing evidence is mixed: while the experimental papers that document violations of Betweenness found deviations in both directions (that is, either preference or aversion to mixing), both Sopher and Narramore (2000) and Dwenger et al. (2013) find explicit evidence in support of Convexity.[35]

[34]Formally, in Gul's model (with $\beta > 0$), Proposition 3 can be strengthen: under the same assumptions, we must have that $C \succ D$ and $E \prec F$ imply $A'_\lambda \succ B'_\lambda$ for all $\lambda \in [0,1]$, where $A'_\lambda = \lambda A + (1-\lambda)C$ and $B'_\lambda = \lambda B + (1-\lambda)D$.

[35]Convexity is another dimension in which our model differs from a popular version of RDU used to address Allais-type behavior. In particular, RDU with a convex distortion f displays *aversion* to mixing (i.e., *lower*, and not upper, contour sets are convex).

Appendix A: Preliminary Results

We begin by proving some preliminary results that will be useful for the proofs of the main results in the text. In the sequel, we denote by $C([w,b])$ the set of all real valued continuous functions on $[w,b]$. Unless otherwise specified, we endow $C([w,b])$ with the topology induced by the supnorm. We denote by $\Delta = \Delta([w,b])$ the set of all Borel probability measures endowed with the topology of weak convergence. We denote by Δ_0 the subset of Δ which contains only the elements with finite support. Since $[w,b]$ is closed and bounded, Δ is compact with respect to this topology and Δ_0 is dense in Δ. Given a binary relation $\succcurlyeq$ on Δ, we define an auxiliary binary relation $\succcurlyeq'$ on Δ by

$$p \succcurlyeq' q \Longleftrightarrow \lambda p + (1-\lambda) r \succcurlyeq \lambda q + (1-\lambda) r \qquad \forall \lambda \in (0,1], \forall r \in \Delta.$$

Lemma 1. *Let $\succcurlyeq$ be a binary relation on Δ that satisfies Weak Order. The following statements are true:*

1. *The relation $\succcurlyeq$ satisfies Negative Certainty Independence if and only if for each $p \in \Delta$ and for each $x \in [w,b]$*

$$p \succcurlyeq \delta_x \Longrightarrow p \succcurlyeq' \delta_x. \qquad \text{(Equivalently } p \not\succcurlyeq' \delta_x \Longrightarrow \delta_x \succ p.\text{)}$$

2. *If $\succcurlyeq$ also satisfies Negative Certainty Independence then $\succcurlyeq$ satisfies Weak Monotonicity if and only if for each $x, y \in [w,b]$*

$$x \geq y \Longleftrightarrow \delta_x \succcurlyeq' \delta_y,$$

that is, $\succcurlyeq'$ satisfies Weak Monotonicity.

Proof. It follows from the definition of $\succcurlyeq'$.

We define

$$\begin{aligned}
\mathcal{V}_{in} &= \{v \in C([w,b]) : v \text{ is increasing}\}, \\
\mathcal{V}_{inco} &= \{v \in C([w,b]) : v \text{ is increasing and concave}\}, \\
\mathcal{U} = \mathcal{V}_{s-in} &= \{v \in C([w,b]) : v \text{ is strictly increasing}\}, \\
\mathcal{U}_{\text{nor}} &= \{v \in C([w,b]) : v(b) - 1 = 0 = v(w)\} \cap \mathcal{V}_{s-in}.
\end{aligned}$$

Consider a binary relation $\succsim^*$ on Δ such that

$$p \succsim^* q \Longleftrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{W} \tag{2}$$

where $\mathcal{W}$ is a subset of $C([w,b])$. Define $\mathcal{W}_{\max}$ as the set of all functions $v \in C([w,b])$ such that $p \succsim^* q$ implies $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$. Define also $\mathcal{W}_{\max-\text{nor}}$ as the set of all functions $v \in \mathcal{U}_{\text{nor}}$ such that $p \succsim^* q$ implies $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$. Clearly, we have that $\mathcal{W}_{\max-\text{nor}} = \mathcal{W}_{\max} \cap \mathcal{U}_{\text{nor}}$ and $\mathcal{W}_{\max-\text{nor}}, \mathcal{W} \subseteq \mathcal{W}_{\max}$.

Proposition 4. *Let $\succsim^*$ be a binary relation represented as in (2) and such that $x \geq y$ if and only if $\delta_x \succsim^* \delta_y$. The following statements are true:*

1. *$\mathcal{W}_{\max}$ and $\mathcal{W}_{\max-\text{nor}}$ are convex and $\mathcal{W}_{\max}$ is closed;*
2. *$\emptyset \neq \mathcal{W}_{\max-\text{nor}}$;*
3. *$\mathcal{W}_{\max} \subseteq \mathcal{V}_{in}$, $\emptyset \neq \mathcal{W}_{\max} \cap \mathcal{V}_{s-in}$, and $cl(\mathcal{W}_{\max} \cap \mathcal{V}_{s-in}) = \mathcal{W}_{\max}$;*
4. *$p \succsim^* q$ if and only if $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for each $v \in \mathcal{W}_{\max-\text{nor}}$;*
5. *If $\mathcal{W}$ is a convex subset of $\mathcal{U}_{\text{nor}}$ that satisfies (2) then $cl(\mathcal{W}) = cl(\mathcal{W}_{\max-\text{nor}})$.*

Proof. 1. Consider $v_1, v_2 \in \mathcal{W}_{\max-\text{nor}}$ (resp., $v_1, v_2 \in \mathcal{W}_{\max}$) and $\lambda \in (0,1)$. Since both functions are continuous, strictly increasing, and normalized (resp., continuous), it follows that $\lambda v_1 + (1-\lambda) v_2$ is continuous, strictly increasing, and normalized (resp., continuous). Since $v_1, v_2 \in \mathcal{W}_{\max-\text{nor}}$ (resp., $v_1, v_2 \in \mathcal{W}_{\max}$), if $p \succsim^* q$ then $\mathbb{E}_p(v_1) \geq \mathbb{E}_q(v_1)$ and $\mathbb{E}_p(v_2) \geq \mathbb{E}_q(v_2)$. This implies that

$$\begin{aligned}
\mathbb{E}_p(\lambda v_1 + (1-\lambda) v_2) &= \lambda \mathbb{E}_p(v_1) + (1-\lambda) \mathbb{E}_p(v_2) \\
&\geq \lambda \mathbb{E}_q(v_1) + (1-\lambda) \mathbb{E}_q(v_2) = \mathbb{E}_q(\lambda v_1 + (1-\lambda) v_2),
\end{aligned}$$

proving that $\mathcal{W}_{\max-\text{nor}}$ (resp., $\mathcal{W}_{\max}$) is convex. Next, consider $\{v_n\}_{n\in\mathbb{N}} \subseteq \mathcal{W}_{\max}$ such that $v_n \to v$. It is immediate to see that v is continuous. Moreover, if $p \succsim^* q$ then $\mathbb{E}_p(v_n) \geq \mathbb{E}_q(v_n)$ for all $n \in \mathbb{N}$. By passing to the limit, we obtain that $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$, that is, that $v \in \mathcal{W}_{\max}$, hence $\mathcal{W}_{\max}$ is closed.

