Centre for
Economic
and Financial
Research
at
New Economic
School

January 2017

Cautious and Globally Ambiguity Averse

Özgür Evren

Working Paper No 236

CEFIR / NES Working Paper series

Cautious and Globally Ambiguity Averse*

Özgür Evren†

January 23, 2017

Abstract

I study ambiguity attitudes in Uzi Segal's recursive non-expected utility model. I show that according to this model, the negative certainty independence axiom over simple lotteries is equivalent to a robust, or global form of ambiguity aversion that requires ambiguity averse behavior irrespective of the number of states and the decision maker's second-order belief. Thus, the recursive cautious expected utility model is the only subclass of Segal's model that robustly predicts ambiguity aversion. Similarly, the independence axiom over lotteries is equivalent to a robust form of ambiguity neutrality. In fact, any non-expected utility preference over lotteries coupled with a suitable second-order belief over three states produces either the Ellsberg paradox or the opposite mode of behavior. Finally, I propose a definition of a mean-preserving spread for second-order beliefs that is equivalent to increasing ambiguity aversion for every recursive preference that satisfies the negative certainty independence axiom.

JEL Classification: D81
Keywords: Ambiguity Aversion, Ellsberg Paradox, Allais Paradox, Negative Certainty Independence, Reduction of Compound Lotteries, Increasing Second-Order Uncertainty

---

*For their comments and suggestions, I thank A. Baillon, S. Cerreia-Vioglio, S.H. Chew, F. Dietrich, D. Dillenberger, D. Freeman, Y. Halevy, M. Marinacci, S. Minardi, E. Ok, P. Ortoleva, E. Ozdenoren, C. Raymond, A. Savochkin, P. Wakker, K. Yıldız, and seminar participants at Bilkent University, Erasmus University, New Economic School, ESEM 2016 and RUD 2016.

†New Economic School, 100 Novaya Street, Skolkovo, Moscow 143025, Russia. Email: oevren@nes.ru

1 Introduction

In the literature on decision making under uncertainty, "ambiguity" refers to situations in which the probabilities of some events are not so clear, whereas the term "risk" is reserved for those situations that involve objectively defined probabilities. Following Ellsberg's (1961) classic examples, which became known as the "Ellsberg paradox," a tendency to prefer risky bets (i.e., lotteries) to ambiguous bets has emerged as a common behavioral pattern in experimental studies. One of the earliest models of such ambiguity averse behavior is Segal's (1987) theory of recursive non-expected utility preferences.

Segal's model predicts certain connections between attitudes towards ambiguity and risk that make this model unique in the related literature. The starting point of the theory is that, absent objectively defined probabilities, the decision maker (DM) deems likely several probability distributions over the states of nature. More specifically, the DM holds a (subjective) second-order belief that attaches a probability to any given (first-order) probability distribution over the states. Then, with the help of this second-order belief, an ambiguous act—a function that assigns prizes to states—is evaluated as if it is a compound lottery, i.e., a lottery over lotteries. Thus, according to this theory, ambiguity aversion is closely related to aversion towards compound risk, as opposed to one-shot or simple risk. In particular, the DM can exhibit Ellsberg-type choices only if she violates the reduction of compound lotteries axiom. Moreover, the reduction of compound lotteries axiom fails only if the independence axiom fails over simple lotteries, implying that Ellsberg-type choices under ambiguity necessitate violations of the independence axiom under risk, as in Allais' (1953) classic examples.[1]

In a related experiment that studies the interplay between attitudes towards ambiguity, and compound and simple risk, Halevy (2007) finds that the behavior of 40 percent of his subjects comply with the predictions of Segal's theory. More recently, Dean and Ortoleva (2015) also report a significant correlation between ambiguity aversion and (i) aversion towards compound risk; (ii) Allais' common ratio and common consequence effects.[2]

Aside from a second-order belief, the only other primitive of Segal's theory is a preference relation over simple lotteries, which directly influences the DM's attitude towards ambiguity. While, in principle, one can incorporate any risk preference into the theory, Segal focuses on Quiggin's (1982) rank dependent utility (RDU) model. One of the main

[1] Seo's (2009) theory of second-order subjective expected utility also explains Ellsberg-type choices with the failure of the reduction of compound lotteries axiom. However, that model assumes expected utility preferences over simple lotteries, unlike Segal's theory. The two models also generate different attitudes towards the timing of the resolution of uncertainty (see Appendix A).

[2] On the other hand, Abdellaoui, Klibanoff and Placido (2015) find only a weak association between attitudes towards compound risk and ambiguity.

findings of Segal (1987, Theorem 4.2) shows that given a risk preference that admits an RDU representation with a convex probability distortion function, under some further assumptions, the theory does indeed generate Ellsberg-type choices over binary bets, which involve only a good prize and a bad one.

More recently, Dillenberger (2010) has shown that according to Segal's model, an axiom about simple lotteries, called "negative certainty independence" (NCI), is equivalent to another axiom about compound lotteries, called "preference for one-shot resolution of uncertainty" (PORU). NCI is a generalization of the independence axiom over simple lotteries that accommodates Allais' common ratio and common consequence effects. In turn, PORU asserts that the reduced form of a compound lottery should be preferred to that lottery. Although Dillenberger is not directly concerned with ambiguity, given the aforementioned relation between ambiguous acts and compound lotteries in Segal's model, the equivalence between NCI and PORU suggests that risk preferences that satisfy NCI may provide a good alternative to the RDU model for the purposes of modeling ambiguity aversion.

Specifically, in light of the role of second-order beliefs in Segal's theory, PORU can be interpreted as saying that the DM would deem uncertain acts more valuable if she were able to convert ambiguity into risk by replacing her second-order belief with the reduced form of that belief. Hence, if PORU holds, for any second-order belief on any state space, the reduced form of that belief will act as a "benchmark" that qualifies the DM as ambiguity averse. Consequently, a risk preference that satisfies PORU (or NCI) possesses a *global ambiguity aversion property* in the sense that it implies ambiguity aversion irrespective of the second-order beliefs or the number of states (Artstein-Avidan and Dillenberger, 2011).

On the other hand, modern definitions of ambiguity aversion do not impose a restriction on first-order distributions that can act as a benchmark (see Epstein, 1999; Ghirardato and Marinacci, 2002; and, Dean and Ortoleva, in press). What implications does this approach entail in Segal's model? Does there exist an ambiguity averse recursive preference with a benchmark that is distinct from the reduced form of the second-order belief in question? If so, is it possible to devise an alternative, non-standard method of reduction that systematically generates benchmarks from second-order beliefs for a given class of risk preferences that may as well violate NCI? After all, is NCI a necessary consequence of the global ambiguity aversion property?

Motivated by these questions, in the first part of the present paper, I provide an in-depth analysis of the relation between NCI and ambiguity aversion in Segal's theory. My first result (Theorem 1) shows that the global ambiguity aversion property does, indeed, imply NCI. This holds true despite the fact that the benchmark for an ambiguity averse preference can actually be distinct from the reduced form of the associated second-order belief. In Section 3.2, I provide an example of this sort where the DM exhibits an Ellsberg paradox with two urns. The known distribution in the risky urn acts as a benchmark and is distinct from the reduced form of the DM's belief about the ambiguous urn, whereas the Ellsberg-type behavior disappears if we change the composition of the risky urn so as to make it consistent with the DM's reduced form belief.

However, it turns out that if a second-order belief possesses a certain property, which I call "uniform separability," then the only possible benchmark for that belief is its reduced form.[3] While the class of uniformly separable beliefs is truly special in many ways, the proof of Theorem 1 shows that, at the same time, this class is rich enough to deduce NCI from the global ambiguity aversion property: If a risk preference induces ambiguity averse behavior for every uniformly separable belief, then it must also satisfy NCI. To summarize, uniformly separable beliefs provide a "sufficiently rich" class of examples where the mode of behavior demanded by NCI is a necessary consequence of ambiguity aversion. On the other hand, in general, unlike NCI or PORU, ambiguity aversion cannot be viewed as a statement on reduced form beliefs.

As noted by Ellsberg himself, global ambiguity aversion may be too demanding from a descriptive point of view (see Machina, Ritzberger, Yannelis and Ellsberg, 2011, Section 2). Indeed, some experimental evidence indicates that people may prefer not to know the exact probabilities when the likelihood of a gain appears to be small (e.g., Kocher, Lahno, and Trautmann, 2016). However, one can think of many applications in which the analyst may want to abstract from such cases. For example, the aforementioned evidence on unlikely gains may not be so relevant for an analysis of a stock market that is expected to perform well with a moderate or high likelihood. Or, to check if ambiguity aversion can lead to another phenomenon in a particular framework, the analyst may want to perform a clear-cut comparison of ambiguity aversion and neutrality/loving that does not depend on the details of second-order beliefs. The main message of Theorem 1 is that, in potential applications of Segal's theory, if we want a model that is guaranteed to generate ambiguity averse behavior, we must select a risk preference that satisfies NCI. Otherwise, the model may well produce non-ambiguity averse behavior (or, only a partial form of ambiguity aversion), depending on the exact structure of the risk preference and the second-order belief selected by the analyst, and how the two primitives of the model interact with each other.

Cerreia-Vioglio, Dillenberger and Ortoleva (2015) have recently provided a utility representation theorem for risk preferences that satisfy NCI, called a "cautious expected utility" (CEU) representation.[4] As a corollary of Theorem 1, it follows that the class of CEU preferences coincides with the class of risk preferences that have the global ambiguity aversion

[3] Roughly speaking, a uniformly separable second-order belief deems possible only a collection of uniform distributions supported over pairwise disjoint sets.

[4] Perhaps the best known non-expected utility model within this class is Gul's (1991) theory of disappointment aversion (see Artstein-Avidan and Dillenberger, 2011).

property. By contrast, as shown by Dillenberger (2010, Proposition 3), the RDU model violates NCI generically. Hence, Theorem 1 also implies that a generic RDU preference lacks the global ambiguity aversion property. To illustrate this point, in Section 3.3, I show that when there are four states, as opposed to the binary case in Segal (1987), the (symmetry, elasticity and convexity) conditions considered by Segal are no longer sufficient for a recursive RDU preference to be ambiguity averse.[5]

Upon combining Theorem 1 with a dual result, it also follows that the class of risk preferences that are globally ambiguity neutral coincides with the class of expected utility preferences. In other words, if a risk preference violates the independence axiom, it will necessarily generate a paradoxical form of ambiguity aversion or the opposite mode of behavior, at least in one instance, depending on the number of states and the specification of a second-order belief. Here, "paradoxical" means that the behavior is incompatible with any first-order belief. Theorem 2 makes precise what sort of a paradox we can expect to observe in practice. The answer turns out to be surprisingly simple: The risk preference coupled with a suitable second-order belief will generate either the Ellsberg paradox or the opposite in a two-urn experiment with three states and binary acts.

Finally, I investigate a dual question: Holding risk preferences constant, how can we manipulate a given second-order belief to obtain a more ambiguity averse recursive preference? Such a manipulation method will find more applications if it functions independently of the details of risk preferences, just as the classical mean-preserving spread operation over monetary lotteries, à la Rothschild and Stiglitz (1970). Assuming NCI as a minimal requirement on risk preferences, Theorem 3 shows that a mean-preserving spread operation over second-order beliefs characterizes such increase in ambiguity aversion. That is, a second-order belief μ is a mean-preserving spread of another second-order belief μ' if and only if for any risk preference that satisfies NCI, μ induces a more ambiguity averse preference than that induced by μ'. Aside from the differences in primitives, the notion of a mean-preserving spread characterized in this theorem is a stronger version of its classical counterpart for monetary lotteries.

Throughout the paper, I utilize a definition of ambiguity aversion that was recently proposed by Dean and Ortoleva (in press). In line with my purposes, this definition is tailored for non-expected utility preferences over lotteries.[6] Moreover, as I noted earlier, the

[5] As usual, this means that no first-order distribution can qualify the preference relation as ambiguity averse, while Dillenberger's (2010) Proposition 3 should be interpreted as a statement on a particular distribution, namely the reduced form of the DM's second-order belief. A recursive RDU preference may fail to pass a test of ambiguity aversion based on the reduced form beliefs—in line with Dillenberger's Proposition 3—and yet, exhibit Ellsberg-type choices. The aforementioned example in Section 3.2 describes such a preference relation. By contrast, the example in Section 3.3 depicts a paradoxical form of ambiguity loving which rules out Ellsberg-type choices.

[6] Specifically, according to Dean and Ortoleva, a preference relation is (absolutely) ambiguity averse if

definition does not impose a restriction on first-order beliefs that can act as a benchmark. This contrasts with Segal's (1987) analysis which focuses on reduced form beliefs.

1.1 Related Literature

Earlier work on Segal's (1987) theory includes Artstein-Avidan and Dillenberger (2011), and Dillenberger and Segal (2015a,b).

The focus of Artstein-Avidan and Dillenberger is mainly on risk, but they also note that NCI implies global ambiguity aversion. The contribution of my Theorem 1 is the converse implication, which allows me to single out the recursive CEU model as the only class within recursive preferences that robustly predicts ambiguity aversion.

The main point of Dillenberger and Segal (2015a) is that the recursive disappointment aversion model—a particular form of recursive CEU preferences—accommodates not only Ellsberg-type choices, but also some related phenomena about non-binary acts that were recently pointed out by Machina (2009, 2014).

Finally, Dillenberger and Segal (2015b) show that the theory of recursive preferences can also be used to model a non-global form of ambiguity aversion, in line with the aforementioned experimental evidence on violations of global ambiguity aversion. Specifically, they model a DM who dislikes ambiguity when the likelihood of a gain is moderate or large, while having the opposite attitude when a gain appears to be unlikely. Naturally, this version of the theory is distinct from the recursive CEU model.

In the next section, I formulate Segal's (1987) theory in the framework of Anscombe and Aumann (1963), which offers a compact way of modelling preferences over simple and compound lotteries as well as subjective acts.[7] Sections 3 and 4 are devoted to global ambiguity aversion and neutrality, respectively. In Section 5, I discuss the notion of a mean-preserving spread for second-order beliefs. The appendix contains all proofs and some further supplementary material.

2 Model

Throughout the paper, $\Delta(A)$ denotes the set of all probability measures on a set A with finite support. Given $l \in \Delta(A)$ and $B \subseteq A$, $l(B)$ represents the probability of B. I write

it is relatively more ambiguity averse than a probabilistically sophisticated preference (i.e., a benchmark), where relative ambiguity aversion is defined as in Ghirardato and Marinacci (2002). In Section 3, I relate this definition to the corresponding definitions of Epstein (1999) and Ghirardato and Marinacci (2002).

[7] While my findings can also be formulated in a more basic setup with purely subjective acts and simple lotteries, I have chosen the Anscombe-Aumann model to be able to relate my findings to earlier results on compound risk.

$l(a)$ in place of $l(\{a\})$. Since the support of l is a finite set, we have $\sum_{a\in A} l(a) = 1$ for every $l \in \Delta(A)$. For $\{l^1, ..., l^n\} \subseteq \Delta(A)$ and $\{\alpha^1, ..., \alpha^n\} \subseteq [0,1]$ with $\sum_{i=1}^n \alpha^i = 1$, the mixture $\sum_{i=1}^n \alpha^i l^i$ is the element of $\Delta(A)$ that attaches the probability $\sum_{i=1}^n \alpha^i l^i(a)$ to any $a \in A$. Similarly, $\Delta^2(A) := \Delta(\Delta(A))$ stands for the set of all probability measures on $\Delta(A)$ with finite support, and the mixture operation on this set is defined analogously.

$X := [x_*, x^*] \subseteq \mathbb{R}$ denotes a set of monetary prizes with $x_* < x^*$. The elements of $\Delta(X)$ represent (simple or one-shot) lotteries. In turn, $\Delta^2(X)$ is considered as the set of compound lotteries.

Let S be a nonempty, finite state space. An *act* refers to a function that maps S into $\Delta(X)$. $\mathcal{H}$ stands for the set of all acts. The primitive of the model is a binary relation $\succsim$ on $\Delta(\mathcal{H})$ that represents the preferences of a DM.

The following table summarizes my notation for some generic objects.

For generic elements of:	X	$\Delta(X)$	$\Delta^2(X)$	$\mathcal{H}$	$\Delta(\mathcal{H})$	S	$\Delta(S)$	$\Delta^2(S)$
I write:	x, y, z	p, q, r	P, Q	f, g	F, G	s	π	μ

The DM does not (necessarily) know the distribution of the states. Rather, she holds a second-order belief $\mu \in \Delta^2(S)$, where $\mu(\pi)$ represents the probability that the "correct" distribution of the states is given by $\pi \in \Delta(S)$.

D_p, D_π and D_f stand for the degenerate elements of $\Delta^2(X)$, $\Delta^2(S)$ and $\Delta(\mathcal{H})$, respectively. D_p attaches probability 1 to the lottery $p \in \Delta(X)$, and similarly for D_π and D_f. In turn, δ_x and δ_s denote the degenerate elements of $\Delta(X)$ and $\Delta(S)$ supported at x and s, respectively.

I identify f with D_f, and p with the constant act that returns p at every state, $\mathbf{1}_S p$. That is, for every $p \in \Delta(X)$ and $f \in \mathcal{H}$,

$$p \equiv \mathbf{1}_S p, \quad f \equiv D_f.$$

Hence, $\Delta(X) \subseteq \mathcal{H} \subseteq \Delta(\mathcal{H})$. By the former inclusion, we also have $\Delta^2(X) \subseteq \Delta(\mathcal{H})$.