2. By Dubra et al. (2004, Proposition 3), it follows that there exists $\hat{v} \in C([w,b])$ such that

$$
\begin{aligned}
& p \sim^* q \Longrightarrow \mathbb{E}_p(\hat{v}) = \mathbb{E}_q(\hat{v}) \\
& \quad \textit{and} \\
& p \succ^* q \Longrightarrow \mathbb{E}_p(\hat{v}) > \mathbb{E}_q(\hat{v}).
\end{aligned}
$$

By assumption, we have that $x \geq y$ if and only if $\delta_x \succsim^* \delta_y$. This implies that $x \geq y$ if and only if $\hat{v}(x) \geq \hat{v}(y)$, proving that $\hat{v}$ is strictly increasing. Since $\hat{v}$ is strictly increasing, by taking a positive and affine transformation, $\hat{v}$ can be chosen to be such that $\hat{v}(w) = 0 = 1 - \hat{v}(b)$. It is immediate to see that $\hat{v} \in \mathcal{W}_{\max-\text{nor}}$.

3. By definition of $\mathcal{W}_{\max}$, we have that if $p \succsim^* q$ then $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for all $v \in \mathcal{W}_{\max}$. On the other hand, by assumption and since $\mathcal{W} \subseteq \mathcal{W}_{\max}$, we have that if $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for all $v \in \mathcal{W}_{\max}$ then $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for all $v \in \mathcal{W}$ which, in turn, implies that $p \succsim^* q$. In other words, $\mathcal{W}_{\max}$ satisfies (2) for $\succsim^*$. By assumption, we can thus conclude that

$$
x \geq y \implies \delta_x \succsim^* \delta_y \implies \mathbb{E}_{\delta_x}(v) \geq \mathbb{E}_{\delta_y}(v) \ \ \forall v \in \mathcal{W}_{\max} \Longrightarrow v(x) \geq v(y) \ \ \forall v \in \mathcal{W}_{\max},
$$

proving that $\mathcal{W}_{\max} \subseteq \mathcal{V}_{in}$. By point 2 and since $\mathcal{W}_{\max-\text{nor}} \subseteq \mathcal{W}_{\max}$, we have that $\emptyset \neq \mathcal{W}_{\max} \cap \mathcal{V}_{s-in}$. Since $\mathcal{W}_{\max} \cap \mathcal{V}_{s-in} \subseteq \mathcal{W}_{\max}$ and the latter is closed, we have that $cl(\mathcal{W}_{\max} \cap \mathcal{V}_{s-in}) \subseteq \mathcal{W}_{\max}$. On the other hand, consider $\dot{v} \in \mathcal{W}_{\max} \cap \mathcal{V}_{s-in}$ and $v \in \mathcal{W}_{\max}$. Define $\{v_n\}_{n\in\mathbb{N}}$ by $v_n = \frac{1}{n}\dot{v} + \left(1 - \frac{1}{n}\right)v$ for all $n \in \mathbb{N}$. Since $v, \dot{v} \in \mathcal{W}_{\max}$ and the latter set is convex, we have that $\{v_n\}_{n\in\mathbb{N}} \subseteq \mathcal{W}_{\max}$. Since $\dot{v}$ is strictly increasing and v is increasing, v_n is strictly increasing for all $n \in \mathbb{N}$, proving that $\{v_n\}_{n\in\mathbb{N}} \subseteq \mathcal{W}_{\max} \cap \mathcal{V}_{s-in}$. Since $v_n \to v$, it follows that $v \in cl(\mathcal{W}_{\max} \cap \mathcal{V}_{s-in})$, proving that $\mathcal{W}_{\max} \subseteq cl(\mathcal{W}_{\max} \cap \mathcal{V}_{s-in})$ and thus the opposite inclusion.

4. By assumption, we have that there exists a subset $\mathcal{W}$ of $C([w,b])$ such that $p \succsim^* q$ if and only if $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for all $v \in \mathcal{W}$. By point 3 and its proof, we can replace first $\mathcal{W}$ with $\mathcal{W}_{\max}$ and then with $\mathcal{W}_{\max} \cap \mathcal{V}_{s-in}$. Consider $v \in \mathcal{W}_{\max} \cap \mathcal{V}_{s-in}$.

Since v is strictly increasing, there exist (unique) $\gamma_1 > 0$ and $\gamma_2 \in \mathbb{R}$ such that $\bar{v} = \gamma_1 v + \gamma_2$ is continuous, strictly increasing, and satisfies $\bar{v}(w) = 0 = 1 - \bar{v}(b)$. For each $v \in \mathcal{W}_{\max} \cap \mathcal{V}_{s-in}$, it is immediate to see that $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ if and only if $\mathbb{E}_p(\bar{v}) \geq \mathbb{E}_q(\bar{v})$. Define $\bar{\mathcal{W}} = \{\bar{v} : v \in \mathcal{W}_{\max} \cap \mathcal{V}_{s-in}\}$. Notice that $\bar{\mathcal{W}} \subseteq \mathcal{U}_{\text{nor}}$. From the previous part, we can conclude that $p \succsim^* q$ if and only if $\mathbb{E}_p(\bar{v}) \geq \mathbb{E}_q(\bar{v})$ for all $\bar{v} \in \bar{\mathcal{W}}$. It is also immediate to see that $\bar{\mathcal{W}} \subseteq \mathcal{W}_{\max - \text{nor}}$. By construction of $\mathcal{W}_{\max - \text{nor}}$, notice that

$$p \succsim^* q \Longrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{W}_{\max - \text{nor}}.$$

On the other hand, since $\bar{\mathcal{W}} \subseteq \mathcal{W}_{\max - \text{nor}}$, we have that

$$\mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{W}_{\max - \text{nor}} \Longrightarrow \mathbb{E}_p(\bar{v}) \geq \mathbb{E}_q(\bar{v}) \qquad \forall \bar{v} \in \bar{\mathcal{W}} \Longrightarrow p \succsim^* q.$$

We can conclude that $\mathcal{W}_{\max - \text{nor}}$ represents $\succsim^*$.