I say that an act f is *purely subjective* if for every state s, $f(s)$ is equal to δ_x for some $x \in X$. $\mathcal{H}_X$ denotes the set of all such acts, which do not involve objective uncertainty.

2.1 Representation Notion

Definition 1. A *certainty equivalence function* c is a map from $\Delta(X)$ onto X, continuous in the topology of weak convergence, and such that:

(i) $p \geq_{\text{fosd}} (>_{\text{fosd}}) q \quad \Rightarrow \quad c(p) \geq (>) c(q)$.

(ii) $c(\delta_x) = x$ for every $x \in X$.

Here, $\geq_{\text{fosd}}$ stands for the first order stochastic dominance relation on $\Delta(X)$, with the asymmetric part $>_{\text{fosd}}$. On the other hand, property (ii) is a normalization condition which implies $c(\delta_{c(p)}) = c(p)$. Thus, if we think of c as a utility function that represents the DM's preferences over simple lotteries, it follows that the utility of the certain prize $c(p)$ is the same as the utility of the lottery p, meaning that $c(p)$ is the certainty equivalent of p. Every transitive, complete and continuous preference relation on $\Delta(X)$ can be represented by such a normalized utility function provided that the relation is also monotonic with respect to $\geq_{\text{fosd}}$. Specifically, given any continuous function $u : \Delta(X) \to \mathbb{R}$ that represents such a preference relation, we can let $c(p) := v^{-1}(u(p))$ for every $p \in \Delta(X)$, where $v : X \to \mathbb{R}$ is defined as $v(x) := u(\delta_x)$ for $x \in X$. It is also clear that c is uniquely defined given the preference relation over $\Delta(X)$.

For $f \in \mathcal{H}_X$ and $\pi \in \Delta(S)$, set

$$\pi_f := \sum_{s \in S} \pi(s) f(s).$$

π_f is the lottery induced by the purely subjective act f in the event that the states are distributed according to π. Our DM assigns the probability $\mu(\pi)$ to this event, and hence, thinks of f as a compound lottery that returns the lottery π_f with probability $\mu(\pi)$. Then, in a recursive fashion, she replaces π_f with its certainty equivalent, $c(\pi_f)$. Thereby, the DM reduces the compound lottery in question into a simple lottery that returns $c(\pi_f)$ with probability $\mu(\pi)$. To summarize, an act $f \in \mathcal{H}_X$ is equivalent to the following lottery in $\Delta(X)$:

$$\mu_f := \sum_{\pi \in \Delta(S)} \mu(\pi) \delta_{c(\pi_f)}.$$

Moreover, under a classical monotonicity assumption (see property (A2) in Appendix A), any $f \in \mathcal{H}$ is equivalent to the purely subjective act that returns $\delta_{c(f(s))}$ at every $s \in S$. Let us denote the latter act as $c \circ f$, and set $\mu_f := \mu_{c \circ f}$ for $f \in \mathcal{H}$.

We end up with the following representation upon extending the recursive method of evaluation above to lotteries over acts.

Definition 2. A *recursive representation* for $\succsim$ consists of a $\mu \in \Delta^2(S)$ and a certainty equivalence function c such that, for every $F, G \in \Delta(\mathcal{H})$,

$$F \succsim G \quad \Leftrightarrow \quad c\left(\sum_{f \in \mathcal{H}} F(f) \delta_{c(\mu_f)}\right) \geq c\left(\sum_{f \in \mathcal{H}} G(f) \delta_{c(\mu_f)}\right).$$

A *recursive preference* refers to a binary relation $\succsim$ on $\Delta(\mathcal{H})$ that can be represented with a pair (μ, c) as above.

Let us write $U(F)$ in place of $c\left(\sum_{f\in\mathcal{H}} F(f)\delta_{c(\mu_f)}\right)$. Observe that $U(D_f) = c\left(\delta_{c(\mu_f)}\right) = c(\mu_f)$ for any $f \in \mathcal{H}$. Moreover, it is easily checked that $f = \mathbf{1}_S p$ implies $\mu_f = \delta_{c(p)}$. Hence, $U(p) \equiv U(D_{\mathbf{1}_S p}) = c(\delta_{c(p)}) = c(p)$ for any $p \in \Delta(X)$. In particular, $p \succsim q$ if and only if $c(p) \geq c(q)$ for any $p, q \in \Delta(X)$. It also follows that $U(f) \equiv U(D_f) = c(\mu_f) = U(\mu_f)$ for any $f \in \mathcal{H}$. Thus, $f \sim \mu_f$, just as I claimed earlier. Finally, for a compound lottery $P \in \Delta^2(X)$, we have $U(P) \equiv c\left(\sum_{p\in\Delta(X)} P(p)\delta_{c(\mu_{\mathbf{1}_S p})}\right) = c\left(\sum_{p\in\Delta(X)} P(p)\delta_{c(p)}\right) = U\left(\sum_{p\in\Delta(X)} P(p)\delta_{c(p)}\right)$. That is,

$$P \sim \sum_{p\in\Delta(X)} P(p)\delta_{c(p)}. \tag{1}$$

This means that the DM reduces a compound lottery into a simple lottery in a recursive fashion, just as in the evaluation of purely subjective acts.

While Segal's (1987) original formulation of recursive preferences focuses on purely subjective acts, the formulation above incorporates the lottery-valued acts into the theory via the (monotonicity) assumption $f \sim c \circ f$. This assumption seems to be coherent with the general logic of recursive preferences. Suppose, for example, that the DM believes that the states are distributed according to a first-order distribution π. Then, $c \circ f \sim \pi_{c\circ f}$, whereas $\pi_{c\circ f} := \sum_{s\in S} \pi(s)\delta_{c(f(s))}$. Moreover, by property (1), the lottery $\sum_{s\in S} \pi(s)\delta_{c(f(s))}$ is equivalent to the compound lottery that returns $f(s)$ with probability $\pi(s)$. Hence, in this case, $f \sim c \circ f$ means that the DM is indifferent between f and the compound lottery that returns $f(s)$ with probability $\pi(s)$.

Let us now turn to the interplay between risk and ambiguity attitudes.

3 Global Ambiguity Aversion

Throughout the remainder of the paper, by a *preference relation* I mean a complete and transitive binary relation. Following Dean and Ortoleva (in press), a preference relation $\succsim$ on $\Delta(\mathcal{H})$ is said to be *ambiguity neutral* if there exists a $\bar{\pi} \in \Delta(S)$ such that

$$f \sim \bar{\pi}_f \quad \forall f \in \mathcal{H}_X. \tag{2}$$

Intuitively, this means that the DM converts subjective uncertainty (or ambiguity) into risk using the distribution $\bar{\pi}$. Alternatively, the DM behaves as if she is probabilistically sophisticated in the sense of Machina and Schmeidler (1992, 1995).[8,9]

For recursive preferences, a natural way to model such behavior is to select a degenerate second-order belief. Indeed, $\mu = D_\pi$ implies $\mu_f = \delta_{c(\pi_f)}$, and hence, $f \sim \pi_f$ for every $f \in \mathcal{H}_X$.

Definition 3. Let $\succsim$ and $\succsim'$ be a pair of preference relations on $\Delta(\mathcal{H})$. $\succsim$ is *more ambiguity averse than* $\succsim'$ (or, equivalently, $\succsim'$ is *more ambiguity loving than* $\succsim$) if:

(i) $f \succsim \delta_x \quad \Rightarrow \quad f \succsim' \delta_x \quad \forall f \in \mathcal{H}_X$ and $x \in X$.

(ii) $p \succsim q \quad \Leftrightarrow \quad p \succsim' q \quad \forall p, q \in \Delta(X)$.

In turn, $\succsim$ is (*absolutely*) *ambiguity averse* (resp. *loving*) if it is more ambiguity averse (resp. loving) than an ambiguity neutral preference.

Conditions (i) and (ii) mean that $\succsim$ displays a weaker desire for uncertain acts than $\succsim'$, while the two relations agree on the ranking of lotteries. This definition of relative ambiguity aversion has become a standard approach since the seminal work of Ghirardato and Marinacci (2002). Defining absolute ambiguity aversion relative to ambiguity neutral preferences, as above, is also a fairly standard practice.[10]

Ambiguity aversion (or neutrality), by itself, does not impose any restriction on risk preferences. However, as we shall see momentarily, the picture changes radically if we demand the DM to be ambiguity averse in a global sense, as follows.

Definition 4. A preference relation $\succsim_c$ on $\Delta(X)$ represented by a certainty equivalence function c has the *global ambiguity aversion property* if the recursive preference represented by (μ, c) is ambiguity averse for any (finite) state space S and any $\mu \in \Delta^2(S)$.

Global ambiguity aversion is a robustness criterion that demands the DM to exhibit ambiguity aversion irrespective of her second-order belief and the number of states. While, in reality, ambiguity aversion may rarely be so robust, as noted in the introduction, in potential applications the analyst may well want to focus on risk preferences that possess this property.

Recall that according to a recursive representation (μ, c), the utility of a compound lottery P is given by $U(P) = c\left(\sum_{p \in \Delta(X)} P(p)\delta_{c(p)}\right)$, which depends only on c. Let $\succsim_c^2$

[8] As a minor difference, Machina and Schmeidler (1995) go a step further and demand property (2) to hold for every $f \in \mathcal{H}$, where $\bar{\pi}_f := \sum_{s \in S} \bar{\pi}(s) f(s)$. This property is too demanding for my purposes because, when combined with the assumption $f \sim c \circ f$, it entails some independence properties for risk preferences.

[9] Conceptually, Dean-Ortoleva definition of ambiguity neutrality agrees with that of Epstein (1999), except that Epstein works in a Savagean setup with purely subjective acts.

[10] As a key difference, Ghirardato and Marinacci (2002) propose to take as ambiguity neutral the class of subjective expected utility preferences. In models of ambiguity with expected utility preferences over lotteries, this approach seems to produce perfectly sensible predictions. For my purposes Dean-Ortoleva approach is more suitable because it is tailored for non-expected utility preferences over lotteries.

denote the preference relation on $\Delta^2(X)$ represented by the same function, $U(\cdot)$. The following two properties about risk preferences were introduced by Dillenberger (2010).

Negative Certainty Independence (NCI). For every $p, q \in \Delta(X)$, $x \in X$ and $\alpha \in [0,1]$,

$$p \succsim_c \delta_x \quad \Rightarrow \quad \alpha p + (1-\alpha)q \succsim_c \alpha\delta_x + (1-\alpha)q.$$

Preference For One-Shot Resolution of Uncertainty (PORU). For every $P \in \Delta^2(X)$,

$$\sum_{p\in\Delta(X)} P(p)p \succsim_c^2 P.$$

As noted by Dillenberger (2010) and Cerreia-Vioglio et al. (2015), NCI accommodates Allais' common ratio and common consequence effects under risk. The main idea that underlies this axiom is that a certain prize have an intrinsic appeal, which disappears upon mixing that prize with another lottery. Hence, if p is better than δ_x despite the certainty appeal of the latter, then a mixture of p with q must also be better than the corresponding mixture of δ_x with q. This is the content of NCI.

To interpret PORU, note that the overall probability of the prize x under the compound lottery P can be computed as $\sum_{p\in\Delta(X)} P(p)p(x)$. Thus, $\sum_{p\in\Delta(X)} P(p)p$ is the *reduced form of* P from a statistical point of view. PORU asserts that the DM should prefer the reduced form a compound lottery to that lottery itself. If a compound lottery is viewed as a dynamic stochastic process with two stages, this axiom describes a preference for one-shot resolution of uncertainty, as opposed to gradual resolution.

Dillenberger (2010) proves that NCI and PORU are equivalent to each other. The next theorem takes a step further: NCI and PORU are equivalent to the global ambiguity aversion property.

Theorem 1. *Let $\succsim_c$ be a preference relation on $\Delta(X)$ represented by a certainty equivalence function c. The following three statements are equivalent.*

(i) *$\succsim_c$ satisfies* NCI*.*

(ii) *$\succsim_c^2$ satisfies* PORU*.*

(iii) *$\succsim_c$ has the global ambiguity aversion property.*

I proceed with some preliminary observations to provide insight into Theorem 1.

Lemma 1. *Given a state space S, let $\succsim$ be a recursive preference on $\Delta(\mathcal{H})$ represented by (μ, c). Then, $\succsim$ is ambiguity averse if and only if there exists a $\bar{\pi} \in \Delta(S)$ such that*

$$\bar{\pi}_f \succsim \mu_f \quad \forall f \in \mathcal{H}_X.$$

Since $\mu_f \sim f$, Lemma 1 means that an ambiguity averse DM would attach larger values to uncertain acts if she were able to form a first-order belief $\bar{\pi}$. Given a recursive preference $\succsim$ that is ambiguity averse, a *benchmark* belief (or distribution) refers to such a $\bar{\pi}$.

Just as the reduced form of a compound lottery, the *reduced form of a second-order belief* μ is the first-order distribution $\bar{\mu}$ defined as

$$\bar{\mu} := \sum_{\pi \in \Delta(S)} \mu(\pi)\pi. \tag{3}$$

Alternatively, for any state s, $\bar{\mu}(s) = \sum_{\pi \in \Delta(S)} \mu(\pi)\pi(s)$ is the expectation of $\pi(s)$ with respect to μ. Thus, $\bar{\mu}$ can also be viewed as the *mean of* μ.

A central issue in this paper is that Definition 3 (or equivalently, Lemma 1) does not impose a restriction on first-order distributions that can act as a benchmark. Following Segal (1987), one can also think of an alternative definition that takes $\bar{\mu}$ as "the" benchmark:

$$\bar{\mu}_f \succsim \mu_f \quad \forall f \in \mathcal{H}_X. \tag{4}$$

In what follows, I say that a recursive preference $\succsim$ represented by (μ, c) is *mean ambiguity averse* if the property (4) holds. In Section 3.2 below, I will show that this alternative definition is more restrictive than Definition 3, both conceptually and behaviorally.

Since $\bar{\mu}$ is the reduced form of the belief μ, PORU—preference for reduced form lotteries—implies a global form of mean ambiguity aversion. This is the content of the related observation of Artstein-Avidan and Dillenberger (2011). Consequently, NCI also implies mean ambiguity aversion, in a global sense. For the sake of completeness, I provide a short proof of this fact in the proof of Theorem 1, in Appendix C.

The contribution of Theorem 1 is the converse implication, that global ambiguity aversion implies NCI. A key observation with regard to this part of the theorem is that for a special class of second-order beliefs, Definition 3 is equivalent to mean ambiguity aversion. Suppose that the state space S can be partitioned into n sets, $S^1, ..., S^n$, each having the same cardinality. Let π^i denote the uniform distribution over S^i and consider a second-order belief μ of the form $\mu = \sum_{i=1}^n \alpha^i D_{\pi^i}$ for some $\{\alpha^1, ..., \alpha^n\} \subseteq [0,1]$ with $\sum_{i=1}^n \alpha^i = 1$. I refer to such a second-order belief μ as *uniformly separable*.

The proof of Theorem 1 shows that for a uniformly separable μ, the only possible benchmark is the mean, $\bar{\mu}$. The following example illustrates this fact in a particular case with four states.

Example 1. Suppose S consists of four distinct points, $s^{11}, s^{12}, s^{21}, s^{22}$. Let $\mu = \frac{1}{2}D_{\pi^1} + \frac{1}{2}D_{\pi^2}$, where π^1 and π^2 are the uniform distributions supported over $S^1 = \{s^{11}, s^{12}\}$ and

$S^2 = \{s^{21}, s^{22}\}$, respectively. For any $S' \subseteq S$, let $f^{S'}$ denote the act defined as $f^{S'}(s) := \delta_{x^*}$ for $s \in S'$, and $f^{S'}(s) := \delta_{x_*}$ for $s \in S \backslash S'$. Then

$$\pi_{f^{S'}} = \pi(S')\delta_{x^*} + (1 - \pi(S'))\, \delta_{x_*} \quad \forall \pi \in \Delta(S), \tag{5}$$

while $\mu_{f^{S'}} = \frac{1}{2}\delta_{c(\pi^1_{f^{S'}})} + \frac{1}{2}\delta_{c(\pi^2_{f^{S'}})}$.

Fix a certainty equivalence function c, and denote by $\succsim$ the recursive preference represented by (μ, c). Let us now show that for any S' with two states, we have

$$\mu_{f^{S'}} \sim \frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*}. \tag{6}$$

Indeed, $S' = S^1$ implies $\pi^1_{f^{S'}} = \delta_{x^*}$ and $\pi^2_{f^{S'}} = \delta_{x_*}$, so that $\mu_{f^{S'}} = \frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*}$. Symmetrically, $S' = S^2$ also implies $\mu_{f^{S'}} = \frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*}$. It remains only one case to consider: $S' \cap S^1 = \{s^{1i}\}$ for some $i \in \{1, 2\}$, while $S' \cap S^2 = \{s^{2j}\}$ for some $j \in \{1, 2\}$. In this case, $\pi^1_{f^{S'}} = \pi^1(S')\delta_{x^*} + (1 - \pi^1(S'))\, \delta_{x_*} = \pi^1(s^{1i})\delta_{x^*} + (1 - \pi^1(s^{1i}))\, \delta_{x_*} = \frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*}$. Similarly, $\pi^2_{f^{S'}}$ also equals $\frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*}$, implying $\mu_{f^{S'}} = \delta_{c(\frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*})} \sim \frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*}$.