5. Consider $v \in \mathcal{W}$. By assumption, v is a strictly increasing and continuous function on $[w, b]$ such that $v(w) = 0 = 1 - v(b)$. Moreover, since $\mathcal{W}$ satisfies (2), it follows that $p \succsim^* q$ implies that $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$. This implies that $v \in \mathcal{W}_{\max - \text{nor}}$. We can conclude that $\mathcal{W} \subseteq \mathcal{W}_{\max - \text{nor}}$, hence, $cl(\mathcal{W}) \subseteq cl(\mathcal{W}_{\max - \text{nor}})$. In order to prove the opposite inclusion, we argue by contradiction. Assume that there exists $v \in cl(\mathcal{W}_{\max - \text{nor}}) \setminus cl(\mathcal{W})$. Since $v \in cl(\mathcal{W}_{\max - \text{nor}})$, we have that $v(w) = 0 = 1 - v(b)$. By Dubra et al. (2004, p. 123–124) and since both $\mathcal{W}$ and $\mathcal{W}_{\max - \text{nor}}$ satisfy (2), we also have that

$$cl\left(cone(\mathcal{W}) + \left\{\theta 1_{[w,b]}\right\}_{\theta \in \mathbb{R}}\right) = cl\left(cone(\mathcal{W}_{\max - \text{nor}}) + \left\{\theta 1_{[w,b]}\right\}_{\theta \in \mathbb{R}}\right).$$

We can conclude that $v \in cl\left(cone(\mathcal{W}) + \left\{\theta 1_{[w,b]}\right\}_{\theta \in \mathbb{R}}\right)$. Observe that there exists $\{\hat{v}_n\}_{n \in \mathbb{N}} \subseteq cone(\mathcal{W}) + \left\{\theta 1_{[w,b]}\right\}_{\theta \in \mathbb{R}}$ such that $\hat{v}_n \to v$. By construction and since $\mathcal{W}$ is convex, there exist $\{\lambda_n\}_{n \in \mathbb{N}} \subseteq [0, \infty)$, $\{v_n\}_{n \in \mathbb{N}} \subseteq \mathcal{W}$, and $\{\theta_n\}_{n \in \mathbb{N}} \subseteq \mathbb{R}$ such that $\hat{v}_n = \lambda_n v_n + \theta_n 1_{[w,b]}$ for all $n \in \mathbb{N}$. It follows that

$$\begin{aligned}
&0 = v(w) = \lim_n \hat{v}_n(w) = \lim_n \left\{\lambda_n v_n(w) + \theta_n 1_{[w,b]}(w)\right\} = \lim_n \theta_n \\
&\quad \textit{and} \\
&1 = v(b) = \lim_n \hat{v}_n(b) = \lim_n \left\{\lambda_n v_n(b) + \theta_n 1_{[w,b]}(b)\right\} = \lim_n \{\lambda_n + \theta_n\}.
\end{aligned}$$

This implies that $\lim_n \theta_n = 0 = 1 - \lim_n \lambda_n$. Without loss of generality, we can thus assume that $\{\lambda_n\}_{n\in\mathbb{N}}$ is bounded away from zero, that is, that there exists $\varepsilon > 0$ such that $\lambda_n \geq \varepsilon > 0$ for all $n \in \mathbb{N}$. Since $\{\theta_n\}_{n\in\mathbb{N}}$ and $\{\hat{v}_n\}_{n\in\mathbb{N}}$ are both convergent, both sequences are bounded, that is, there exists $k > 0$ such that

$$\|\hat{v}_n\| \leq k \text{ and } |\theta_n| \leq k \qquad \forall n \in \mathbb{N}.$$

It follows that

$$\begin{aligned}
\varepsilon \|v_n\| \leq \lambda_n \|v_n\| = \|\lambda_n v_n\| &= \left\|\lambda_n v_n + \theta_n 1_{[w,b]} - \theta_n 1_{[w,b]}\right\| \\
&\leq \left\|\lambda_n v_n + \theta_n 1_{[w,b]}\right\| + \left\|-\theta_n 1_{[w,b]}\right\| \\
&\leq \left\|\lambda_n v_n + \theta_n 1_{[w,b]}\right\| + |\theta_n| \\
&\leq \|\hat{v}_n\| + |\theta_n| \leq 2k \qquad \forall n \in \mathbb{N},
\end{aligned}$$

that is, $\|v_n\| \leq \frac{2k}{\varepsilon}$ for all $n \in \mathbb{N}$. We can conclude that

$$\begin{aligned}
\|v - v_n\| = \|v - \hat{v}_n + \hat{v}_n - v_n\| &\leq \|v - \hat{v}_n\| + \|\hat{v}_n - v_n\| \\
&= \|v - \hat{v}_n\| + \left\|\lambda_n v_n + \theta_n 1_{[w,b]} - v_n\right\| \\
&\leq \|v - \hat{v}_n\| + |\lambda_n - 1| \|v_n\| + |\theta_n| \\
&\leq \|v - \hat{v}_n\| + |\lambda_n - 1| \frac{2k}{\varepsilon} + |\theta_n| \qquad \forall n \in \mathbb{N}.
\end{aligned}$$

Passing to the limit, it follows that $v_n \to v$, that is, $v \in cl(\mathcal{W})$, a contradiction. ■

We next provide a characterization of $\succcurlyeq'$ which is due to Cerreia-Vioglio (2009). Here, it is further specialized to the particular case where $\succcurlyeq$ satisfies Weak Monotonicity and NCI in addition to Weak Order and Continuity. Before proving the statement, we need to introduce a piece of terminology. We will say that $\succcurlyeq''$ is an integral stochastic order if and only if there exists a set $\mathcal{W}'' \subseteq C([w,b])$ such that

$$p \succcurlyeq'' q \Longleftrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{W}''.$$

Proposition 5. *Let $\succcurlyeq$ be a binary relation on Δ that satisfies Weak Order, Continuity, Weak Monotonicity, and Negative Certainty Independence. The following statements are true:*

(a) *There exists a set* $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$ *such that* $p \succsim' q$ *if and only if* $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ *for all* $v \in \mathcal{W}$.

(b) *For each* $p, q \in \Delta$ *if* $p \succsim' q$ *then* $p \succsim q$.

(c) *If* $\succsim''$ *is an integral stochastic order that satisfies (b) then* $p \succsim'' q$ *implies* $p \succsim' q$.

(d) *If* $\succsim''$ *is an integral stochastic order that satisfies (b) and such that* $\mathcal{W}''$ *can be chosen to be a subset of* $\mathcal{U}_{\text{nor}}$ *then* $\overline{\text{co}}(\mathcal{W}) \subseteq \overline{\text{co}}(\mathcal{W}'')$.

Proof. (a). By Cerreia-Vioglio (2009, Proposition 22), there exists a set $\mathcal{W} \subseteq C([w,b])$ such that $p \succsim' q$ if and only if $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for all $v \in \mathcal{W}$. By Lemma 1, we also have that $x \geq y$ if and only if $\delta_x \succsim' \delta_y$. By point 4 of Proposition 4, if $\succsim^* = \succsim'$ then $\mathcal{W}$ can be chosen to be $\mathcal{W}_{\max-\text{nor}}$.

(b), (c), and (d). The statements follow from Cerreia-Vioglio (2009, Proposition 22 and Lemma 35). ■

The next proposition clarifies what is the relation between our assumption of Weak Monotonicity and First Order Stochastic Dominance. Given $p, q \in \Delta$, we write $p \succsim_{FSD} q$ if and only if p dominates q with respect to First Order Stochastic Dominance.