By (5) and (6), we have $\bar{\pi}_{f^{S'}} \succsim \mu_{f^{S'}}$ if and only if $\bar{\pi}(S') \geq 1/2$. In turn, if the latter inequality holds for every S' with two states, then $\bar{\pi}$ is necessarily equal to the uniform distribution over S, which is nothing but $\bar{\mu} := \frac{1}{2}\pi^1 + \frac{1}{2}\pi^2$. $\square$

A risk preference $\succsim_c$ with the global ambiguity aversion property must also induce an ambiguity averse preference for every uniformly separable belief μ. As we just discussed, for such a belief, the only possible benchmark is $\bar{\mu}$. Thus, if $\succsim_c$ has the global ambiguity aversion property, then

$$\bar{\mu}_f \succsim_c \mu_f \quad \text{for any } f \in \mathcal{H}_X \text{ and any uniformly separable } \mu. \tag{7}$$

The main step in the remainder of the proof of Theorem 1 is to show that the class of uniformly separable beliefs is rich enough to derive NCI from property (7). Specifically, given some arbitrarily selected $p, q \in \Delta(X)$ and $\alpha \in [0, 1]$, if $p(x)$ is a rational number for every x, there exist a uniformly separable belief μ on a suitably selected set S and an act $f \in \mathcal{H}_X$ such that

$$\bar{\mu}_f = \alpha p + (1 - \alpha)q \quad \text{and} \quad \mu_f = \alpha\delta_{c(p)} + (1 - \alpha)q.$$

By (7), this yields $\alpha p + (1 - \alpha)q \succsim_c \alpha\delta_{c(p)} + (1 - \alpha)q$, while monotonicity with respect to $\geq_{\text{fosd}}$ implies $\alpha\delta_{c(p)} + (1 - \alpha)q \succsim_c \alpha\delta_x + (1 - \alpha)q$ for any x with $p \succsim_c \delta_x$. It follows that $\alpha p + (1 - \alpha)q \succsim_c \alpha\delta_x + (1 - \alpha)q$ for any such x, as demanded by NCI.

3.1 Cautious Expected Utility

Under some mild regularity assumptions, Cerreia-Vioglio et al. (2015) show that a preference relation on the closure of $\Delta(X)$ satisfies NCI if and only if it can be represented by a certainty equivalence function of the form $c(p) = \inf_{v \in W} v^{-1}(E_p(v))$ for some nonempty $W \subseteq C_\uparrow(X)$, where $C_\uparrow(X)$ is the set of all continuous, real functions on X that are strictly increasing, while $E_p(v) := \sum_{x \in X} p(x)v(x)$. This representation depicts a DM who behaves as if she is unsure how to evaluate a given lottery. The DM has in mind several von Neumann-Morgenstern functions, v; computes the certainty equivalent, $v^{-1}(E_p(v))$, of a given lottery p according to each v; and then, in a cautious way, she selects the worst certainty equivalent to evaluate the lottery. Hence the name *cautious expected utility* (CEU) representation.

We obtain the following result as a corollary of Theorem 1 and the representation theorem of Cerreia-Vioglio et al. (2015).

Corollary 1. *Let $\succsim_c$ be a preference relation on $\Delta(X)$ represented by a certainty equivalence function c that is uniformly continuous on $\Delta(X)$.*[11] *The following two statements are equivalent.*

(i) *$\succsim_c$ has the global ambiguity aversion property.*

(ii) *There exists a nonempty set $W \subseteq C_\uparrow(X)$ such that $c(p) = \inf_{v \in W} v^{-1}(E_p(v))$ for every $p \in \Delta(X)$.*

Here, c is demanded to be uniformly continuous because we need to extend this function (continuously) to the closure of $\Delta(X)$ to be able to apply the representation theorem of Cerreia-Vioglio et al. (2015).[12] For brevity, I omit the proof of Corollary 1, which boils down to showing that if c satisfies NCI on $\Delta(X)$ and is uniformly continuous, then the extension of c to the closure of $\Delta(X)$ also satisfies NCI on this larger domain.

3.2 Ambiguity Aversion vs Mean Ambiguity Aversion

In general, given a DM who does not reduce compound lotteries in a standard way, there seems to be no reason to attach a special importance to the reduced form of the DM's belief. On the other hand, it is not so clear what we gain in practice from the generality

[11] Recall that the topology of weak convergence is metrizable. Let d denote a compatible metric on $\Delta(X)$. c is *uniformly continuous on* $\Delta(X)$ if for each $\varepsilon > 0$ there exists a $\gamma > 0$ such that $d(p, q) < \gamma$ implies $|c(p) - c(q)| < \varepsilon$ for every $p, q \in \Delta(X)$. The existence of such a uniformly continuous certainty equivalence function can be characterized along the lines of Kopylov (2016).

[12] The proof of Cerreia-Vioglio et al. builds upon the expected multi-utility theorem of Dubra, Maccheroni and Ok (2004). The latter theorem focuses on a continuous preorder on the closure of $\Delta(X)$, and the compactness of this set is crucial for the theorem (see Evren, 2008).

embodied in Definition 3. In this section, I will show that there do exist ambiguity averse recursive representations (μ, c) with a benchmark that is distinct from $\bar{\mu}$. Since NCI implies mean ambiguity aversion, in such cases, the risk preference $\succsim_c$ must violate NCI. By Theorem 1, then, such c will induce non-ambiguity averse behavior in some other context, with a different second-order belief. To summarize, it follows that the difference between Definition 3 and mean ambiguity aversion matters at a local level, but not at a global level.

I will illustrate the local distinction between the two definitions with the *rank dependent utility* (RDU) model. Given a $p \in \Delta(X)$, let $x^1, ..., x^m$ denote the points in the support of p, where $x^1 \leq x^2 \leq \cdots \leq x^m$. An RDU functional $u : \Delta(X) \to \mathbb{R}$ is defined by a pair of strictly increasing functions $v : X \to \mathbb{R}$ and $\Psi : [0,1] \to [0,1]$, with $\Psi(0) = 0$ and $\Psi(1) = 1$, such that

$$u(p) = v(x^1) + \sum_{j=2}^{m} \left(v(x^j) - v(x^{j-1})\right) \Psi\left(\sum_{i=j}^{m} p(x^i)\right).$$

In turn, $c(p) := v^{-1}(u(p))$ gives the corresponding certainty equivalence function.

The function Ψ reflects how the DM distorts probabilities in her mind. When this function is convex, the DM distorts probabilities in a pessimistic way, by effectively increasing the probability of smaller prizes.

Suppose there are two states of nature. Let $\succsim$ denote a recursive preference on $\Delta(\mathcal{H})$ represented by (μ, c) where c is as in the RDU model above. Assuming that μ is symmetric around its mean $\bar{\mu}$, Segal's (1987) Theorem 4.2 provides sufficient conditions on the function Ψ that imply mean ambiguity aversion. Specifically, this theorem demands Ψ to be a convex function such that (i) the elasticity of Ψ is non-decreasing in the sense that $\Psi(\lambda)\Psi(\lambda') \leq \Psi(\lambda\lambda')$ for every $\lambda, \lambda' \in [0,1]$; and (ii) the elasticity of the function $\lambda \to 1 - \Psi(1-\lambda)$ is non-increasing for $\lambda \in [0,1]$.

As Segal also notes, it is easy to construct a convex function Ψ that violates the elasticity assumptions above. In such cases, the conclusion of Segal's Theorem 4.2 may fail, meaning that $\succsim$ need not be mean ambiguity averse. In Appendix B, I give an example of this sort, where $\bar{\mu}$ is distinct from the uniform distribution over the two states. Let $\bar{\pi}$ denote the uniform distribution. The preference relation $\succsim$ in this example satisfies the property $\bar{\pi}_f \succsim \mu_f$ for every $f \in \mathcal{H}_X$. (In fact, $\bar{\pi}_f \succ \mu_f$ whenever f is not constant across the states.) Thus, the relation $\succsim$ is ambiguity averse but not mean ambiguity averse.

To see what this may entail in practice, consider an Ellsberg-type experiment with two urns each containing a given number of balls. At a later stage, the experimenter will randomly extract a ball from each urn. Each ball is either blue (b) or orange (o). The composition of urn 1 (the *risky urn*) is known to the subject, and the fraction of blue balls in this urn is a control variable, denoted as β. Urn 2 is *ambiguous* in the sense that its exact composition is unknown. Yet, the subject has a bit of further information. Given a small number ε in the interval $(0, 1/6)$, the subject is told that:

(I) The fraction of blue balls in urn 2 is greater than or equal to 2ε.

As usual, each state s represents the event that the ball extracted from the ambiguous urn is of color s. Thus, $S := \{b, o\}$. Pick a pair of prizes x and x' with $x > x'$. For $i = 1, 2$, let f^{bi} denote the bet that pays x if the color of the ball extracted from urn i is b, and that pays x' otherwise. f^{oi} is defined analogously. Put differently,

$$f^{b2}(s) := \begin{cases} \delta_x & \text{if } s = b \\ \delta_{x'} & \text{if } s = o \end{cases}, \quad f^{o2}(s) := \begin{cases} \delta_x & \text{if } s = o \\ \delta_{x'} & \text{if } s = b \end{cases},$$

while

$$f^{b1} := \beta\delta_x + (1-\beta)\delta_{x'}, \quad f^{o1} := (1-\beta)\delta_x + \beta\delta_{x'}.$$

The example that I construct in Appendix B is based on a particular second-order belief μ. Specifically, $\mu = \frac{1}{2}D_{\pi^1} + \frac{1}{2}D_{\pi^2}$, where the distributions π^1 and π^2 are as in the following table.

	b	o
π^1	$\frac{1}{2} + 3\varepsilon$	$\frac{1}{2} - 3\varepsilon$
π^2	$\frac{1}{2} - \varepsilon$	$\frac{1}{2} + \varepsilon$

Since $\mu(\pi^1) = \mu(\pi^2) = 1/2$, the mean $\bar{\mu}$ is the distribution that sits in the middle of the interval between π^1 and π^2. Thus, $\bar{\mu}(b) = \frac{1}{2}\pi^1(b) + \frac{1}{2}\pi^2(b) = \frac{1}{2} + \varepsilon$, while $\bar{\mu}(o) = \frac{1}{2} - \varepsilon$. Note that $\bar{\mu}(b)$ also sits in the middle of the interval $[2\varepsilon, 1]$, which is quite reasonable given the property (I).

As before, let $\bar{\pi}$ denote the uniform distribution, so that $\bar{\pi}(b) = \bar{\pi}(o) = 1/2$. The convex function Ψ that I describe in Appendix B implies

$$f^{o2} \succ \bar{\mu}_{f^{o2}} \quad \text{and} \quad f^{s2} \prec \bar{\pi}_{f^{s2}} \text{ for } s = b, o. \tag{8}$$

Observe that given a specific value of β and the distribution $\pi \in \Delta(S)$ with $\pi(b) = \beta$, we have $\pi_{f^{s2}} = f^{s1}$ for $s = b, o$. In particular, if $\beta = \bar{\mu}(b)$, then $\bar{\mu}_{f^{o2}} = f^{o1}$. Thus, the left hand side of (8) means that when the fraction of blue balls in the risky urn equals $\bar{\mu}(b)$, i.e., $\frac{1}{2} + \varepsilon$, then for bets on orange balls, the subject strictly prefers the ambiguous urn to the risky one. This follows from the absence of mean ambiguity aversion.

Similarly, with $\beta = \bar{\pi}(b)$, we have $\bar{\pi}_{f^{s2}} = f^{s1}$ for $s = b, o$. Hence, the right hand side of (8) means that when the risky urn contains an equal number of blue and orange balls, then—just as in a classical Ellsberg paradox with two urns—the subject strictly prefers the risky urn to the ambiguous one, irrespective of the color that she is betting on.

Finally, note that $f^{b1} \succsim \bar{\pi}_{f^{b2}}$ for $\beta \geq 1/2$, while $f^{o1} \succsim \bar{\pi}_{f^{o2}}$ for $\beta \leq 1/2$. Thus, the right hand side of (8) also implies that

$$\nexists \beta \in [0,1] \text{ with } f^{s2} \succ f^{s1} \text{ for } s = b, o. \tag{9}$$

In other words, the subject will not exhibit the opposite of the Ellsberg paradox, irrespective of the composition of the risky urn. (On a related note, the expression (9) remains true upon replacing $\succ$ with $\succsim$, which means that the subject is not ambiguity loving according to Definition 3.)

More generally, in the context of similar experiments with two urns, if the subject weakly prefers any given bet on the risky urn to the corresponding bet on the ambiguous urn, we can immediately conclude that the subject is ambiguity averse according to Definition 3. This holds true irrespective of the composition of the risky urn, and precludes the possibility that an alternative composition may lead to a paradoxical form of ambiguity loving behavior. By contrast, Ellsberg-type choices can be taken as evidence of mean ambiguity aversion only if the composition of the *risky* urn coincides with the DM's mean belief, $\bar{\mu}$, about the *ambiguous* urn. In light of the example above, even a classical form of the Ellsberg paradox with a 50-50 distributed risky urn does not guarantee mean ambiguity aversion if there is a reason to suspect a mean belief about the ambiguous urn that is distinct from the 50-50 distribution. (Indeed, the only role of property (I) above is to motivate such a mean belief.) Thus, it seems fair to conclude that, at a local level, the notion of mean ambiguity aversion is too demanding.

3.3 Non-Robustness of RDU

The discussion above also attests to the fact that, at a local level, there is no shortage of ambiguity averse preferences within the recursive RDU model. Yet, according to Theorem 1, the recursive CEU model is the only subclass of recursive preferences that robustly generates ambiguity averse behavior. In this section, I will give an example which shows that an RDU preference may induce non-ambiguity averse behavior, even if it satisfies all assumptions in Segal's (1987) Theorem 4.2. As usual, the absence of ambiguity aversion means that the preference relation does not admit any benchmark. To this end, following the proof of Theorem 1, I will utilize a uniformly separable belief, which also necessitates four states, as opposed to the case of binary states considered by Segal.

Example 2. Consider a modified version of the experiment in Section 3.2 so that there are four different colors: dark blue (db), light blue (lb), dark yellow (dy) and light yellow (ly). As before, the exact composition of the risky urn (urn 1) is known to the subject.

Further, the subject is told that urn 2 contains four balls in total, an equal number of dark and light blue balls, and an equal number of dark and light yellow balls. However, she is not given any information about the ratio of the total number of blue balls to that of yellow balls, which makes urn 2 ambiguous. The set of states is $S := \{db, lb, dy, ly\}$, and each state represents the event that the corresponding color will be extracted from urn 2.

The subject's preference relation admits a recursive representation (μ, c), where c is as in the RDU model defined in Section 3.2. More specifically, Ψ is a strictly convex function, which may as well satisfy the aforementioned elasticity assumptions in Theorem 4.2 of Segal (1987).

The information about the ambiguous urn leaves three cases to consider: (B) all balls in this urn may be blue, two of them being dark blue and two of them light blue; (Y) the symmetric case in which all balls are yellow; (M) the urn may contain exactly one ball of each color. Let π^K denote the distribution on S that corresponds to case K, so that

$$\pi^B(db) = \pi^B(lb) = \frac{1}{2}, \quad \pi^Y(dy) = \pi^Y(ly) = \frac{1}{2} \quad \text{and} \quad \pi^M(s) = \frac{1}{4} \ \forall s \in S.$$

The subject happens to attach zero probability to π^M. So, $\mu := \frac{1}{2}D_{\pi^B} + \frac{1}{2}D_{\pi^Y}$, which is a uniformly separable belief. Hence, as we have seen in Example 1 above, the only distribution that can act as a benchmark is $\bar{\mu} := \frac{1}{2}\pi^B + \frac{1}{2}\pi^Y = \pi^M$. It remains to show that $\bar{\mu}$ cannot be a benchmark either.

To this end, let $x := c\left(\frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*}\right)$, and define a non-binary bet f^* on urn 2 as

$$f^*(s) = \begin{cases} x & \text{if } s \in \{db, lb\}, \\ x^* & \text{if } s = dy, \\ x_* & \text{if } s = ly. \end{cases}$$

Set $p := \frac{1}{2}\delta_{x^*} + \frac{1}{2}\delta_{x_*}$. Observe that for any $\pi \in \Delta(S)$, we have $\pi_{f^*} = \pi(dy)\delta_{x^*} + \pi(\{db, lb\})\delta_x + \pi(ly)\delta_{x_*}$. Thus,

$$\pi^B_{f^*} = \delta_x, \quad \pi^Y_{f^*} = p, \quad \text{and} \quad \pi^M_{f^*} = \frac{1}{4}\delta_{x^*} + \frac{1}{2}\delta_x + \frac{1}{4}\delta_{x_*}.$$

By the first two equalities, $\mu_{f^*} = \frac{1}{2}\delta_{c(\delta_x)} + \frac{1}{2}\delta_{c(p)}$. Since $x = c(p)$, it follows that $\mu_{f^*} = \delta_x$.

Moreover, $x = c(p)$ also implies $v(x) = u(p) = v(x_*) + (v(x^*) - v(x_*))\,\Psi(\frac{1}{2})$. Using the latter equality and the definition of $u(\pi^M_{f^*})$, it can easily be seen that

$$v(x) - u(\pi^M_{f^*}) = (v(x^*) - v(x_*))\left(\Psi(\tfrac{1}{2})\left(1 - \Psi(\tfrac{3}{4})\right) - \left(1 - \Psi(\tfrac{1}{2})\right)\Psi(\tfrac{1}{4})\right). \tag{10}$$

Note that, by strict convexity of Ψ,

$$\frac{\Psi(1/2)}{0.5} > \frac{\Psi(1/4)}{0.25} \quad \text{and} \quad \frac{1-\Psi(3/4)}{0.25} > \frac{1-\Psi(1/2)}{0.5}.$$

Upon multiplying these two inequalities and canceling the term 0.5×0.25, we get

$$\Psi(1/2)(1-\Psi(3/4)) > \Psi(1/4)(1-\Psi(1/2)).$$

Thus, equation (10) implies $v(x) > u(\pi^M_{f^*})$, which means $\delta_x \succ \pi^M_{f^*}$. Since $\mu_{f^*} = \delta_x$ and $\bar{\mu}_{f^*} = \pi^M_{f^*}$, it follows that $\mu_{f^*} \succ \bar{\mu}_{f^*}$. So, $\bar{\mu}$ is not a benchmark either. $\square$

As in Section 3.2, in this example, we can think of the color distribution in the risky urn as a potential benchmark. Hence, the absence of ambiguity aversion means that irrespective of the specification of the risky urn, there always exists a bet on the ambiguous urn that the subject strictly prefers to the corresponding bet on the risky urn. This behavior is not compatible with any first-order belief about the states. Indeed, such a belief would make the DM indifferent between the two urns at least for one specification of the risky urn, namely, the one that coincides with the DM's belief about the ambiguous urn. In this sense, the subject exhibits a paradoxical form of ambiguity loving.[13]

It is also useful to compare Example 2 with Dillenberger's (2010) Proposition 3. The latter result shows that any RDU preference over lotteries that violates the independence axiom must also violate NCI. Given the link between NCI and the mean ambiguity aversion property, this result can be interpreted as saying that the recursive preference induced by an RDU functional will violate the mean ambiguity aversion property, at least for some second-order beliefs. However, by itself, this cannot be taken as evidence of a form of ambiguity loving, or the absence of ambiguity aversion. Indeed, such a preference relation can even exhibit a classical form of the Ellsberg paradox as we have seen in Section 3.2. In Example 2, a uniformly separable belief closes the gap between Definition 3 and mean ambiguity aversion. It is this feature of the example that leads to a paradoxical form of ambiguity loving from the absence of mean ambiguity aversion.

4 Global Ambiguity Neutrality

The following is a straightforward extension of Definition 4.

Definition 5. A preference relation $\succsim_c$ on $\Delta(X)$ represented by a certainty equivalence function c has the *global ambiguity neutrality* (resp. *loving*) *property* if the recursive pref-

[13] This observation can also be viewed as an additional argument as to why Definition 3 outperforms the concept of mean ambiguity aversion, locally.

erence represented by (μ, c) is ambiguity neutral (resp. loving) for any state space S and any $\mu \in \Delta^2(S)$.

In this section, as a first order of business, I shall utilize Theorem 1 to obtain a characterization of global ambiguity neutrality. Specifically, we will see that this property is equivalent to the independence axiom. To this end, the first point to note is that:

Lemma 2. *Given a state space S, a recursive preference on $\Delta(\mathcal{H})$ is ambiguity neutral if and only if it is ambiguity averse and loving.*

This lemma is a simple consequence of the fact that a recursive preference is monotonic with respect to $\geq_{\text{fosd}}$ over $\Delta(X)$. (For the details, see the proof of Lemma 2 in Appendix C.)

While Theorem 1 focuses on global ambiguity aversion, a dual of this result can also be established with symmetric arguments. That is, *the following three statements are equivalent.*

(i) $\succsim_c$ *satisfies the dual of* NCI: $\delta_x \succsim_c p \quad \Rightarrow \quad \alpha\delta_x + (1-\alpha)q \succsim_c \alpha p + (1-\alpha)q$.

(ii) $\succsim_c^2$ *satisfies the dual of* PORU: $P \succsim_c^2 \sum_{p\in\Delta(X)} P(p)p$.

(iii) $\succsim_c$ *has the global ambiguity loving property.*

By Lemma 2, global ambiguity neutrality is equivalent to the conjunction of global ambiguity aversion and loving. In turn, NCI and its dual in statement (i) above are jointly equivalent to the classical **independence axiom**:

$$p \succsim_c r \quad \Rightarrow \quad \alpha p + (1-\alpha)q \succsim_c \alpha r + (1-\alpha)q.$$

Indeed, by monotonicity w.r.t. $\geq_{\text{fosd}}$, it is plain that this axiom is equivalent to the condition $\alpha p + (1-\alpha)q \sim_c \alpha\delta_{c(p)} + (1-\alpha)q$, and that the latter property is equivalent to the conjunction of NCI with its dual. Finally, note that PORU and its dual are jointly equivalent to the **reduction of compound lotteries axiom**:

$$P \sim_c^2 \sum_{p\in\Delta(X)} P(p)p.$$

To summarize, we obtain the following characterization of global ambiguity neutrality.

Corollary 2. *Let $\succsim_c$ be a preference relation on $\Delta(X)$ represented by a certainty equivalence function c. The following three statements are equivalent.*

(i) $\succsim_c$ *satisfies the independence axiom.*

(ii) $\succsim_c^2$ *satisfies the reduction of compound lotteries axiom.*

(iii) $\succsim_c$ *has the global ambiguity neutrality property.*

Given a recursive representation (μ, c), if $\succsim_c$ satisfies the independence axiom, then μ can be replaced with $D_{\bar{\mu}}$ (or, with the mean belief $\bar{\mu}$) without altering the associated preference relation over $\Delta(\mathcal{H})$. Thus, the recursive expected utility model characterized in Corollary 2 is nothing but Anscombe and Aumann's (1963) expected utility theory.

As noted in the discussion of Example 2, the absence of ambiguity aversion, according to Definition 3, is equivalent to a paradoxical form of ambiguity loving, "paradoxical" in the sense that the behavior cannot be explained with any first-order belief. Analogously, the failure of the global ambiguity neutrality property will necessarily lead to a paradoxical mode of behavior. In principle, such behavior may involve complicated acts on a large state space. I shall next show that we can, in fact, focus on a classical paradox in a simple environment to check the global ambiguity neutrality property.

4.1 An Experimental Characterization

Consider a modified version of the experiment in Section 3.2 with three colors: blue (b), orange (o), and white (w). Both urns contain m balls. The composition of urn 2 is unknown, while urn 1 contains exactly m^k balls of color k, where $m^b + m^o + m^w = m$. Set $\gamma^k := m^k/m$ for $k = b, o, w$. Furthermore, for $i = 1, 2$ and $k = b, o, w$, let f^{ki} denote the bet that pays x^* if the color of the ball extracted from urn i is k, and that pays x_* otherwise.

Given a certainty equivalence function c on $\Delta(X)$, by a *recursive subject of type* c I mean a subject who evaluates the bets described above according to a recursive representation (μ, c), in line with the available information. Specifically, the subject takes $S := \{b, o, w\}$ as the state space, her preference relation $\succsim$ over $\Delta(\mathcal{H})$ admits a recursive representation (μ, c), and she identifies the bets as follows:

$$f^{k1} := \gamma^k \delta_{x^*} + \left(1 - \gamma^k\right) \delta_{x_*} \quad \text{for } k = b, o, w,$$

$$f^{k2}(s) := \begin{cases} \delta_{x^*} & \text{if } s = k \\ \delta_{x_*} & \text{if } s \in S \backslash \{k\} \end{cases} \quad \text{for } k = b, o, w.$$

I say that the subject exhibits the *Ellsberg paradox* if

$$f^{k1} \succ f^{k2} \quad \text{for } k = b, o, w.$$

In turn, the subject exhibits the *anti-Ellsberg paradox* if

$$f^{k1} \prec f^{k2} \quad \text{for } k = b, o, w.$$

Observe that given a $\pi \in \Delta(S)$, we have $\pi_{f^{k2}} = \pi(k)\delta_{x^*} + (1 - \pi(k))\delta_{x_*}$. Thus, $\pi_{f^{k2}} \succsim f^{k1}$

if and only if $\pi(k) \geq \gamma^k$. Since $\pi(b) + \pi(o) + \pi(w) = 1 = \gamma^b + \gamma^o + \gamma^w$, a subject with a first-order belief π cannot exhibit either paradox. In particular, if the subject's preferences over $\Delta(X)$ satisfies the independence axiom, she may not exhibit either paradox. The next result establishes the converse of this statement, in a global sense.

Theorem 2. *Let $\succsim_c$ be a preference relation on $\Delta(X)$ represented by a certainty equivalence function c. The following two statements are equivalent.*

(i) *$\succsim_c$ satisfies the independence axiom.*

(ii) *A recursive subject of type c exhibits neither Ellsberg nor anti-Ellsberg paradox for any $\mu \in \Delta^2(S)$ and for any specification of the parameters m, m^b, m^o and m^w in the experiment above.*

Since Segal (1987), it is well-known that the independence axiom over $\Delta(X)$ entails ambiguity neutrality within the class of recursive preferences. By establishing the converse, Corollary 2 and Theorem 2 provide full characterizations of the interplay between the independence axiom and ambiguity neutrality. The message of Theorem 2 is much stronger: The failure of the independence axiom (or the global ambiguity neutrality property) will lead to a classical paradox with only three states and binary acts, for at least one second-order belief.[14]

5 Increasing Ambiguity Aversion

This section studies how we can manipulate a given second-order belief to increase the strength of ambiguity aversion, irrespective of the details of risk preferences. As a minimal requirement, I will focus on risk preferences that satisfy NCI because selecting a larger class is likely to trivialize the problem.

Consider a fixed state space S, and a pair of second-order beliefs, $\mu, \mu' \in \Delta^2(S)$.

Definition 6. μ is a *mean-preserving spread of* μ' if there exists an $\alpha \in [0, 1]$ such that

$$\mu' = \alpha\mu + (1 - \alpha)D_{\bar{\mu}}. \tag{11}$$

Observe that the mean of $D_{\bar{\mu}}$ is simply $\bar{\mu}$. Since the expectation operator is linear, equation (11) implies $\bar{\mu}' = \alpha\bar{\mu} + (1 - \alpha)\bar{\mu} = \bar{\mu}$. That is, μ' and μ have the same mean. Moreover, $\mu'(\pi) = \alpha\mu(\pi) \leq \mu(\pi)$ for any first-order distribution π that is distinct from $\bar{\mu}$. Thus, μ can be obtained from μ' by transferring some mass from $\bar{\mu}$ to other first-order

[14] On a related note, the recursive CEU model cannot be characterized in a way that is directly analogous to Theorem 2. Indeed, by Segal's (1987) Theorem 4.2, a subclass of the recursive RDU model also predicts ambiguity aversion consistently, insofar as the *binary* acts are concerned.

distributions. In this sense, μ embodies a larger amount of second-order uncertainty than μ' does. As a peculiar feature, equation (11) also implies that μ and μ' induce the same relative likelihood for any pair of first-order distributions that are distinct from $\bar{\mu}$.

To see why Definition 6 entails "increasing ambiguity aversion," fix a certainty equivalence function c. Let $\succsim$ and $\succsim'$ denote the preference relations on $\Delta(\mathcal{H})$ represented by (μ, c) and (μ', c), respectively. Along the lines of Lemma 1, it can be shown that $\succsim$ is more ambiguity averse than $\succsim'$ if and only if

$$\mu'_f \succsim \mu_f \quad \forall f \in \mathcal{H}_X.$$

Fix an $f \in \mathcal{H}_X$, and note that equation (11) implies

$$\begin{aligned} \mu'_f := \sum_{\pi \in \Delta(S)} \mu'(\pi)\delta_{c(\pi_f)} &= \sum_{\pi \in \Delta(S)} \left(\alpha\mu(\pi) + (1-\alpha)D_{\bar{\mu}}(\pi)\right)\delta_{c(\pi_f)} \\ &= \alpha \sum_{\pi \in \Delta(S)} \mu(\pi)\delta_{c(\pi_f)} + (1-\alpha) \sum_{\pi \in \Delta(S)} D_{\bar{\mu}}(\pi)\delta_{c(\pi_f)} \\ &= \alpha\mu_f + (1-\alpha)\delta_{c(\bar{\mu}_f)}. \end{aligned}$$

If c satisfies NCI, then, $\bar{\mu}_f \succsim \mu_f$ as we have seen earlier. Equivalently, $\delta_{c(\bar{\mu}_f)} \succsim \mu_f$. Furthermore, as noted by Dillenberger (2010, Lemma 2), NCI implies convexity, meaning that $\alpha q + (1-\alpha)p \succsim q$ whenever $p \succsim q$. It follows that $\mu'_f = \alpha\mu_f + (1-\alpha)\delta_{c(\bar{\mu}_f)} \succsim \mu_f$. So, $\succsim$ is more ambiguity averse than $\succsim'$, as we sought.

The converse implication requires a further assumption. I say that μ' is *regular* if

$$\mu'(\delta_s) = 0 \quad \forall s \in S. \tag{12}$$

Geometrically, this amounts to saying that the vertices of the simplex $\Delta(S)$ have zero probability, which is a rather mild assumption. I will clarify the role of this assumption momentarily.

In what follows, the properties of a certainty equivalence function c (such as NCI, risk aversion or loving) refer to the corresponding properties of the risk preference $\succsim_c$.

Theorem 3. *Let $\mu, \mu' \in \Delta^2(S)$, and suppose μ' is regular. Then, the following two statements are equivalent.*

(i) *μ is a mean-preserving spread of μ'.*

(ii) *The recursive preference represented by (μ, c) is more ambiguity averse than that represented by (μ', c) for every certainty equivalence function c that satisfies* NCI.

Moreover, the same conclusion obtains if statement (ii) *is restricted to risk averse* (*resp. loving*) *certainty equivalence functions that satisfy* NCI.

Theorem 3 shows that, under the regularity assumption (12), taking a mean-preserving spread of a second-order belief is equivalent to increasing ambiguity aversion embodied in that belief for every risk-preference that satisfies NCI.

As shown by Cerreia-Vioglio et al. (2015, Theorem 3), risk averse preference relations over lotteries that satisfy NCI correspond to CEU representations with a set of utility indices, W, that consists of concave functions on X. Similarly, risk loving CEU preferences are characterized by convex utility indices.[15] The final statement in Theorem 3 means that both of these subclasses lead to the same characterization of increasing ambiguity aversion as the class of all CEU preferences.

In the proof of Theorem 3, to show that (ii) implies (i), I consider several CEU representations, each giving a particular certainty equivalence function. That μ and μ' must have the same mean follows from expected utility functionals. The remaining links between μ and μ' are established by further functional forms, carefully selected in relation to the distributions in the supports of μ and μ'.

The regularity condition (12) is indispensable in the statement of Theorem 3 because, as a consequence of the time-neutrality property of recursive preferences, the degenerate belief $D_{\bar{\mu}}$ cannot be distinguished behaviorally from a dual belief $\bar{\bar{\mu}}$ that attaches the probability $\bar{\mu}(s)$ to the degenerate first-order distribution δ_s, for every $s \in S$.[16] For risk preferences that satisfy NCI, it can be shown that a convex combination of $D_{\bar{\mu}}$ and $\bar{\bar{\mu}}$ is also indistinguishable from these two beliefs.

One way to deal with this issue may be to replace equation (11) with an expression of the form

$$\mu' \in \alpha\mu + (1-\alpha)[D_{\bar{\mu}}],$$

where $[D_{\bar{\mu}}]$ stands for the set of all second-order beliefs that are behaviorally equivalent to $D_{\bar{\mu}}$. However, the present approach based on equation (11) has several advantages. First, it seems to be easier to interpret (11) as a definition of a mean-preserving spread. Indeed, aside from the differences in primitives, this definition is a special version of the classical mean-preserving spread operation over monetary lotteries. A direct analogue of the classical definition would qualify $\alpha\hat{\mu} + (1-\alpha)\mu$ as a mean-preserving spread of $\alpha\hat{\mu} + (1-\alpha)D_{\bar{\mu}}$ for every $\hat{\mu}, \mu \in \Delta^2(S)$.[17] Since $\mu = \alpha\mu + (1-\alpha)\mu$, upon letting $\hat{\mu} = \mu$, we see that Definition 6

[15] Risk aversion (resp. loving) refers to a negative (resp. positive) attitude towards the classical mean-preserving spread operation over monetary lotteries.

[16] If we think of a second-order belief as a dynamic stochastic process with two stages, time-neutrality entails that as long as only one stage involves uncertainty, it does not matter whether that stage is the first one or the second. (See Appendix A for a formal statement of the time-neutrality property.)

[17] Ergin and Gul (2009) apply this classical definition, formulated in a slightly different way, to compound lotteries, in order to characterize second-order risk aversion in their model of second-order probabilistic sophistication.

is a particular case of this classical formula.