Proposition 6. *If* $\succsim$ *is a binary relation on* Δ *that satisfies Weak Order, Continuity, Weak Monotonicity, and Negative Certainty Independence then*

$$p \succsim_{FSD} q \Longrightarrow p \succsim q.$$

Proof. By Proposition 5, there exists $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$ such that

$$p \succsim' q \Longleftrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{W}.$$

By Proposition 5 and since $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}} \subseteq \mathcal{V}_{in}$, it follows that

$$\begin{aligned} p \succsim_{FSD} q &\Longrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{V}_{in} \\ &\Longrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{W} \\ &\Longrightarrow p \succsim' q \Longrightarrow p \succsim q, \end{aligned}$$

proving the statement. ■

Appendix B: Proof of the Results in the Text

Proof of Theorem 1. Before starting, we point out that in proving (i) implies (ii), we will prove the existence of a Continuous Cautious Expected Utility representation $\mathcal{W}$ which is convex and normalized, that is, a subset of $\mathcal{U}_{\text{nor}}$. This will turn out to be useful in the proofs of other results in this section. The normalization of $\mathcal{W}$ will play no role in proving (ii) implies (i).

(i) implies (ii). We proceed by steps.

Step 1. There exists a continuous certainty equivalent utility function $V : \Delta \to \mathbb{R}$.

Proof of the Step. Since $\succcurlyeq$ satisfies Weak Order and Continuity, there exists a continuous function $\bar{V} : \Delta \to \mathbb{R}$ such that $\bar{V}(p) \geq \bar{V}(q)$ if and only if $p \succcurlyeq q$. By Weak Monotonicity, we have that

$$x \geq y \iff \delta_x \succcurlyeq \delta_y \iff \bar{V}(\delta_x) \geq \bar{V}(\delta_y). \tag{3}$$

Next, observe that $\delta_b \succcurlyeq_{FSD} q \succcurlyeq_{FSD} \delta_w$ for all $q \in \Delta$. By Proposition 6 and since $\succcurlyeq$ satisfies Weak Order, Continuity, Weak Monotonicity, and NCI, this implies that

$$\delta_b \succcurlyeq q \succcurlyeq \delta_w \qquad \forall q \in \Delta. \tag{4}$$

Consider a generic $q \in \Delta$ and the sets

$$\{\delta_x : \delta_x \succcurlyeq q\} = \{p \in \Delta : p \succcurlyeq q\} \cap \{\delta_x\}_{x \in [w,b]}$$

and

$$\{\delta_x : q \succcurlyeq \delta_x\} = \{p \in \Delta : q \succcurlyeq p\} \cap \{\delta_x\}_{x \in [w,b]}.$$

By (4), Continuity, and Aliprantis and Border (2005, Theorem 15.8), both sets are nonempty and closed. Since $\succcurlyeq$ satisfies Weak Order, it follows that the sets

$$\{x \in [w,b] : \delta_x \succcurlyeq q\} \text{ and } \{x \in [w,b] : q \succcurlyeq \delta_x\}$$

are nonempty, closed, and their union coincides with $[w,b]$. Since $[w,b]$ is connected, there exists an element x_q in their intersection. In other words, there exists $x_q \in [w,b]$ such that $\delta_{x_q} \sim q$. Since q was chosen to be generic and by (3) and (4), such element is unique and we further have that

$$\bar{V}(\delta_b) \geq \bar{V}(q) = \bar{V}\left(\delta_{x_q}\right) = \bar{V}(q) \geq \bar{V}(\delta_w) \qquad \forall q \in \Delta. \tag{5}$$

Next, define $f : [w,b] \to \mathbb{R}$ by $f(x) = \bar{V}(\delta_x)$ for all $x \in [w,b]$. By (3), Aliprantis and Border (2005, Theorem 15.8), and (5), f is strictly increasing, continuous, and such that $f([w,b]) = \bar{V}(\Delta)$. It follows that $V : \Delta \to \mathbb{R}$ defined by $u = f^{-1} \circ \bar{V}$ is a well defined continuous function such that $p \succcurlyeq q$ if and only if $V(p) \geq V(q)$ and $V(\delta_x) = x$ for all $x \in [w,b]$, proving the statement. □

Step 2. $\succcurlyeq'$ *is represented by a set* $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$*, that is,*

$$p \succcurlyeq' q \Longleftrightarrow c(p,v) \geq c(q,v) \qquad \forall v \in \mathcal{W}. \tag{6}$$

Proof of the Step. It follows by point (a) of Proposition 5. Recall that $\mathcal{W}$ can be chosen to be $\mathcal{W}_{\max-\text{nor}}$ for $\succcurlyeq'$. □

Step 3. For each $p \in \Delta$ *we have that* $\inf_{v\in\mathcal{W}} c(p,v) \in [w,b]$.

Proof of the Step. Fix $p \in \Delta$. By construction, we have that $b \geq c(p,v) \geq w$ for all $v \in \mathcal{W}$. It follows that $c = \inf_{v\in\mathcal{W}} c(p,v)$ is a real number in $[w,b]$.

Step 4. For each $p \in \Delta$ *we have that*

$$V(p) \leq \inf_{v\in\mathcal{W}} c(p,v).$$

Proof of the Step. Fix $p \in \Delta$. By Step 3, $c = \inf_{v\in\mathcal{W}} c(p,v)$ is a real number in $[w,b]$. Since $V(\Delta) = [w,b]$, if $c = b$ then we have that $V(p) \leq b = c$. Otherwise, pick d such that $b > d > c$. Since $d > c$, we have that there exists $\tilde{v} \in \mathcal{W}$ such that

$$c(p,\widetilde{v}) < d = c(\delta_d,\widetilde{v}).$$

By Step 2, it follows that $p \not\succcurlyeq' \delta_d$. By Lemma 1, this implies that $\delta_d \succ p$, that is, $V(p) < V(\delta_d) = d$. Since d was chosen to be generic and strictly greater than c, we have that $V(p) \leq c$, proving the statement. □

Step 5. For each $p \in \Delta$ *we have that*

$$V(p) \geq \inf_{v\in\mathcal{W}} c(p,v).$$

Proof of the Step. Fix $p \in \Delta$. By Step 3, $c = \inf_{v \in \mathcal{W}} c(p, v)$ is a real number in $[w, b]$. By construction, we have that

$$c(p, v) \geq c = c(\delta_c, v) \qquad \forall v \in \mathcal{W}.$$

By Step 2, it follows that $p \succcurlyeq' \delta_c$. By Proposition 5 point (b), this implies that $p \succcurlyeq \delta_c$, that is, $V(p) \geq V(\delta_c) = c$, proving the statement. □

The implication follows from Steps 1, 2, 4, and 5.