Moreover, the present approach can also be useful in alternative models which do not possess the time-neutrality property, such as the smooth ambiguity model of Klibanoff, Marinacci and Mukerji (2005), and the second-order subjective expected utility model of Seo (2009). In fact, in these models, even the classical formula described above can be utilized to increase the strength of ambiguity aversion. Thus, it can be shown that in a certain sense, equation (11) describes the most general mean-preserving spread operation for second-order beliefs that functions well *both* in Segal's theory (assuming NCI) and the aforementioned theories based on a second-order expected utility operator.

6 Concluding Remarks

In this paper, I have studied Segal's (1987) theory of recursive preferences with a focus on the relations between (i) risk and ambiguity attitudes; (ii) the structure of second-order beliefs and the strength of ambiguity aversion.

While second-order beliefs is a natural starting point, Segal's theory can be extended in a straightforward way to include higher order beliefs, say, in $\Delta^n(S) := \Delta(\Delta^{n-1}(S))$ for $n = 3, 4, ...$ It is also a simple exercise to verify that within the recursive CEU model, the strength of ambiguity aversion increases with the degree of the belief, n, holding constant the expectations with lower degrees. In fact, following Dillenberger (2010, Proposition 6), it may be possible to model even extreme forms of ambiguity aversion by increasing the degree of beliefs arbitrarily. Thus, in potential applications, it may also be worthwhile to consider higher order beliefs.

As another venue for future research, an axiomatic description of Segal's model in the Anscombe-Aumann setup is not yet available. Specifically, it is an open problem to provide a behavioral description of a DM who converts acts into compound lotteries with the help of a second-order belief. To this end, the main difficulty is that the Anscombe-Aumann setup does not accommodate second-order acts that assign prizes to first-order distributions, as opposed to states. In particular, absent second-order acts, one cannot formulate direct analogues of Machina and Schmeidler's (1992) axioms that characterize first-order probabilistic sophistication. Another notable approach to this problem is due to Ergin and Gul (2009), who consider a richer state space which effectively equips the model with acts that are comparable to the second-order acts that I just described. A more detailed discussion of the axiomatic features of Segal's model can be found in Appendix A below.

Appendix
A. More on Behavioral Properties of Recursive Preferences

The following property formalizes the idea that the DM converts acts into compound lotteries with the help of a second-order belief.

A0: Second-Order Probabilistic Sophistication. There exists a $\mu \in \Delta^2(S)$ such that $f \sim \sum_{\pi \in \Delta(S)} \mu(\pi) D_{\pi_f}$ for every $f \in \mathcal{H}_X$.

As I noted earlier, a behavioral description of (A0) in the Anscombe-Aumann setup is not yet available. Apart from Segal's (1987) recursive preferences, this property also underlies the second-order subjective expected utility model of Seo (2009).

Along the lines of Segal (1990), it can easily be shown that the recursive representation in Definition 2 is characterized by the following six properties, in addition to (A0).

A1: Time-Neutrality. $p \sim \sum_{x \in X} p(x) D_{\delta_x}$ for any $p \in \Delta(X)$.

A2: Monotonicity. For any $f, g \in \mathcal{H}$, if $f(s) \succsim g(s)$ for every $s \in S$, then $f \succsim g$.

A3: Recursivity. For any $p, q \in \Delta(X)$, $P \in \Delta^2(X)$, and $\alpha \in (0, 1]$,

$$p \succsim q \quad \Leftrightarrow \quad \alpha D_p + (1-\alpha)P \succsim \alpha D_q + (1-\alpha)P.$$

A4: More Is Better. For any $x, y \in X$, $\delta_x \succsim \delta_y$ if and only if $x \geq y$.

A5: Continuity. $\{F : F \succsim G\}$ and $\{F : F \precsim G\}$ are closed subsets of $\Delta(\mathcal{H})$ for any $G \in \Delta(\mathcal{H})$, where $\Delta(\mathcal{H})$ is endowed with the topology of weak convergence associated with the product topology on $\mathcal{H}$.

A6: Weak-Order. $\succsim$ is transitive and complete.

Time-neutrality asserts that a compound lottery that is degenerate in the second stage is equivalent to the reduced form of that lottery. When the two stages involved in a compound lottery represent different points in time, this property entails indifference towards the timing of the resolution of uncertainty, which is a key difference between the models of Segal (1987) and Seo (2009).

As noted in Section 2.1, the most important implication of the monotonicity axiom is that $f \sim c \circ f$ for every $f \in \mathcal{H}$. So, this axiom allows us to focus on purely subjective acts in place of the lottery-valued acts.

Finally, recursivity requires preferences over compound lotteries to be consistent with preferences over simple lotteries. Alternatively, when the elements of $\Delta(X)$ are considered as the prizes associated with compound lotteries, this property can also be seen as a monotonicity condition with respect to the first order stochastic dominance relation over compound lotteries. Segal (1990) utilizes a slightly different form of this axiom, called "compound independence."

B. Ambiguity Aversion Does Not Imply Mean Ambiguity Aversion

In this appendix, I construct an example of a recursive RDU preference with the properties described in Section 3.2.

Given any $\varepsilon \in (0, 1/6)$, define a function Ψ as

$$\Psi(\lambda) := \begin{cases} 0 & \text{for } 0 \leq \lambda \leq \frac{1}{2} - \varepsilon, \\ \frac{2}{1+2\varepsilon}\left(\lambda - (\frac{1}{2} - \varepsilon)\right) & \text{for } \frac{1}{2} - \varepsilon < \lambda \leq 1. \end{cases}$$

Ψ is a piecewise linear, convex and weakly increasing function that maps $[0, 1]$ onto $[0, 1]$. (We will shortly see that strictly increasing and strictly convex functions that are close to Ψ also possess the properties that we seek.)

Pick any $v : X \to \mathbb{R}$ that is strictly increasing. Let u denote the RDU functional on $\Delta(X)$ defined by v and Ψ, and set $c(p) := v^{-1}(u(p))$ for $p \in \Delta(X)$.

As in Section 3.2, $S := \{b, o\}$ and $\mu := \frac{1}{2}D_{\pi^1} + \frac{1}{2}D_{\pi^2}$, where π^1 and π^2 are the elements of $\Delta(S)$ with $\pi^1(b) := \frac{1}{2} + 3\varepsilon$ and $\pi^2(b) := \frac{1}{2} - \varepsilon$. In turn, $\succsim$ stands for the recursive preference on $\Delta(\mathcal{H})$ represented by (μ, c).

For $s \in S$, set $\mathcal{H}^s := \{f \in \mathcal{H}_X : f(s) \succ f(s')\}$, where s' denotes the element of S that is distinct from s. Given an $f \in \mathcal{H}^s$, x and x' stand for the prizes returned by f in states s and s', respectively. That is, $f(s) = \delta_x$ and $f(s') = \delta_{x'}$.

In what follows, I focus on acts in the set $\mathcal{H}^b \cup \mathcal{H}^o$. Indeed, for a purely subjective act f that does not belong to $\mathcal{H}^b \cup \mathcal{H}^o$, we have $f(b) = f(o)$, which implies $\pi_f = \mu_f$ for every $\pi \in \Delta(S)$.

Observe that

$$u(\pi_f) = v(x') + (v(x) - v(x'))\,\Psi(\pi(s)) \quad \forall s \in S,\ f \in \mathcal{H}^s \text{ and } \pi \in \Delta(S). \tag{13}$$

It follows that $u(\pi^1_f) > u(\pi^2_f)$ for every $f \in \mathcal{H}^b$ because $\Psi(\pi^1(b)) = \Psi(\frac{1}{2} + 3\varepsilon) > 0 = \Psi(\frac{1}{2} - \varepsilon) = \Psi(\pi^2(b))$. Hence,

$$u(\mu_f) = u(\pi^2_f) + \left(u(\pi^1_f) - u(\pi^2_f)\right)\Psi(\tfrac{1}{2}) \quad \forall f \in \mathcal{H}^b. \tag{14}$$

Given the values of $u(\pi^i_f)$ from equation (13), after some algebra, equation (14) reduces to

$$u(\mu_f) = v(x') + (v(x) - v(x'))\left(\Psi(\pi^2(b)) + (\Psi(\pi^1(b)) - \Psi(\pi^2(b)))\,\Psi(\tfrac{1}{2})\right) \quad \forall f \in \mathcal{H}^b.$$

If we compare this equation with (13) and invoke the definitions of $\pi^i(b)$, we get the following characterization for any $f \in \mathcal{H}^b$ and $\pi \in \Delta(S)$:

$$\pi_f \succsim \mu_f \quad \Leftrightarrow \quad \Psi(\pi(b)) \geq \Psi(\tfrac{1}{2} - \varepsilon) + \big(\Psi(\tfrac{1}{2} + 3\varepsilon) - \Psi(\tfrac{1}{2} - \varepsilon)\big)\, \Psi(\tfrac{1}{2}). \tag{B}$$

Similarly, for any $f \in \mathcal{H}^o$ and $\pi \in \Delta(S)$:

$$\pi_f \succsim \mu_f \quad \Leftrightarrow \quad \Psi(\pi(o)) \geq \Psi(\tfrac{1}{2} - 3\varepsilon) + \big(\Psi(\tfrac{1}{2} + \varepsilon) - \Psi(\tfrac{1}{2} - 3\varepsilon)\big)\, \Psi(\tfrac{1}{2}) \tag{O}$$

It is easily checked that for the uniform distribution $\bar{\pi}$, with $\bar{\pi}(b) = \bar{\pi}(o) = 1/2$, the inequality on the right hand side of (B) holds strictly. Indeed, $\Psi(\frac{1}{2} - \varepsilon) = 0 < \Psi(\frac{1}{2})$, while $\Psi(\frac{1}{2} + 3\varepsilon) < 1$. By the same logic, the inequality on the right hand side of (O) also holds strictly for the distribution $\bar{\pi}$. Thus, $\bar{\pi}_f \succ \mu_f$ for every $f \in \mathcal{H}^b \cup \mathcal{H}^a$, which implies that $\succsim$ is ambiguity averse.

Moreover, $\Psi(\bar{\mu}(o)) = 0 < \Psi(\frac{1}{2}+\varepsilon)\Psi(\frac{1}{2})$ because $\bar{\mu}(o) = \frac{1}{2}\pi^1(o) + \frac{1}{2}\pi^2(o) = \frac{1}{2} - \varepsilon$. Hence, with $\pi = \bar{\mu}$, from (O) it follows that $\bar{\mu}_f \prec \mu_f$ for every $f \in \mathcal{H}^o$. So, $\succsim$ is not mean ambiguity averse, as we sought.

An examination of the arguments above reveals that the following three features of Ψ are of key importance: (i) $\Psi(\frac{1}{2} - \varepsilon) = 0$; (ii) $\Psi(\frac{1}{2}) > 0$; (iii) $\Psi(\frac{1}{2} + 3\varepsilon) < 1$.[18] While the latter two conditions will be satisfied by any strictly increasing function, an approximate version of condition (i) would work equally well as this condition. Specifically, the arguments above remain valid for any strictly increasing function Ψ with $\Psi\left(\frac{1}{2} - \varepsilon\right) \approx 0$ and $\Psi\left(\frac{1}{2}\right)/\Psi\left(\frac{1}{2} - \varepsilon\right) \approx \infty$. In particular, any neighborhood (with respect to the sup-norm) of the original function that I constructed also contains strictly convex and strictly increasing functions that possess all the properties that we seek.

C. Proofs

Recall that, given a recursive preference $\succsim$ represented by (μ, c), we have $f \sim \mu_f$ and $p \sim \delta_{c(p)}$ for every $f \in \mathcal{H}_X$ and $p \in \Delta(X)$. Since $\succsim$ is transitive, it also follows that $f \sim \delta_{c(\mu_f)}$ for every $f \in \mathcal{H}_X$. I frequently utilize these observations throughout the proofs.

Proof of Lemma 1. Let $\succsim$ be a recursive preference on $\Delta(\mathcal{H})$ represented by (μ, c). Assume first that $\succsim$ is ambiguity averse. Then, there exists an ambiguity neutral preference relation $\succsim'$ on $\Delta(\mathcal{H})$ s.t. $\succsim$ and $\succsim'$ satisfy properties (i) and (ii) in Definition 3. Moreover, by ambiguity neutrality of $\succsim'$, there exists a $\bar{\pi} \in \Delta(S)$ such that $f \sim' \bar{\pi}_f$ for every $f \in \mathcal{H}_X$. Pick any $f \in \mathcal{H}_X$. Then, by property (i), $f \sim \delta_{c(\mu_f)}$ implies $f \succsim' \delta_{c(\mu_f)}$, which also means $\bar{\pi}_f \succsim' \delta_{c(\mu_f)}$. From property (ii), it then follows that $\bar{\pi}_f \succsim \delta_{c(\mu_f)}$, i.e., $\bar{\pi}_f \succsim \mu_f$.

[18] Also note that the elasticity of Ψ is not non-decreasing because $\Psi(\frac{1}{2})^2 > \Psi(\frac{1}{4})$.

For the converse implication, suppose now that there exists a $\bar{\pi} \in \Delta(S)$ such that $\bar{\pi}_f \succsim \mu_f$ for every $f \in \mathcal{H}_X$. Let $\succsim'$ denote the recursive preference on $\Delta(\mathcal{H})$ represented by (μ', c), where $\mu' := D_{\bar{\pi}}$. Then, for any $p, q \in \Delta(X)$, we have $p \succsim' q$ iff $c(p) \geq c(q)$ iff $p \succsim q$, which verifies property (ii) in Definition 3. Moreover, $\mu'_f := \delta_{c(\bar{\pi}_f)}$, and hence, $f \sim' \delta_{c(\bar{\pi}_f)} \sim' \bar{\pi}_f$ for every $f \in \mathcal{H}_X$. This shows that $\succsim'$ is ambiguity neutral. It remains to verify property (i) in Definition 3. Fix any $f \in \mathcal{H}_X$, and note that because $\bar{\pi}_f \succsim \mu_f \sim f$, we have $\bar{\pi}_f \succsim \delta_x$ for any $x \in X$ with $f \succsim \delta_x$. In turn, $\bar{\pi}_f \succsim \delta_x$ means $\bar{\pi}_f \succsim' \delta_x$, while $f \sim' \bar{\pi}_f$ as noted earlier. Thus, $f \succsim' \delta_x$ for any $x \in X$ with $f \succsim \delta_x$, as we sought. $\square$

In what follows, given any $S' \subseteq S$ and $x, y \in X$, I denote by $x_{S'}y$ the act f such that $f(s) = \delta_x$ for $s \in S'$ and $f(s) = \delta_y$ for $s \in S \backslash S'$.

Proof of Theorem 1. By Proposition 1 of Dillenberger (2010), it suffices to show that (i) is equivalent to (iii). In turn, that (i) implies (iii) has been noted by Artstein-Avidan and Dillenberger (2011), without proof. For the sake of completeness, I start with a proof of this fact.

Fix a state space S, and let $\succsim$ be a recursive preference on $\Delta(\mathcal{H})$ represented by (μ, c) for some $\mu \in \Delta^2(S)$ and a certainty equivalence function c. By definitions, for any $f \in \mathcal{H}_X$ we have

$$
\begin{aligned}
\bar{\mu}_f = \sum_{s \in S} \bar{\mu}(s) f(s) = \sum_{s \in S} \left(\sum_{\pi \in \Delta(S)} \mu(\pi)\pi(s) \right) f(s) &= \sum_{\pi \in \Delta(S)} \mu(\pi) \left(\sum_{s \in S} \pi(s) f(s) \right) \\
&= \sum_{\pi \in \Delta(S)} \mu(\pi)\pi_f.
\end{aligned} \tag{15}
$$

Moreover, if $\succsim_c$ satisfies NCI, applying this axiom successively for every π in the support of μ yields

$$\sum_{\pi \in \Delta(S)} \mu(\pi)\pi_f \succsim_c \sum_{\pi \in \Delta(S)} \mu(\pi)\delta_{c(\pi_f)}$$

because $\pi_f \succsim_c \delta_{c(\pi_f)}$ for every $\pi \in \Delta(S)$. Since $\mu_f := \sum_{\pi \in \Delta(S)} \mu(\pi)\delta_{c(\pi_f)}$ and $\succsim_c = \succsim$ over $\Delta(X)$, it follows that $\bar{\mu}_f \succsim \mu_f$. Thus, if $\succsim_c$ satisfies NCI, it exhibits a special form of the global ambiguity aversion property: For any S and $\mu \in \Delta^2(S)$, the recursive preference represented by (μ, c) is mean ambiguity averse.

To prove that (iii) implies (i), let $\succsim_c$ be a preference relation on $\Delta(X)$ represented by a certainty equivalence function c, and suppose that $\succsim_c$ has the global ambiguity aversion property. Then, by Lemma 1, for any state space S and any $\mu \in \Delta^2(S)$ the recursive preference $\succsim$ on $\Delta(\mathcal{H})$ represented by (μ, c) admits a $\bar{\pi} \in \Delta(S)$ such that $\bar{\pi}_f \succsim \mu_f$ for every

$f \in \mathcal{H}_X$.