(ii) implies (i). Assume there exists a set $\mathcal{W} \subseteq \mathcal{U}$ such that $V : \Delta \to \mathbb{R}$, defined by

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

is a continuous utility function for $\succcurlyeq$. Since $\succcurlyeq$ is represented by a continuous utility function, it follows that it satisfies Weak Order and Continuity. By construction, it is also immediate to see that $V(\delta_x) = x$ for all $x \in [w, b]$. In light of this fact, Weak Monotonicity follows immediately. Finally, consider $p \in \Delta$ and $x \in [w, b]$. Assume that $p \succcurlyeq \delta_x$. It follows that for each $\lambda \in [0, 1]$ and for each $q \in \Delta$

$$\begin{aligned}
c(p, v) \geq V(p) \geq V(\delta_x) = x = c(\delta_x, v) \qquad \forall v \in \mathcal{W} \\
&\Longrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_{\delta_x}(v) \qquad \forall v \in \mathcal{W} \\
&\Longrightarrow \mathbb{E}_{\lambda p + (1-\lambda) q}(v) \geq \mathbb{E}_{\lambda \delta_x + (1-\lambda) q}(v) \qquad \forall v \in \mathcal{W} \\
&\Longrightarrow c(\lambda p + (1-\lambda) q, v) \geq c(\lambda \delta_x + (1-\lambda) q, v) \qquad \forall v \in \mathcal{W} \\
&\Longrightarrow V(\lambda p + (1-\lambda) q) \geq V(\lambda \delta_x + (1-\lambda) q) \\
&\Longrightarrow \lambda p + (1-\lambda) q \succcurlyeq \lambda \delta_x + (1-\lambda) q,
\end{aligned}$$

proving that $\succcurlyeq$ satisfies NCI. ■

Proof of Proposition 1. Consider $\mathcal{W}$ and $\mathcal{W}'$ in $\mathcal{U}_{\text{nor}}$ such that $\overline{\text{co}}(\mathcal{W}) = \overline{\text{co}}(\mathcal{W}')$. Notice first that if both $\mathcal{W}$ and $\mathcal{W}'$ are convex, the proposition follows trivially. To prove the proposition, it will therefore suffice to show that for each $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}}$ we have that

$$\inf_{v \in \mathcal{W}} c(p, v) = \inf_{v \in co(\mathcal{W})} c(p, v) \qquad \forall p \in \Delta.$$

Consider $p \in \Delta$. It is immediate to see that

$$\inf_{v \in \mathcal{W}} c(p, v) \geq \inf_{v \in co(\mathcal{W})} c(p, v).$$

Conversely, consider $\bar{v} \in co(\mathcal{W})$. It follows that there exist $\{v_i\}_{i=1}^n \subseteq \mathcal{W}$ and $\{\lambda_i\}_{i=1}^n \subseteq [0,1]$ such that $\sum_{i=1}^n \lambda_i = 1$ and $\sum_{i=1}^n \lambda_i v_i = \bar{v}$. Define $x \in [w, b]$ and $\{x_i\}_{i=1}^n \subseteq [w, b]$ by $x = c(p, \bar{v})$ and $x_i = c(p, v_i)$ for all $i \in \{1, ..., n\}$. By contradiction, assume that $x < \min_{i \in \{1,...,n\}} x_i$. Since $\{v_i\}_{i=1}^n \subseteq \mathcal{W} \subseteq \mathcal{V}_{s-in}$, we have that

$$\mathbb{E}_p(\bar{v}) = \mathbb{E}_p\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i \mathbb{E}_p(v_i) = \sum_{i=1}^n \lambda_i v_i(x_i) > \sum_{i=1}^n \lambda_i v_i(x) = \bar{v}(x),$$

that is, $x = c(p, \bar{v}) > x$, a contradiction. This implies that

$$c(p, \bar{v}) = x \geq \min_{i \in \{1,...,n\}} x_i = \min_{i \in \{1,...,n\}} c(p, v_i) \geq \inf_{v \in \mathcal{W}} c(p, v).$$

Since $\bar{v}$ was chosen to be generic in $co(\mathcal{W})$, we can conclude that

$$c(p, \bar{v}) \geq \inf_{v \in \mathcal{W}} c(p, v) \qquad \forall \bar{v} \in co(\mathcal{W}),$$

proving that $\inf_{v \in \mathcal{W}} c(p, v) \leq \inf_{v \in co(\mathcal{W})} c(p, v)$ and thus the statement. ■

Proof of Theorem 2. By the proof of Theorem 1 (Steps 1, 2, 4, and 5), we have that there exists a set $\widehat{\mathcal{W}} \subseteq \mathcal{U}_{\text{nor}}$ such that

$$p \succsim' q \Longleftrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \widehat{\mathcal{W}} \tag{7}$$

and such that $V : \Delta \to \mathbb{R}$, defined by

$$V(p) = \inf_{v \in \widehat{\mathcal{W}}} c(p, v) \qquad \forall p \in \Delta, \tag{8}$$

is a continuous utility function for $\succsim$. This proves points (i) and (iii). Next consider a subset $\mathcal{W}$ of $\mathcal{U}_{\text{nor}}$ such that the function $V : \Delta \to \mathbb{R}$ defined by $V(p) = \inf_{v \in \mathcal{W}} c(p, v)$ for all $p \in \Delta$ represents $\succsim$. Define $\succsim''$ by

$$p \succsim'' q \Longleftrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{W}.$$

It is immediate to see that if $p \succsim'' q$ then $p \succsim q$. By point (d) of Proposition 5, this implies that $\overline{\text{co}}\left(\widehat{\mathcal{W}}\right) \subseteq \overline{\text{co}}\left(\mathcal{W}\right)$, proving point (ii).

Finally, consider two sets $\widehat{\mathcal{W}}_1$ and $\widehat{\mathcal{W}}_2$ in $\mathcal{U}_{\text{nor}}$ that satisfy (7) and (8). By point 5 of Proposition 4, it follows that $\overline{\text{co}}(\widehat{\mathcal{W}}_1) = \overline{\text{co}}(\widehat{\mathcal{W}}_2)$. ■

Proof of Theorem 3. We just prove point (i) since point (ii) follows by an analogous argument. Given $p \in \Delta$, we denote by $e(p)$ its expected value. We say that $p \succsim_{MPS} q$ if and only if q is a mean preserving spread of p.[36] Recall that $\succsim$ is risk averse if and only if $p \succsim_{MPS} q$ implies $p \succsim q$. Assume that $\succsim$ is risk averse. Let $p, q \in \Delta_0$.[37] Since Δ_0 is dense in Δ and $\succsim$ satisfies Weak Order and Continuity, we have that

$$\begin{aligned} p \succsim_{MPS} q &\Longrightarrow \lambda p + (1-\lambda) r \succsim_{MPS} \lambda q + (1-\lambda) r \qquad \forall \lambda \in (0,1], \forall r \in \Delta_0 \\ &\Longrightarrow \lambda p + (1-\lambda) r \succsim \lambda q + (1-\lambda) r \qquad \forall \lambda \in (0,1], \forall r \in \Delta_0 \\ &\Longrightarrow \lambda p + (1-\lambda) r \succsim \lambda q + (1-\lambda) r \qquad \forall \lambda \in (0,1], \forall r \in \Delta \\ &\Longrightarrow p \succsim' q. \end{aligned}$$

This implies that

$$p \succsim_{MPS} q \Longrightarrow p \succsim' q \Longrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \widehat{\mathcal{W}}.$$

We can conclude that each v in $\widehat{\mathcal{W}}$ is concave. For the other direction, assume that each v in $\widehat{\mathcal{W}}$ is concave. Since $\widehat{\mathcal{W}} \subseteq \mathcal{V}_{inco}$, we have that

$$\begin{aligned} p \succsim_{MPS} q &\Longrightarrow e(p) = e(q) \text{ and } \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \mathcal{V}_{inco} \\ &\Longrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \widehat{\mathcal{W}} \Longrightarrow p \succsim' q \Longrightarrow p \succsim q, \end{aligned}$$

proving that $\succsim$ is risk averse. ■

[36] Recall that, by Rothschild and Stiglitz (1970), if p and q are elements of Δ_0 and q is a mean preserving spread of p, then p and q have the same mean and they give the same probability to each point in their support with the exception of four ordered points $x_1 < x_2 < x_3 < x_4$. There the following relations hold:

$$\begin{gathered} q(x_1) - p(x_1) = p(x_2) - q(x_2) \geq 0 \\ \textit{and} \\ q(x_4) - p(x_4) = p(x_3) - q(x_3) \geq 0. \end{gathered}$$

[37] Recall that Δ_0 is the subset of Δ which contains just the elements with finite support.

Proof of Theorem 4. Before proceeding, we make a few remarks. Fix $i \in \{1,2\}$. By the proof of Theorem 1 and since $\succsim_i$ satisfies Weak Order, Continuity, Weak Monotonicity, and NCI, it follows that $\mathcal{W}^i_{\max-\text{nor}}$ for $\succsim'_i$ constitutes a Continuous Cautious Expected Utility representation of $\succsim_i$. Since $\mathcal{W}^i_{\max-\text{nor}}$ is convex, if $\widehat{\mathcal{W}}_i$ is chosen as in Theorem 2 then we have that $\overline{\text{co}}(\widehat{\mathcal{W}}_i)$ coincides with the closure of $\mathcal{W}^i_{\max-\text{nor}}$. Also recall that for each $p \in \Delta$, we denote by x^i_p the element in $[w,b]$ such that $p \sim_i \delta_{x^i_p}$. We also have that $V_i : \Delta \to \mathbb{R}$, defined by

$$V_i(p) = \inf_{v \in \mathcal{W}^i_{\max-\text{nor}}} c(p,v) = \inf_{v \in \widehat{\mathcal{W}}_i} c(p,v) = \inf_{v \in \mathcal{W}_i} c(p,v) \qquad \forall p \in \Delta,$$

represents $\succsim_i$, yielding that $x^i_p = V_i(p)$.

(i) implies (ii) and (i) implies (iii). Since $\succsim_1$ is more risk averse than $\succsim_2$, we have that $p \sim_1 \delta_{x^1_p}$ implies $p \succsim_2 \delta_{x^1_p}$. Since $\succsim_2$ satisfies Weak Order and Weak Monotonicity, it follows that $x^2_p \geq x^1_p$ for all $p \in \Delta$. This implies that

$$V_1(p) = \min\{V_1(p), V_2(p)\} = \inf_{v \in \mathcal{W}_1 \cup \mathcal{W}_2} c(p,v) \qquad \forall p \in \Delta,$$

that is, $\mathcal{W}_1 \cup \mathcal{W}_2$ is a Continuous Cautious Expected Utility representation of $\succsim_1$. By the remark in Section 3, it follows that $\widehat{\mathcal{W}_1 \cup \mathcal{W}_2}$ is also a Continuous Cautious Expected Utility representation of $\succsim_1$. By the initial part of the proof, we can conclude that $\overline{\text{co}}(\widehat{\mathcal{W}}_1) = cl\left(\mathcal{W}^1_{\max-\text{nor}}\right) = \overline{\text{co}}(\widehat{\mathcal{W}_1 \cup \mathcal{W}_2})$.

(iii) implies (i). Since $\overline{\text{co}}(\widehat{\mathcal{W}}_1) = cl\left(\mathcal{W}^1_{\max-\text{nor}}\right) = \overline{\text{co}}(\widehat{\mathcal{W}_1 \cup \mathcal{W}_2})$, it follows that

$$\begin{aligned} V_1(p) &= \inf_{v \in \widehat{\mathcal{W}}_1} c(p,v) = \inf_{v \in \widehat{\mathcal{W}_1 \cup \mathcal{W}_2}} c(p,v) \\ &= \inf_{v \in \mathcal{W}_1 \cup \mathcal{W}_2} c(p,v) = \min\{V_1(p), V_2(p)\} \ \leq \ V_2(p) \quad \forall p \in \Delta, \end{aligned}$$

proving that $x^2_p \geq x^1_p$ for all $p \in \Delta$. It follows that $\succsim_1$ is more risk averse than $\succsim_2$.

(ii) implies (i). Since $\mathcal{W}_1 \cup \mathcal{W}_2$ is a Continuous Cautious Expected Utility representation of $\succsim_1$, it follows that

$$V_1(p) = \inf_{v \in \mathcal{W}_1 \cup \mathcal{W}_2} c(p,v) \leq \inf_{v \in \mathcal{W}_2} c(p,v) \leq V_2(p) \qquad \forall p \in \Delta,$$

proving that $x^2_p \geq x^1_p$ for all $p \in \Delta$. It follows that $\succsim_1$ is more risk averse than $\succsim_2$. ■

Proof of Proposition 2. (i) We first prove necessity. By Theorem 2 and since $\succsim_1$ and $\succsim_2$ satisfy Weak Order, Continuity, Weak Monotonicity, and NCI, we have that

$$p \succsim_i' q \Longleftrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \widehat{\mathcal{W}}_i \Longleftrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v) \qquad \forall v \in \overline{\text{co}}(\widehat{\mathcal{W}}_i). \quad (9)$$

By Proposition 5 point (b) and since $\succsim_1$ is more indecisive than $\succsim_2$, we have that

$$p \succsim_1' q \Longrightarrow p \succsim_2' q \Longrightarrow p \succsim_2 q.$$

By Proposition 5 point (d) and (9), we can conclude that $\overline{\text{co}}(\widehat{\mathcal{W}}_2) \subseteq \overline{\text{co}}(\widehat{\mathcal{W}}_1)$. We next prove sufficiency. By (9) and since $\overline{\text{co}}(\widehat{\mathcal{W}}_2) \subseteq \overline{\text{co}}(\widehat{\mathcal{W}}_1)$, we have that

$$p \succsim_1' q \Longrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v)\ \forall v \in \overline{\text{co}}(\widehat{\mathcal{W}}_1) \Longrightarrow \mathbb{E}_p(v) \geq \mathbb{E}_q(v)\ \forall v \in \overline{\text{co}}(\widehat{\mathcal{W}}_2) \Longrightarrow p \succsim_2' q,$$

proving point (i).