Consider a state space S that consists of $n \times m$ points, where n and m are arbitrarily selected natural numbers. Partition S into n sets, $S^1, ..., S^n$, each having m points. For $i = 1, ..., n$, let $\pi^i \in \Delta(S)$ denote the uniform distribution over S^i, so that $\pi^i(s) = 1/m$ for $s \in S^i$ and $\pi^i(s) = 0$ for $s \in S \backslash S^i$. Pick any $\{\alpha^1, ..., \alpha^n\} \subseteq [0,1]$ with $\sum_{i=1}^n \alpha^i = 1$ and set $\mu := \sum_{i=1}^n \alpha^i D_{\pi^i}$. I claim that $\bar{\pi}$, delivered by Lemma 1, is equal to $\bar{\mu} := \sum_{i=1}^n \alpha^i \pi^i$. That is, $\bar{\pi}(s) = \alpha^i/m$ for $s \in S^i$ and $i = 1, ..., n$.

The first step is to show that

$$\bar{\pi}(S^i) = \alpha^i \quad \text{for } i = 1, ..., n. \tag{16}$$

Since $\sum_{i=1}^n \alpha^i = 1 = \bar{\pi}(S) = \sum_{i=1}^n \bar{\pi}(S^i)$, equation (16) holds iff $\bar{\pi}(S^i) \geq \alpha^i$ for every $i \in \{1, ..., n\}$. Pick any $i \in \{1, ..., n\}$ and $x, y \in X$ with $x > y$. Set $f := x_{S^i} y$ and observe that $\bar{\pi}_f = \bar{\pi}(S^i)\delta_x + (1 - \bar{\pi}(S^i))\delta_y$. Moreover, $\pi^i_f = \delta_x$ while $\pi^j_f = \delta_y$ for every $j \in \{1, ..., n\}$ with $j \neq i$. Thus, $\mu_f = \sum_{j=1}^n \mu(\pi^j)\delta_{c(\pi^j_f)} = \sum_{j=1}^n \alpha^j \delta_{c(\pi^j_f)} = \alpha^i \delta_x + (1 - \alpha^i)\delta_y$. Hence, $\bar{\pi}_f \succsim \mu_f$ means $\bar{\pi}(S^i)\delta_x + (1 - \bar{\pi}(S^i))\delta_y \succsim \alpha^i \delta_x + (1 - \alpha^i)\delta_y$. It follows that $\bar{\pi}(S^i) \geq \alpha^i$ because $\succsim$ is monotonic w.r.t. $\geq_{\text{fosd}}$. This proves equation (16).

For every $i \in \{1, ..., n\}$, pick an arbitrary point $s^i \in S^i$, and set $S' := \{s^1, ..., s^n\}$. Let us now show that

$$\bar{\pi}(S') = 1/m. \tag{17}$$

Put $f := x_{S'} y$ for some $x, y \in X$ with $x > y$. By construction, $S' \cap S^i = \{s^i\}$ for $i = 1, ..., n$. Hence, by definition of π^i, we have $\pi^i(S') = \pi^i(s^i) = \frac{1}{m}$, implying that

$$\pi^i_f = \pi^i(S')\delta_x + (1 - \pi^i(S'))\,\delta_y = \tfrac{1}{m}\delta_x + \left(1 - \tfrac{1}{m}\right)\delta_y \quad \text{for } i = 1, ..., n.$$

Thus,

$$\mu_f = \sum_{i=1}^n \mu(\pi^i)\delta_{c\left(\pi^i_f\right)} = \sum_{i=1}^n \mu(\pi^i)\delta_{c\left(\frac{1}{m}\delta_x + \left(1 - \frac{1}{m}\right)\delta_y\right)} = \delta_{c\left(\frac{1}{m}\delta_x + \left(1 - \frac{1}{m}\right)\delta_y\right)}.$$

In particular, $\mu_f \sim \frac{1}{m}\delta_x + \left(1 - \frac{1}{m}\right)\delta_y$. On the other hand, $\bar{\pi}_f = \bar{\pi}(S')\delta_x + (1 - \bar{\pi}(S'))\delta_y$. Hence, as in the proof of the previous step, $\bar{\pi}_f \succsim \mu_f$ implies $\bar{\pi}(S') \geq \frac{1}{m}$. Similarly, we get $\bar{\pi}(S') \leq \frac{1}{m}$ upon letting $f' := y_{S'} x$ and invoking the condition $\bar{\pi}_{f'} \succsim \mu_{f'}$. This completes the proof of (17).

Note that S' in equation (17) is an arbitrary set that contains exactly one element of every S^i. Fix an $i \in \{1, ..., n\}$ and a point $\bar{s}^j \in S^j$ for every $j \in \{1, ..., n\} \backslash \{i\}$. Put $\bar{S}_{-i} := \bigcup_{j \neq i} \{\bar{s}^j\}$ and $S'_i(s) := \{s\} \cup \bar{S}_{-i}$ for any $s \in S^i$. Then, equation (17) and additivity of the probability measure $\bar{\pi}$ imply

$$\bar{\pi}(s)+\bar{\pi}\left(\bar{S}_{-i}\right)=\bar{\pi}\left(S_i'(s)\right)=\tfrac{1}{m}=\bar{\pi}\left(S_i'(\hat{s})\right)=\bar{\pi}\left(\hat{s}\right)+\bar{\pi}\left(\bar{S}_{-i}\right)\quad \forall s,\hat{s}\in S^i.$$

It follows that $\bar{\pi}(s)=\bar{\pi}\left(\hat{s}\right)$ for every $s,\hat{s}\in S^i$. In view of equation (16), we must then have $\bar{\pi}(s)=\alpha^i/m$ for every $s\in S^i$ because $\bar{\pi}$ is additive and S^i contains m points. That is, $\bar{\pi}=\bar{\mu}$, as we sought.

Let $\Delta_0(X):=\{p\in\Delta(X):p(x)\text{ is a rational number }\forall x\in X\}$. Pick any $p\in\Delta_0(X)$, $q\in\Delta(X)$ and $\alpha\in[0,1]$. Suppose that the support of p consists of K distinct points, $x^1,...,x^K$, whereas the support of q consists of I distinct points, $y^1,...,y^I$. As p belongs to $\Delta_0(X)$, for every $k\in\{1,...,K\}$, there exist natural numbers a^k,b^k such that $p(x^k)=\frac{a^k}{b^k}$. Set $m:=\prod_{k=1}^{K}b^k$ and $m^k:=m\frac{a^k}{b^k}$ for $k=1,...,K$. Let $n:=I+1$ and consider a state space S that consists of $n\times m$ points. Finally, define $S^1,...,S^n$, $\pi^1,...,\pi^n$ just as before, and let $\mu:=\sum_{i=1}^n\alpha^iD_{\pi^i}$, where $\alpha^i:=(1-\alpha)q(y^i)$ for $i=1,...,I$ and $\alpha^n:=\alpha$.

Observe that m^k is a natural number for every k and that $\sum_{k=1}^K m^k=m\sum_{k=1}^K p(x^k)=m$. Moreover, by construction, S^n contains m points. Thus, this set can be partitioned into K subsets, $S^{n1},...,S^{nK}$, such that the cardinality of S^{nk} equals m^k for every $k=1,...,K$.

Define a purely subjective act f as

$$f(s):=\begin{cases}\delta_{y^i} & \text{if } s\in S^i \text{ for some } i=1,...,I,\\ \delta_{x^k} & \text{if } s\in S^{nk} \text{ for some } k=1,...,K.\end{cases}$$

As we have seen earlier, $\bar{\mu}_f\succsim\mu_f$ because $\succsim_c$ has the global ambiguity aversion property. Observe that

$$\begin{aligned}\bar{\mu}_f&=\sum_{i=1}^I\bar{\mu}(S^i)\delta_{y^i}+\sum_{k=1}^K\bar{\mu}(S^{nk})\delta_{x^k}\\ &=\sum_{i=1}^I\alpha^i\delta_{y^i}+\sum_{k=1}^K\alpha^n\frac{m^k}{m}\delta_{x^k}\\ &=\sum_{i=1}^I(1-\alpha)q(y^i)\delta_{y^i}+\sum_{k=1}^K\alpha p(x^k)\delta_{x^k}=(1-\alpha)q+\alpha p.\end{aligned}$$

Moreover, $\pi_f^i=\delta_{y^i}$ for $i=1,...,I$, while

$$\pi_f^n=\sum_{k=1}^K\pi^n(S^{nk})\delta_{x^k}=\sum_{k=1}^K\frac{m^k}{m}\delta_{x^k}=\sum_{k=1}^K p(x^k)\delta_{x^k}=p.$$

Thus,

$$\mu_f = \sum_{i=1}^{I} \mu(\pi^i)\delta_{c(\pi^i_f)} + \mu(\pi^n)\delta_{c(\pi^n_f)} = \sum_{i=1}^{I} \alpha^i \delta_{y^i} + \alpha^n \delta_{c(p)} = (1-\alpha)q + \alpha\delta_{c(p)}.$$

Hence, $\bar{\mu}_f \succsim \mu_f$ means $(1-\alpha)q + \alpha p \succsim (1-\alpha)q + \alpha\delta_{c(p)}$. Clearly, in the latter expression we can replace p with an arbitrary element of $\Delta(X)$ because c is a continuous function on $\Delta(X)$, and $\Delta_0(X)$ is a dense subset of $\Delta(X)$. Thus, $(1-\alpha)q + \alpha p \succsim (1-\alpha)q + \alpha\delta_{c(p)}$ for every $p, q \in \Delta(X)$ and $\alpha \in [0,1]$. This is equivalent to saying that $\succsim_c$ satisfies NCI because $\succsim = \succsim_c$ over $\Delta(X)$ and $\succsim_c$ is monotonic w.r.t. to $\geq_{\text{fosd}}$. $\square$

Proof of Lemma 2. By Definition 3, every ambiguity neutral preference is trivially ambiguity averse and loving. To establish the converse, let $\succsim$ be a recursive preference on $\Delta(\mathcal{H})$ represented by (μ, c), and suppose that $\succsim$ is both ambiguity averse and loving. Then, Lemma 1 and an obvious, dual property imply that there exist a pair of distributions $\bar{\pi}, \pi' \in \Delta(S)$ such that

$$\bar{\pi}_f \succsim \mu_f \succsim \pi'_f \quad \forall f \in \mathcal{H}_X. \tag{18}$$

I claim that $\bar{\pi} = \pi'$. Indeed, if $\bar{\pi}$ and π' were distinct, there would exist a set $S' \subseteq S$ such that $\pi'(S') > \bar{\pi}(S')$. Let $f := x_{S'}y$ for some $x, y \in X$ with $x > y$, so that $\pi'_f = \pi'(S')\delta_x + (1-\pi'(S'))\delta_y$, while $\bar{\pi}_f = \bar{\pi}(S')\delta_x + (1-\bar{\pi}(S'))\delta_y$. As $\succsim$ is monotonic w.r.t. $\geq_{\text{fosd}}$, $\pi'(S') > \bar{\pi}(S')$ implies $\pi'_f \succ \bar{\pi}_f$, which contradicts (18). So, $\bar{\pi} = \pi'$, and hence, property (18) yields $\bar{\pi}_f \sim \mu_f$ for every $f \in \mathcal{H}_X$. Since $f \sim \mu_f$, this means that $\succsim$ is ambiguity neutral. $\square$

I omit the proofs of Corollary 1 and 2. (The main arguments regarding the proof of Corollary 2 can be found in Section 4.)

Proof of Theorem 2. As noted in Section 4.1, if $\succsim_c$ satisfies the independence axiom, a recursive subject of type c cannot exhibit either paradox. To establish the converse, suppose that a recursive subject of type c does not exhibit either paradox for any $\mu \in \Delta^2(S)$ and any specification of m and m^k for $k = b, o, w$. By varying the latter parameters, in urn 1 we can obtain any distribution of the fractions $\gamma^b, \gamma^o, \gamma^w$, subject to the requirement that γ^k is a rational number for every k. Let $\succsim$ stand for the preference relation of a generic subject of type c.

Given the defining properties of a certainty equivalence function, it is a routine exercise to show that for every $p \in \Delta(X)$, there exists a unique number $u(p) \in [0,1]$ such that $u(p)\delta_{x^*} + (1-u(p))\delta_{x_*} \sim_c p$. In fact, since $\succsim_c$ is monotonic w.r.t. $\geq_{\text{fosd}}$, the function u represents $\succsim_c$. That is, $u(p) \geq u(q)$ iff $p \succsim_c q$ for any $p, q \in \Delta(X)$. It is also easy to see that u is continuous on $\Delta(X)$. In the remainder of the proof, I shall show that u is an expected utility functional.

Claim 1. *For any $\mu \in \Delta^2(S)$, we have $u(\mu_{f^{b2}}) + u(\mu_{f^{o2}}) + u(\mu_{f^{w2}}) = 1$.*

Proof. If $u(\mu_{f^{b2}}) + u(\mu_{f^{o2}}) + u(\mu_{f^{w2}}) < 1$ for some $\mu \in \Delta^2(S)$, we can select the fractions γ^k such that $\gamma^k > u(\mu_{f^{k2}})$ for $k = b, o, w$. Observe that $\gamma^k = u(\gamma^k \delta_{x^*} + (1-\gamma^k)\delta_{x_*})$. Thus, $\gamma^k > u(\mu_{f^{k2}})$ means $\gamma^k\delta_{x^*} + (1-\gamma^k)\delta_{x_*} \succ \mu_{f^{k2}}$, or equivalently, $f^{k1} \succ f^{k2}$ for $k = b, o, w$. The latter statement is simply the Ellsberg paradox. Similarly, if $u(\mu_{f^{b2}}) + u(\mu_{f^{o2}}) + u(\mu_{f^{w2}}) > 1$ for some $\mu \in \Delta^2(S)$, we can select γ^k such that $\gamma^k < u(\mu_{f^{k2}})$ for $k = b, o, w$. This, in turn, implies the anti-Ellsberg paradox. $\square$

Claim 2. *For any $n \in \mathbb{N}$ and $i = 1, ..., n$, let $\lambda^i, \beta^i, \alpha^i, \alpha^{i1}, \alpha^{i2}$ be numbers in $[0,1]$ such that $\sum_{i=1}^n \lambda^i = 1$ and $\alpha^i + \max\{\alpha^{i1}, \alpha^{i2}\} \leq 1$ for every i. Set*

$$
\begin{aligned}
p^b &:= \textstyle\sum_{i=1}^n \lambda^i \delta_{c(\alpha^i\delta_{x^*} + (1-\alpha^i)\delta_{x_*})} \\
p^o &:= \textstyle\sum_{i=1}^n \lambda^i \left(\beta^i \delta_{c(\alpha^{i1}\delta_{x^*} + (1-\alpha^{i1})\delta_{x_*})} + (1-\beta^i)\delta_{c(\alpha^{i2}\delta_{x^*} + (1-\alpha^{i2})\delta_{x_*})}\right) \\
p^w &:= \textstyle\sum_{i=1}^n \lambda^i \left(\beta^i \delta_{c((1-\alpha^i-\alpha^{i1})\delta_{x^*} + (\alpha^i+\alpha^{i1})\delta_{x_*})} + (1-\beta^i)\delta_{c((1-\alpha^i-\alpha^{i2})\delta_{x^*} + (\alpha^i+\alpha^{i2})\delta_{x_*})}\right).
\end{aligned}
$$

Then, $u(p^b) + u(p^o) + u(p^w) = 1$.

Proof. For every $i = 1, ..., n$ and $j = 1, 2$, let π^{ij} denote the element of $\Delta(S)$ defined as $\pi^{ij}(b) = \alpha^i$, $\pi^{ij}(o) = \alpha^{ij}$ and $\pi^{ij}(w) = 1 - \alpha^i - \alpha^{ij}$. Set $\mu := \sum_{i=1}^n \lambda^i (\beta^i D_{\pi^{i1}} + (1-\beta^i) D_{\pi^{i2}})$ so that

$$\mu_{f^{k2}} = \sum_{i=1}^n \lambda^i \left(\beta^i \delta_{c(\pi^{i1}_{f^{k2}})} + (1-\beta^i)\delta_{c(\pi^{i2}_{f^{k2}})}\right) \text{ for } k = b, o, w.$$

Using the fact that $\pi^{ij}_{f^{k2}} = \pi^{ij}(k)\delta_{x^*} + (1-\pi^{ij}(k))\delta_{x_*}$ for every i, j and k, it can easily be checked that $\mu_{f^{k2}} = p^k$ for $k = b, o, w$. Thus, the proof follows from Claim 1. $\square$

In what follows, set $v(x) := u(\delta_x)$ for $x \in X$. Observe that

$$c\left(v(x)\delta_{x^*} + (1-v(x))\delta_{x_*}\right) = x \quad \forall x \in X$$

because $u\left(v(x)\delta_{x^*} + (1-v(x))\delta_{x_*}\right) = v(x) = u(\delta_x)$. I proceed with two implications of Claim 2.

Pick any $\theta \in [0,1]$ and some $x, y \in X$ with $v(x) + v(y) \leq 1$. In the statement of Claim 2, let $n = 1, \lambda^1 = 1$, $\beta^1 = \theta$, $\alpha^1 = 1 - v(x) - v(y)$, $\alpha^{11} = v(x)$ and $\alpha^{12} = v(y)$. Then,

$$
\begin{aligned}
p^b &= \delta_{c((1-v(x)-v(y))\delta_{x^*} + (v(x)+v(y))\delta_{x_*})} \\
p^o &= \theta\delta_{c(v(x)\delta_{x^*} + (1-v(x))\delta_{x_*})} + (1-\theta)\delta_{c(v(y)\delta_{x^*} + (1-v(y))\delta_{x_*})} = \theta\delta_x + (1-\theta)\delta_y \\
p^w &= \theta\delta_{c(v(y)\delta_{x^*} + (1-v(y))\delta_{x_*})} + (1-\theta)\delta_{c(v(x)\delta_{x^*} + (1-v(x))\delta_{x_*})} = \theta\delta_y + (1-\theta)\delta_x.
\end{aligned}
\tag{19}
$$

By definitions, the first equality in (19) implies $u(p^b) = 1 - v(x) - v(y)$. Hence, the latter two equalities in (19) combined with Claim 2 yield, for every $\theta \in [0,1]$ and $x, y \in X$ with $v(x)+v(y) \leq 1$,

$$u\left(\theta\delta_x + (1-\theta)\delta_y\right) + u\left(\theta\delta_y + (1-\theta)\delta_x\right) = v(x) + v(y). \tag{20}$$

To uncover the second implication of Claim 2, pick any $\theta \in [0,1]$ and $x, y \in X$. In the statement of the claim, let $n = 1$, $\lambda^1 = 1$, $\beta^1 = \theta$, $\alpha^1 = 0$, $\alpha^{11} = v(x)$ and $\alpha^{12} = v(y)$. Then, $p^b = \delta_{x_*}$, and hence, $u(p^b) = 0$. Moreover, $p^o = \theta\delta_x + (1-\theta)\delta_y$ as in equation (19), while $p^w = \theta\delta_{c((1-v(x))\delta_{x^*}+v(x)\delta_{x_*})} + (1-\theta)\delta_{c((1-v(y))\delta_{x^*}+v(y)\delta_{x_*})}$. Thus, by Claim 2, for every $\theta \in [0,1]$ and $x, y \in X$ we have

$$u\left(\theta\delta_x + (1-\theta)\delta_y\right) + u\left(\theta\delta_{x'} + (1-\theta)\delta_{y'}\right) = 1, \tag{21}$$

where $x' := c((1-v(x))\delta_{x^*} + v(x)\delta_{x_*})$ and $y' := c((1-v(y))\delta_{x^*} + v(y)\delta_{x_*})$.

Claim 3. *$u\left(\theta\delta_x + (1-\theta)\delta_y\right) = \theta v(x) + (1-\theta)v(y)$ for any $\theta \in [0,1]$ and $x, y \in X$.*

Proof. By continuity of u, it suffices to establish the claim for all dyadic θ. For every $L = 0, 1, 2, ...$, set $\mathbb{Q}_L := \{\frac{\ell}{2^L} : \ell = 0, 1, ..., 2^L\}$. Observe that the claim trivially holds for $\theta \in \mathbb{Q}_0 = \{0, 1\}$. Inductively, given an $L \geq 0$, suppose that the claim holds for every $\theta \in \mathbb{Q}_L$. Pick any $\theta \in \mathbb{Q}_{L+1}$. By changing the roles of θ and $1-\theta$ if necessary, we can assume $\theta \leq 1/2$. Then, there exists an integer ℓ such that $0 \leq \ell \leq 2^L$ and $\theta = \frac{\ell}{2^{L+1}}$. Pick any $x, y \in X$, and set $q := \theta\delta_x + (1-\theta)\delta_y$.

First, assume $v(x)+v(y) \leq 1$ and $v(y) \leq v(x)$, so that $v(y) \leq 1/2$. In the statement of Claim 2, let $n = 2$, and

$$\begin{pmatrix} \lambda^1 \\ \beta^1 \\ \alpha^1 \\ \alpha^{11} \\ \alpha^{12} \end{pmatrix} = \begin{pmatrix} 2\theta \\ \frac{1}{2} \\ 1 - v(x) - v(y) \\ v(x) \\ v(y) \end{pmatrix}, \quad \begin{pmatrix} \lambda^2 \\ \beta^2 \\ \alpha^2 \\ \alpha^{21} \\ \alpha^{22} \end{pmatrix} = \begin{pmatrix} 1-2\theta \\ 1 \\ 1 - 2v(y) \\ v(y) \\ v(y) \end{pmatrix}.$$

Then, it is easily checked that

$$\begin{aligned} p^b &= 2\theta\delta_{c((1-v(x)-v(y))\delta_{x^*}+(v(x)+v(y))\delta_{x_*})} + (1-2\theta)\delta_{c((1-2v(y))\delta_{x^*}+2v(y)\delta_{x_*})} \\ p^o &= 2\theta\left(\tfrac{1}{2}\delta_x + \tfrac{1}{2}\delta_y\right) + (1-2\theta)\delta_y = q \\ p^w &= 2\theta\left(\tfrac{1}{2}\delta_y + \tfrac{1}{2}\delta_x\right) + (1-2\theta)\delta_y = q. \end{aligned} \tag{22}$$

Observe that $2\theta = \frac{\ell}{2^L}$ belongs to $\mathbb{Q}_L$. Thus, the first equation in (22) and the induction hypothesis imply $u(p^b) = 2\theta v^1 + (1-2\theta)v^2$, where

$$\begin{aligned} v^1 &:= v\left(c((1-v(x)-v(y))\delta_{x^*} + (v(x)+v(y))\delta_{x_*})\right) \\ v^2 &:= v\left(c((1-2v(y))\delta_{x^*} + 2v(y)\delta_{x_*})\right). \end{aligned}$$

Moreover, by definitions, $v\left(c(\lambda\delta_{x^*} + (1-\lambda)\delta_{x_*})\right) = \lambda$ for every $\lambda \in [0,1]$. In particular, $v^1 = 1 - v(x) - v(y)$ while $v^2 = 1 - 2v(y)$. Hence,

$$u(p^b) = 2\theta\left(1 - v(x) - v(y)\right) + (1-2\theta)\left(1 - 2v(y)\right) = 1 - 2\left(\theta v(x) + (1-\theta)v(y)\right).$$

In turn, if we substitute into the conclusion of Claim 2 the equation above together with the last two equations in (22), we obtain

$$u(q) = \frac{1}{2}\left(1 - u(p^b)\right) = \theta v(x) + (1-\theta)v(y).$$

This establishes the claim for the case $v(x) + v(y) \leq 1$ and $v(y) \leq v(x)$.

Suppose now that $v(x)+v(y) \leq 1$ and $v(y) > v(x)$. Then, by the first part of the proof, $u(\theta\delta_y + (1-\theta)\delta_x) = \theta v(y) + (1-\theta)v(x)$, and hence, the desired conclusion easily follows from equation (20).

Finally, assume $v(x) + v(y) > 1$, and let $x' := c((1-v(x))\delta_{x^*} + v(x)\delta_{x_*})$ and $y' := c((1-v(y))\delta_{x^*} + v(y)\delta_{x_*})$. Then, $v(x') + v(y') = 1 - v(x) + 1 - v(y) < 1$, and hence, $u(\theta\delta_{x'} + (1-\theta)\delta_{y'}) = \theta v(x') + (1-\theta)v(y')$ as we have just seen. Thus, in this case, the desired conclusion follows from equation (21). $\square$

Claim 4. *$u\left(\sum_{i=1}^n \theta^i\delta_{x^i}\right) = \sum_{i=1}^n \theta^i v(x^i)$ for any $n \in \mathbb{N}$, $\{x^1, ..., x^n\} \subseteq X$ and $\{\theta^1, ..., \theta^n\} \subseteq [0,1]$ with $\sum_{i=1}^n \theta^i = 1$.*

Proof. For $n = 2$, the desired conclusion follows from Claim 3. Inductively, fix an integer $n \geq 2$, and assume that the desired conclusion holds for n. Pick any $\{x^1, ..., x^{n+1}\} \subseteq X$ and $\{\theta^1, ..., \theta^{n+1}\} \subseteq (0,1)$ with $\sum_{i=1}^{n+1} \theta^i = 1$. Without loss of generality assume $x^1 \leq x^2 \leq \cdots \leq x^{n+1}$.

Given the number n, in the statement of Claim 2, let $\alpha^1 = v(x^{n+1}) - v(x^1)$, $\alpha^i = 0$ for $i = 2, ..., n$, and

$$\begin{pmatrix} \lambda^i \\ \beta^i \\ \alpha^{i1} \\ \alpha^{i2} \end{pmatrix} = \begin{pmatrix} \theta^i \\ 1 \\ v(x^i) \\ v(x^i) \end{pmatrix} \text{ for } i = 1, ..., n-1, \quad \begin{pmatrix} \lambda^n \\ \beta^n \\ \alpha^{n1} \\ \alpha^{n2} \end{pmatrix} = \begin{pmatrix} \theta^n + \theta^{n+1} \\ \frac{\theta^n}{\theta^n + \theta^{n+1}} \\ v(x^n) \\ v(x^{n+1}) \end{pmatrix}.$$

Observe that, by construction, $\alpha^1 \geq 0$ and $\alpha^1 + \alpha^{1j} = v(x^{n+1}) \leq 1$ for $j = 1, 2$. Clearly, the analogous inequalities also hold for $i \geq 2$, while $\sum_{i=1}^{n} \lambda^i = \sum_{i=1}^{n+1} \theta^i = 1$. So, Claim 2 is applicable with the selected parameters.

By definitions of λ^1 and $\alpha^1, ..., \alpha^n$, we have

$$p^b = \theta^1 \delta_{c((v(x^{n+1})-v(x^1))\delta_{x^*}+(1-v(x^{n+1})+v(x^1))\delta_{x_*})} + (1-\theta^1)\delta_{x_*}.$$

Thus, Claim 3 implies

$$u(p^b) = \theta^1 \left(v(x^{n+1}) - v(x^1)\right) + \left(1-\theta^1\right) v(x_*) = \theta^1 \left(v(x^{n+1}) - v(x^1)\right). \tag{23}$$

Moreover,

$$\begin{aligned} p^w &= \theta^1 \delta_{c((1-v(x^{n+1}))\delta_{x^*}+v(x^{n+1})\delta_{x_*})} + \textstyle\sum_{i=2}^{n-1} \theta^i \delta_{c((1-v(x^i))\delta_{x^*}+v(x^i)\delta_{x_*})} \\ &\qquad + \left(\theta^n + \theta^{n+1}\right) \left(\tfrac{\theta^n}{\theta^n+\theta^{n+1}} \delta_{c((1-v(x^n))\delta_{x^*}+v(x^n)\delta_{x_*})} + \tfrac{\theta^{n+1}}{\theta^n+\theta^{n+1}} \delta_{c((1-v(x^{n+1}))\delta_{x^*}+v(x^{n+1})\delta_{x_*})}\right) \\ &= \theta^1 \delta_{c((1-v(x^{n+1}))\delta_{x^*}+v(x^{n+1})\delta_{x_*})} + \textstyle\sum_{i=2}^{n} \theta^i \delta_{c((1-v(x^i))\delta_{x^*}+v(x^i)\delta_{x_*})} \\ &\qquad + \theta^{n+1} \delta_{c((1-v(x^{n+1}))\delta_{x^*}+v(x^{n+1})\delta_{x_*})}. \end{aligned}$$

Observe that the same degenerate lottery appears in the first and the last terms on the right hand side of the latter equality. So, this equality can be rewritten as

$$p^w = \left(\theta^1 + \theta^{n+1}\right) \delta_{c((1-v(x^{n+1}))\delta_{x^*}+v(x^{n+1})\delta_{x_*})} + \sum_{i=2}^{n} \theta^i \delta_{c((1-v(x^i))\delta_{x^*}+v(x^i)\delta_{x_*})}.$$

Hence, by the induction hypothesis,

$$u(p^w) = \left(\theta^1 + \theta^{n+1}\right) \left(1 - v(x^{n+1})\right) + \sum_{i=2}^{n} \theta^i \left(1 - v(x^i)\right). \tag{24}$$

From equations (23) and (24), it easily follows that $1 - u(p^b) - u(p^w) = \sum_{i=1}^{n+1} \theta^i v(x^i)$. Finally, note that

$$p^o = \sum_{i=1}^{n-1} \theta^i \delta_{x^i} + \left(\theta^n + \theta^{n+1}\right) \left(\frac{\theta^n}{\theta^n + \theta^{n+1}} \delta_{x^n} + \frac{\theta^{n+1}}{\theta^n + \theta^{n+1}} \delta_{x^{n+1}}\right) = \sum_{i=1}^{n+1} \theta^i \delta_{x^i}.$$

Thus, Claim 2 implies $u\left(\sum_{i=1}^{n+1} \theta^i \delta_{x^i}\right) = 1 - u(p^b) - u(p^w) = \sum_{i=1}^{n+1} \theta^i v(x^i)$. This proves Claim 4, which also completes the proof of Theorem 2. $\square$

The following lemma will be useful in the proof of Theorem 3.

Lemma 3. *Given a state space S, let $\succsim$ and $\succsim'$ denote a pair of recursive preferences on $\Delta(\mathcal{H})$ represented by (μ, c) and (μ', c'), respectively. $\succsim$ is more ambiguity averse than $\succsim'$*

if and only if $c = c'$ *and* $\mu'_f \succsim \mu_f$ *for every* $f \in \mathcal{H}_X$.

Proof. Suppose $\succsim$ is more ambiguity averse than $\succsim'$. Given any $p \in \Delta(X)$, $p \sim' \delta_{c'(p)}$ implies $p \sim \delta_{c'(p)}$ by part (ii) of Definition 3. As $p \sim \delta_{c(p)}$, it then follows that $\delta_{c(p)} \sim \delta_{c'(p)}$. This means $c(p) = c'(p)$ because $\succsim$ is monotonic w.r.t. $\geq_{\text{fosd}}$ over $\Delta(X)$. Since p is an arbitrary element of $\Delta(X)$, we conclude that $c = c'$.

Now, fix any $f \in \mathcal{H}_X$. As $f \sim \delta_{c(\mu_f)}$, part (i) of Definition 3 implies $f \succsim' \delta_{c(\mu_f)}$. Since $f \sim' \delta_{c'(\mu'_f)} = \delta_{c(\mu'_f)}$, we then see that $\delta_{c(\mu'_f)} \succsim' \delta_{c(\mu_f)}$. That is, $\mu'_f \succsim' \mu_f$, or equivalently, $\mu'_f \succsim \mu_f$.

Conversely, suppose $c = c'$ and $\mu'_f \succsim \mu_f$ for every $f \in \mathcal{H}_X$. Part (ii) of Definition 3 immediately follows from the fact that $c = c'$ represents both $\succsim$ and $\succsim'$ over $\Delta(X)$. As for part (i), fix any $f \in \mathcal{H}_X$. Then, $\mu'_f \succsim \mu_f \sim f$, and hence, $\mu'_f \succsim \delta_x$ for any $x \in X$ with $f \succsim \delta_x$. In turn, $\mu'_f \succsim \delta_x$ means $\mu'_f \succsim' \delta_x$. Since $f \sim' \mu'_f$, it follows that $f \succsim' \delta_x$ for any $x \in X$ with $f \succsim \delta_x$. $\square$

Proof of Theorem 3. Fix a pair $\mu, \mu' \in \Delta^2(S)$. Given a generic certainty equivalence function c, let $\succsim$ and $\succsim'$ denote the preference relations on $\Delta(\mathcal{H})$ represented by (μ, c) and (μ', c), respectively.

By Lemma 3, $\succsim$ is more ambiguity averse than $\succsim'$ if and only if $\mu'_f \succsim \mu_f$ for every $f \in \mathcal{H}_X$. In Theorem 3, that the statement (i) implies (ii) follows from this observation and the related arguments in Section 5.

For the converse implication, suppose $\succsim$ is more ambiguity averse than $\succsim'$ for every specification of c that is risk averse and that satisfies NCI. Observe that this is a weak form of statement (ii) that is restricted to risk averse certainty equivalence functions.

Applying Lemma 3 once again yields $\mu'_f \succsim \mu_f$ for every $f \in \mathcal{H}_X$ and every risk averse c that satisfies NCI. Equivalently, $c(\mu'_f) \geq c(\mu_f)$ for every such f and c.

Claim 5. $\bar{\mu}' = \bar{\mu}$.

Proof. Let $c(p) := \sum_{x \in X} p(x)x$ for $p \in \Delta(X)$, so that c is the risk neutral expectation functional. Fix any $\hat{s} \in S$, and set $f := x_{\{\hat{s}\}}y$ for some $x, y \in X$ with $x > y$. Then, $\pi_f = \pi(\hat{s})\delta_x + (1 - \pi(\hat{s}))\delta_y$ for every $\pi \in \Delta(S)$. Hence,

$$\begin{aligned} c(\mu_f) &= c\left(\sum_{\pi \in \Delta(S)} \mu(\pi)\delta_{c(\pi_f)}\right) = \sum_{\pi \in \Delta(S)} \mu(\pi)c(\pi_f) \\ &= \sum_{\pi \in \Delta(S)} \mu(\pi)\left(\pi(\hat{s})x + (1 - \pi(\hat{s}))y\right) \\ &= \bar{\mu}(\hat{s})x + (1 - \bar{\mu}(\hat{s}))y, \end{aligned}$$

where the last equality follows from the definition of $\bar{\mu}(\hat{s})$. Similarly, $c(\mu'_f) = \bar{\mu}'(\hat{s})x + (1 - \bar{\mu}'(\hat{s}))\,y$. So, $c(\mu'_f) \geq c(\mu_f)$ means $\bar{\mu}'(\hat{s}) \geq \bar{\mu}(\hat{s})$. This implies $\bar{\mu}' = \bar{\mu}$ because $\hat{s}$ is an arbitrary point in S. $\square$

I proceed with a bit of notation. C_w stands for the set of all (strictly) increasing, continuous and concave functions on X. In turn, C_{st} is the set of all strictly concave functions that belong to C_w. Given a $v \in C_w$ and a nonempty, open interval J contained in (x_*, x^*), let v_J denote the continuous function on X that is linear on J and equal to v on $X \backslash J$. Specifically,

$$v_J(x) := \begin{cases} v(x) & \text{if } x \in X \backslash J, \\ v(\inf J) + (x - \inf J)\frac{v(\sup J) - v(\inf J)}{\sup J - \inf J} & \text{if } x \in J. \end{cases}$$

Note that v_J is concave and increasing because of the corresponding properties of v. Moreover, if v is strictly concave on the interval J, then $v(x) > v_J(x)$ for every $x \in J$.