(ii) By (9) and Proposition 4, we have that $\overline{\text{co}}(\widehat{\mathcal{W}}_i) = cl\left(\mathcal{W}^i_{\max-\text{nor}}\right)$ for $i \in \{1,2\}$. Since $\succsim_1$ is more indecisive than $\succsim_2$, it follows that $cl\left(\mathcal{W}^2_{\max-\text{nor}}\right) \subseteq cl\left(\mathcal{W}^1_{\max-\text{nor}}\right)$. By definition of $\mathcal{W}^1_{\max-\text{nor}}$ and $\mathcal{W}^2_{\max-\text{nor}}$, it follows that $\mathcal{W}^2_{\max-\text{nor}} \subseteq \mathcal{W}^1_{\max-\text{nor}}$. By the proof of Theorem 1, this implies that

$$V_1(p) = \inf_{v \in \mathcal{W}^1_{\max-\text{nor}}} c(p,v) \leq \inf_{v \in \mathcal{W}^2_{\max-\text{nor}}} c(p,v) = V_2(p) \qquad \forall p \in \Delta.$$

Since each V_i is a continuous certainty equivalent utility function, it follows that $\succsim_1$ is more risk averse than $\succsim_2$. ■

Proof of Theorem 5. Let $\succsim'$ be a reflexive and transitive binary relation on Δ that satisfies Sequential Continuity, Weak Monotonicity, and Independence. We first prove the existence of a Cautious Completion. In doing this, we show that this completion has a Cautious Expected Utility representation. By Dubra et al. (2004), there exists a set $\mathcal{W} \subseteq C([w,b])$ such that $p \succsim' q$ if and only if $\mathbb{E}_p(v) \geq \mathbb{E}_q(v)$ for all $v \in \mathcal{W}$. By Proposition 4, without loss of generality, we can assume that $\mathcal{W} \subseteq \mathcal{U}_{\text{nor}} \subseteq \mathcal{U}$.

Next define the binary relation $\hat{\succsim}$ as

$$p \mathrel{\hat{\succsim}} q \Longleftrightarrow \inf_{v \in \mathcal{W}} c(p,v) \geq \inf_{v \in \mathcal{W}} c(q,v). \quad (10)$$

Notice that $\hat{\succsim}$ is well defined and it satisfies Weak Order, Weak Monotonicity, and clearly for each $p \in \Delta$ there exists $x \in [w,b]$ such that $p \hat{\sim} \delta_x$. Next, we show $\hat{\succsim}$ is a completion of $\succsim'$. Since each $v \in \mathcal{W}$ is strictly increasing, we have that

$$p \succsim' q \Longleftrightarrow c(p,v) \geq c(q,v) \qquad \forall v \in \mathcal{W}. \tag{11}$$

This implies that if $p \succsim' q$ then $\inf_{v\in\mathcal{W}} c(p,v) \geq \inf_{v\in\mathcal{W}} c(q,v)$, that is, if $p \succsim' q$ then $p \hat{\succsim} q$. Finally, let x be an element of $[w,b]$ and $p \in \Delta$ such that $p \not\succsim' \delta_x$. By (11), it follows that there exists $\tilde{v} \in \mathcal{W}$ such that $c(\delta_x, \widetilde{v}) = x > c(p, \widetilde{v})$. By (10), this implies that $\inf_{v\in\mathcal{W}} c(\delta_x, v) = x > \inf_{v\in\mathcal{W}} c(p,v)$, hence $\delta_x \hat{\succ} p$. This concludes the proof of the existence of a Cautious Completion.

We are left with proving uniqueness. Let $\succsim^\circ$ be a Cautious Completion of $\succsim'$. By point 1 of Definition 5, $\succsim^\circ$ satisfies Weak Order, Weak Monotonicity, and for each $p \in \Delta$ there exists $x \in [w,b]$ such that $p \sim^\circ \delta_x$. This implies that there exists $V : \Delta \to \mathbb{R}$ such that V represents $\succsim^\circ$ and $V(\delta_x) = x$ for all $x \in [w,b]$. Moreover, we have that $V(\Delta) = [w,b]$. Let $p \in \Delta$. Define $c = \inf_{v\in\mathcal{W}} c(p,v) \in [w,b]$. If $c = b$ then $V(p) \leq b = c = \inf_{v\in\mathcal{W}} c(p,v)$. If $c < b$ then for each $d \in (c,b)$ there exists $\tilde{v} \in \mathcal{W}$ such that $c(\delta_d, \widetilde{v}) = d > c(p, \widetilde{v})$, yielding that $p \not\succsim' \delta_d$. By point 3 of Definition 5, we can conclude that $\delta_d \succ^\circ p$, that is, $d = V(\delta_d) > V(p)$. Since d was arbitrarily chosen in (c,b), it follows that $V(p) \leq c = \inf_{v\in\mathcal{W}} c(p,v)$. Finally, by definition of c and (11), we have that $c(p,v) \geq c = c(\delta_c, v)$ for all $v \in \mathcal{W}$, that is, $p \succsim' \delta_c$. By point 2 of Definition 5, it follows that $p \succsim^\circ \delta_c$, that is, $V(p) \geq V(\delta_c) = c = \inf_{v\in\mathcal{W}} c(p,v)$. In other words, we have shown that $V(p) = \inf_{v\in\mathcal{W}} c(p,v)$ for all $p \in \Delta$. By (10) and since V represents $\succsim^\circ$, we can conclude that $\succsim^\circ = \hat{\succsim}$, proving the statement. ■

References

Agranov, M. and P. Ortoleva (2014): "Stochastic Choice and Hedging," Mimeo California Institute of Technology.

Aliprantis, C. D. and K. C. Border (2005): *Infinite Dimensional Analysis*, Springer, Berlin Heidelberg New York.

Andreoni, J. and W. Harbaugh (2010): "Unexpected utility: Experimental tests of five key questions about preferences over risk," Mimeo University of California San Diego.

Andreoni, J. and C. Sprenger (2010): "Certain and Uncertain Utility: The Allais Paradox and Five Decision Theory Phenomena," Mimeo University of California San Diego.

——— (2012): "Uncertainty Equivalents: Linear Tests of the Independence Axiom," Mimeo, University of California San Diego.

Artstein-Avidan, S. and D. Dillenberger (2011): "Dynamic Disappointment Aversion," Mimeo, University of Pennsylvania.

Battalio, R. C., J. H. Kagel, and K. Jiranyakul (1990): "Testing between alternative models of choice under uncertainty: Some initial results," *Journal of risk and uncertainty*, 3, 25–50.

Bell, D. E. and P. C. Fishburn (2003): "Probability weights in rank-dependent utility with binary even-chance independence," *Journal of Mathematical Psychology*, 47, 244–258.

Bernasconi, M. (1994): "Nonlinear preferences and two-stage lotteries: Theories and evidence," *The Economic Journal*, 54–70.

Blavatskyy, P. R. (2010): "Reverse common ratio effect," *Journal of risk and uncertainty*, 40, 219–241.

Burke, M. S., J. R. Carter, R. D. Gominiak, and D. F. Ohl (1996): "An experimental note on the allais paradox and monetary incentives," *Empirical Economics*, 21, 617–632.

Camerer, C. F. (1989): "An experimental test of several generalized utility theories," *Journal of Risk and uncertainty*, 2, 61–104.

——— (1992): "Recent tests of generalizations of expected utility theory," *Utility theories: Measurements and applications*, 207–251.

——— (1995): "Individual decision making," in *Handbook of Experimental Economics*, ed. by A. R. J. Kagel, Princeton University Press, Princeton, vol. 2.

Camerer, C. F. and T.-H. Ho (1994): "Violations of the betweenness axiom and nonlinearity in probability," *Journal of risk and uncertainty*, 8, 167–196.

Cerreia-Vioglio, S. (2009): "Maxmin Expected Utility on a Subjective State Space: Convex Preferences under Risk," Mimeo, Bocconi University.

Cerreia-Vioglio, S., P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi (2011): "Rational Preferences under Ambiguity," *Economic Theory*, 48, 341–375.

Chatterjee, K. and V. R. Krishna (2011): "A nonsmooth approach to nonexpected utility theory under risk," *Mathematical Social Sciences*, 62, 166–175.

Chew, S. H. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, 51, 1065–1092.

——— (1989): "Axiomatic utility theories with the betweenness property," *Annals of Operations Research*, 19, 273–298.

Chew, S. H. and L. G. Epstein (1989): "A unifying approach to axiomatic non-expected utility theories," *Journal of Economic Theory*, 49, 207–240.

Chew, S. H. and W. S. Waller (1986): "Empirical tests of weighted utility theory," *Journal of Mathematical Psychology*, 30, 55–72.

Cohen, M. and J. Jaffray (1988): "Certainty effect versus probability distortion: An experimental analysis of decision making under risk." *Journal of Experimental Psychology: Human Perception and Performance*, 14, 554.

Conlisk, J. (1989): "Three variants on the Allais example," *The American Economic Review*, 79, 392–407.

Dean, M. and P. Ortoleva (2014): "Allais, Ellsberg, and Preferences for Hedging," Mimeo, Columbia University.

Dekel, E. (1986): "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom," *Journal of Economic Theory*, 40, 304–318.

Dillenberger, D. (2010): "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior," *Econometrica*, 78, 1973–2004.

Dillenberger, D. and S. Erol (2013): "Negative certainty independence without Betweenness," *Economics Letters (forthcoming).*

Dubra, J., F. Maccheroni, and E. A. Ok (2004): "Expected utility theory without the completeness axiom," *Journal of Economic Theory*, 115, 118–133.

Dwenger, N., D. Kübler, and G. Weizsäcker (2013): "Flipping a Coin: Theory and Evidence," Mimeo Humboldt-Universität zu Berlin.

Evren, Ö. (2014): "Scalarization methods and expected multi-utility representations," *Journal of Economic Theory*, 151, 30–63.

Fan, C.-P. (2002): "Allais paradox in the small," *Journal of Economic Behavior & Organization*, 49, 411–421.

Galaabaatar, T. and E. Karni (2013): "Subjective Expected Utility with Incomplete Preferences," *Econometrica*, 81, 255–284.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004): "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory*, 118, 133–173.

Gilboa, I. (1988): "A combination of expected utility and maxmin decision criteria," *Journal of Mathematical Psychology*, 32, 405–420.

Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Gilboa, I., D. Schmeidler, F. Maccheroni, and M. Marinacci (2010): "Objective and subjective rationality in a multi-prior model," *Econometrica*, 78, 755–770.

Gul, F. (1991): "A theory of disappointment aversion," *Econometrica*, 59, 667–686.

Harless, D. W. and C. F. Camerer (1994): "The predictive utility of generalized expected utility theories," *Econometrica*, 1251–1289.

Holt, C. and S. Laury (2002): "Risk aversion and incentive effects," *American Economic Review*, 92, 1644–1655.

Huck, S. and W. Müller (2012): "Allais for all: Revisiting the paradox in a large representative sample," *Journal of Risk and Uncertainty*, 44, 261–293.

Humphrey, S. J. (2000): "The common consequence effect: Testing a unified explanation of recent mixed evidence," *Journal of Economic Behavior & Organization*, 41, 239–262.

Ikefuji, M., R. J. Laeven, J. R. Magnus, and C. Muris (2012): "Pareto utility," *Theory and Decision*, 1–15.

Jaffray, J. Y. (1988): "Choice under risk and the security factor: an axiomatic model," *Theory and Decision*, 24, 169–200.

Kahneman, D. and A. Tversky (1979): "Prospect theory: an analysis of choice under risk," *Econometrica*, 47, 263–291.

Maccheroni, F. (2002): "Maxmin under risk," *Economic Theory*, 19, 823–831.

MacCrimmon, K. R. and S. Larsson (1979): "Utility theory: Axioms versus paradoxes," in *Expected utility hypotheses and the Allais paradox*, Springer, 333–409.

Machina, M. J. (1982): "Expected Utility" Analysis without the Independence Axiom," *Econometrica*, 50, 277–323.

——— (1987): "Choice Under Uncertainty: Problems Solved and Unsolved," *The Journal ofEconomic Perspectives*, 50, 121–154.

Ok, E., P. Ortoleva, and G. Riella (2012): "Incomplete Preferences under Uncertainty: Indecisiveness in Beliefs vs. Tastes," *Econometrica*, 80, 1791–1808.

Prelec, D. (1990): "A Pseudo-endowment effect, and its implications for some recent nonexpected utility models," *Journal of Risk and Uncertainty*, 3, 247–259.

Quiggin, J. (1982): "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 3, 323–343.

Riella, G. (2013): "On the Representation of Incomplete Preferences under Uncertainty with Indecisiveness in Tastes," Mimeo, Universidade de Brasília.

Röell, A. (1987): "Risk aversion in Quiggin and Yaari's rank-order model of choice under uncertainty," *The Economic Journal*, 143–159.

Rothschild, M. and J. E. Stiglitz (1970): "Increasing risk: I. A definition," *Journal of Economic theory*, 2, 225–243.

Saha, A. (1993): "Expo-power utility: A 'flexible' form for absolute and relative risk aversion," *American Journal of Agricultural Economics*, 75, 905–913.

Schmidt, U. (1998): "A Measurement of the Certainty Effect," *Journal of Mathematical Psychology*, 42, 32–47.

Sopher, B. and G. Gigliotti (1993): "A test of generalized expected utility theory," *Theory and Decision*, 35, 75–106.

Sopher, B. and J. M. Narramore (2000): "Stochastic choice and consistency in decision making under risk: An experimental study," *Theory and Decision*, 48, 323–350.

Starmer, C. (2000): "Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk," *Journal of Economic Literature*, 38, 332–382.

Tversky, A. and D. Kahneman (1992): "Advances in prospect theory: cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Wakker, P., I. Erev, and E. U. Weber (1994): "Comonotonic independence: The critical test between classical and rank-dependent utility theories," *Journal of Risk and Uncertainty*, 9, 195–230.

Wu, G. (1994): "An empirical test of ordinal independence," *Journal of Risk and Uncertainty*, 9, 39–60.

Wu, G. and R. Gonzalez (1998): "Common consequence conditions in decision making under risk," *Journal of Risk and Uncertainty*, 16, 115–139.

Wu, G., J. Zhang, and M. Abdellaoui (2005): "Testing prospect theories using probability tradeoff consistency," *Journal of Risk and Uncertainty*, 30, 107–131.