Π denotes the set of all relevant first-order distributions. That is,

$$\Pi := \{\bar{\mu}\} \cup \{\pi \in \Delta(S) : \mu(\pi) > 0\} \cup \{\pi \in \Delta(S) : \mu'(\pi) > 0\} \cup \{\delta_s : s \in S\}.$$

Since S is finite, there exists a one-to-one function $\mathbf{y}$ that maps S into (x_*, x^*). Let f stand for the purely subjective act defined as $f(s) := \delta_{\mathbf{y}(s)}$ for $s \in S$. Then,

$$\pi_f = \sum_{s \in S} \pi(s)\delta_{\mathbf{y}(s)} \quad \text{and} \quad E_{\pi_f}(v) = \sum_{s \in S} \pi(s)v(\mathbf{y}(s)) \quad \forall(\pi, v) \in \Delta(S) \times C_w.$$

Claim 6. *There exists a $v \in C_{st}$ such that $E_{\pi_f}(v) \neq E_{\hat{\pi}_f}(v)$ for every $\pi, \hat{\pi} \in \Pi$ with $\pi \neq \hat{\pi}$.*

Proof. Since Π is a finite set, by an obvious, inductive argument, it suffices to show that given any $v^1 \in C_{st}$, if $E_{\pi^1_f}(v^1) = E_{\pi^2_f}(v^1)$ for some $\pi^1, \pi^2 \in \Pi$ with $\pi^1 \neq \pi^2$, then there exists a $v^2 \in C_{st}$ that satisfies the following two properties:

(a) $E_{\pi^1_f}(v^2) \neq E_{\pi^2_f}(v^2)$.

(b) $E_{\pi_f}(v^1) \neq E_{\hat{\pi}_f}(v^1)$ implies $E_{\pi_f}(v^2) \neq E_{\hat{\pi}_f}(v^2)$ for every $\pi, \hat{\pi} \in \Pi$.

Note that any $v^2 \in C_{st}$ that is sufficiently close to v^1 (in sup-norm) satisfies the condition (b) because Π is a finite set, and the expectation operator $E_p(v)$ is continuous in v for every $p \in \Delta(X)$.

Fix a $v^1 \in C_{st}$ and suppose $E_{\pi^1_f}(v^1) = E_{\pi^2_f}(v^1)$ for some $\pi^1, \pi^2 \in \Pi$ with $\pi^1 \neq \pi^2$. It remains to show that any neighborhood of v^1 contains a $v^2 \in C_{st}$ that satisfies the condition (a).

Since $\pi^1 \neq \pi^2$, there exists an $\bar{s} \in S$ such that $\pi^1(\bar{s}) > \pi^2(\bar{s})$. Recall that $\mathbf{y}$ is a one-to-one function with values in (x_*, x^*). Hence, there exists an open interval $J \subseteq (x_*, x^*)$ such that $\mathbf{y}(\bar{s}) \in J$ and $\mathbf{y}(s) \notin J$ for every $s \in S \backslash \{\bar{s}\}$. By definition of v^1_J, the latter statement implies $v^1_J(\mathbf{y}(s)) = v^1(\mathbf{y}(s))$ for every $s \in S\backslash\{\bar{s}\}$. Thus,

$$E_{\pi^i_f}(v^1) - E_{\pi^i_f}(v^1_J) = \pi^i(\bar{s})\left(v^1(\mathbf{y}(\bar{s})) - v^1_J(\mathbf{y}(\bar{s}))\right) \quad \text{for } i = 1, 2.$$

Moreover, $v^1(\mathbf{y}(\bar{s})) > v^1_J(\mathbf{y}(\bar{s}))$ because v^1 is strictly concave and $\mathbf{y}(\bar{s}) \in J$. Since $\pi^1(\bar{s}) > \pi^2(\bar{s})$, it follows that $E_{\pi^1_f}(v^1) - E_{\pi^1_f}(v^1_J) > E_{\pi^2_f}(v^1) - E_{\pi^2_f}(v^1_J)$. We must then have $E_{\pi^1_f}(v^1_J) < E_{\pi^2_f}(v^1_J)$ because $E_{\pi^1_f}(v^1) = E_{\pi^2_f}(v^1)$ by assumption.

Set $v^2 := \lambda v^1 + (1-\lambda)v^1_J$ for an arbitrarily selected $\lambda \in (0,1)$. Since $v^1 \in C_{st}$ and $v^1_J \in C_w$, it is clear that $v^2 \in C_{st}$. Moreover,

$$\begin{aligned} E_{\pi^1_f}(v^2) - E_{\pi^2_f}(v^2) &= \lambda\left(E_{\pi^1_f}(v^1) - E_{\pi^2_f}(v^1)\right) + (1-\lambda)\left(E_{\pi^1_f}(v^1_J) - E_{\pi^2_f}(v^1_J)\right) \\ &= (1-\lambda)\left(E_{\pi^1_f}(v^1_J) - E_{\pi^2_f}(v^1_J)\right). \end{aligned}$$

It follows that $E_{\pi^1_f}(v^2) \neq E_{\pi^2_f}(v^2)$, as we sought. This completes the proof because v^2 converges to v^1 as $\lambda \to 1$. $\square$

Fix a function $v \in C_{st}$ as in Claim 6. Let m and $n+m$ denote the cardinality of the sets S and Π, respectively. Label the elements of Π as $\pi^1, ..., \pi^{n+m}$ in such a way that $\{\pi^{n+1}, ..., \pi^{n+m}\} = \{\delta_s : s \in S\}$. Observe that $\bar{\mu}' = \delta_s$ would imply $\mu' = D_{\delta_s}$, which contradicts the regularity assumption. Hence, $n \geq 1$, and $\bar{\mu}' = \bar{\mu}$ belongs to $\{\pi^1, ..., \pi^n\}$. Let $\bar{\imath}$ denote the particular index such that $\pi^{\bar{\imath}} = \bar{\mu}$.

For $i = 1, ..., n+m$, set $x^i := v^{-1}(E_{\pi^i_f}(v))$. Since v separates $\pi^1_f, ..., \pi^{n+m}_f$ (as in Claim 6), the points $x^1, ..., x^{n+m}$ are pairwise distinct. It is also clear that $x^i \in (x_*, x^*)$ for every i because π^i_f is supported in (x_*, x^*). Hence, there exists a collection of pairwise disjoint open intervals $J_1, ..., J_{n+m}$ contained in (x_*, x^*) such that $x^i \in J_i$ for every $i = 1, ..., n+m$.

Fix a $k \in \{1, ..., n\}$ that is distinct from $\bar{\imath}$. Given any pair of numbers $\lambda^k, \lambda^{\bar{\imath}} \geq 0$ with $\lambda^k + \lambda^{\bar{\imath}} \leq 1$ define

$$\hat{v} := \lambda^k v_{J_k} + \lambda^{\bar{\imath}} v_{J_{\bar{\imath}}} + (1 - \lambda^k - \lambda^{\bar{\imath}})v.$$

Since v_{J_k}, $v_{J_{\bar{\imath}}}$ and v belong to C_w, so does $\hat{v}$. Set

$$c(p) := \min\left\{v^{-1}(E_p(v)), \hat{v}^{-1}(E_p(\hat{v}))\right\} \quad \forall p \in \Delta(X).$$

This is a risk averse certainty equivalence function that satisfies NCI because v and $\hat{v}$ both belong to C_w (see Cerreia-Vioglio et al., 2015, Theorem 3).

As we shall see momentarily, for suitably selected values of λ^k and $\lambda^{\bar{\imath}}$, we will have

$$\mu(\pi^k)(v(x^k) - \hat{v}(x^k)) \geq (1 - \mu(\pi^{\bar{\imath}}))(v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}})). \tag{25}$$

The next claim uncovers a key implication of this inequality.

Claim 7. (25) *implies*

$$(\mu(\pi^k) - \mu'(\pi^k))(v(x^k) - \hat{v}(x^k)) + (\mu(\pi^{\bar{\imath}}) - \mu'(\pi^{\bar{\imath}}))(v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}})) \geq 0. \tag{26}$$

Proof. By construction, $\hat{v}(x^i) = v(x^i)$ for any $i = 1, ..., n+m$ with $i \notin \{k, \bar{\imath}\}$. In particular, $\hat{v}(x^i) = v(x^i)$ for every $i \geq n+1$. Moreover, it is easily checked that $\{x^{n+1}, ..., x^{n+m}\} = \{\mathbf{y}(s) : s \in S\}$ because $\{\pi_f^{n+1}, ..., \pi_f^{n+m}\} = \{\delta_{\mathbf{y}(s)} : s \in S\}$. It follows that $\hat{v}(\mathbf{y}(s)) = v(\mathbf{y}(s))$ for every $s \in S$, and hence,

$$E_{\pi_f}(\hat{v}) = E_{\pi_f}(v) \quad \forall \pi \in \Delta(S). \tag{27}$$

Also note that the range of v, $v(X)$, is the same as that of $\hat{v}$ because $\hat{v}(x_*) = v(x_*)$ and $\hat{v}(x^*) = v(x^*)$. Furthermore, $v(x) \geq \hat{v}(x)$ for every $x \in X$ by construction. It easily follows that $\hat{v}^{-1}(a) \geq v^{-1}(a)$ for every $a \in v(X)$.

Combining the latter inequality with (27) yields $\hat{v}^{-1}(E_{\pi_f}(\hat{v})) \geq v^{-1}(E_{\pi_f}(v))$ for every $\pi \in \Delta(S)$, which means $c(\pi_f) = v^{-1}(E_{\pi_f}(v))$. In particular, $c(\pi_f^i) = x^i$ for $i = 1, ..., n+m$, and hence, $\mu_f = \sum_{i=1}^{n+m} \mu(\pi^i)\delta_{c(\pi_f^i)} = \sum_{i=1}^{n+m} \mu(\pi^i)\delta_{x^i}$. Similarly, $\mu'_f = \sum_{i=1}^{n+m} \mu'(\pi^i)\delta_{x^i}$. Since $v(x^i) = \hat{v}(x^i)$ for any $i \notin \{k, \bar{\imath}\}$, it follows that

$$\begin{array}{rcccl} E_{\mu_f}(v) - E_{\mu_f}(\hat{v}) & = & \mu(\pi^k)(v(x^k) - \hat{v}(x^k)) & + & \mu(\pi^{\bar{\imath}})(v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}})), \\ E_{\mu'_f}(v) - E_{\mu'_f}(\hat{v}) & = & \mu'(\pi^k)(v(x^k) - \hat{v}(x^k)) & + & \mu'(\pi^{\bar{\imath}})(v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}})). \end{array} \tag{28}$$

Thus, (26) is equivalent to the following inequality

$$(E_{\mu_f}(v) - E_{\mu_f}(\hat{v})) - (E_{\mu'_f}(v) - E_{\mu'_f}(\hat{v})) \geq 0. \tag{29}$$

Let $p := \sum_{i=1}^{n+m} \mu(\pi^i)\pi_f^i$. Note that $E_{\mu_f}(v) = \sum_{i=1}^{n+m} \mu(\pi^i)v(x^i) = \sum_{i=1}^{n+m} \mu(\pi^i)E_{\pi_f^i}(v) = E_p(v)$. Here, the second equation follows from the definition of x^i, and the last one from the linearity of the expectation operator $q \to E_q(v)$. Moreover, equation (15) implies $\bar{\mu}_f = p$. Hence, $E_{\mu_f}(v) = E_{\bar{\mu}_f}(v)$, and similarly, $E_{\mu'_f}(v) = E_{\bar{\mu}'_f}(v)$. As $\bar{\mu} = \bar{\mu}' = \pi^{\bar{\imath}}$, we conclude that

$$E_{\mu_f}(v) = E_{\mu'_f}(v) = E_{\pi_f^{\bar{\imath}}}(v) = v(x^{\bar{\imath}}). \tag{30}$$

Since the first and last terms of (30) are equal to each other, we also have

$$(\hat{v}(x^{\bar{\imath}}) - v(x^{\bar{\imath}})) + (E_{\mu_f}(v) - E_{\mu_f}(\hat{v})) = \hat{v}(x^{\bar{\imath}}) - E_{\mu_f}(\hat{v}). \tag{31}$$

By the first equation in (28), it is easily verified that the left hand side of (31) equals

$\mu(\pi^k)(v(x^k) - \hat{v}(x^k)) - (1-\mu(\pi^{\bar{\imath}}))(v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}}))$. In turn, (25) amounts to saying that this number is nonnegative. Then, the right hand side of (31) must also be nonnegative. That is, $\hat{v}(x^{\bar{\imath}}) \geq E_{\mu_f}(\hat{v})$, or equivalently, $x^{\bar{\imath}} \geq \hat{v}^{-1}(E_{\mu_f}(\hat{v}))$. Also note that $x^{\bar{\imath}} = v^{-1}(E_{\mu_f}(v))$ by (30). Thus, $v^{-1}(E_{\mu_f}(v)) \geq \hat{v}^{-1}(E_{\mu_f}(\hat{v}))$, which means $c(\mu_f) = \hat{v}^{-1}(E_{\mu_f}(\hat{v}))$. Moreover, $\hat{v}^{-1}(E_{\mu'_f}(\hat{v})) \geq c(\mu'_f)$ by definition of c. Hence, $c(\mu'_f) \geq c(\mu_f)$ implies $\hat{v}^{-1}(E_{\mu'_f}(\hat{v})) \geq \hat{v}^{-1}(E_{\mu_f}(\hat{v}))$, which is equivalent to saying $E_{\mu'_f}(\hat{v}) \geq E_{\mu_f}(\hat{v})$. Finally, (29) follows from this inequality and the first equation in (30). $\square$

Observe that $v_{J_k}(x^{\bar{\imath}}) = v(x^{\bar{\imath}})$ because $x^{\bar{\imath}} \notin J_k$. Similarly, $v_{J_{\bar{\imath}}}(x^k) = v(x^k)$. Substituting these equations into the definition of $\hat{v}$ yields

$$v(x^i) - \hat{v}(x^i) = \lambda^i(v(x^i) - v_{J_i}(x^i)) \quad \text{for } i = \bar{\imath}, k. \tag{32}$$

Also recall that $v(x^i) - v_{J_i}(x^i)$ is strictly positive for every i because v is strictly concave. Hence, by (32), $v(x^i) - \hat{v}(x^i) > (=)0$ iff $\lambda^i > (=)0$ for $i = \bar{\imath}, k$.

Let $\lambda^{\bar{\imath}} = 0$ and $\lambda^k > 0$. As I just noted, $\lambda^{\bar{\imath}} = 0$ implies $v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}}) = 0$. Then, the right hand side of (25) also equals 0, while the left hand side is always nonnegative. So, (25) holds. In turn, (26) reduces to

$$(\mu(\pi^k) - \mu'(\pi^k))(v(x^k) - \hat{v}(x^k)) \geq 0. \tag{33}$$

Moreover, $v(x^k) - \hat{v}(x^k) > 0$ because $\lambda^k > 0$. Hence, (33) implies $\mu(\pi^k) \geq \mu'(\pi^k)$. Since k is an arbitrary element of $\{1, ..., n\} \backslash \{\bar{\imath}\}$, it also follows that

$$1 - \mu(\pi^{\bar{\imath}}) \geq \sum_{i \neq \bar{\imath}}^{n} \mu(\pi^i) \geq \sum_{i \neq \bar{\imath}}^{n} \mu'(\pi^i) = 1 - \mu'(\pi^{\bar{\imath}}), \tag{34}$$

where the last equation holds due to regularity of μ'.

Note that if $\mu'(\pi^{\bar{\imath}}) = 1$, then $\mu' = D_{\bar{\mu}}$. In this case, it is trivially true that μ is a mean-preserving spread of μ'. Assume therefore that $\mu'(\pi^{\bar{\imath}}) < 1$. Then, by (34), $1 - \mu(\pi^{\bar{\imath}})$ and $1 - \mu'(\pi^{\bar{\imath}})$ are both strictly positive.

The next step is to show that

$$\frac{\mu(\pi^i)}{1 - \mu(\pi^{\bar{\imath}})} \geq \frac{\mu'(\pi^i)}{1 - \mu'(\pi^{\bar{\imath}})} \quad \forall i \in \{1, ..., n\} \backslash \{\bar{\imath}\}. \tag{35}$$

This inequality holds trivially for any i with $\mu'(\pi^i) = 0$. Fix a $k \in \{1, ..., n\} \backslash \{\bar{\imath}\}$ with $\mu'(\pi^k) > 0$. Then, $\mu(\pi^k) > 0$ because $\mu(\pi^k) \geq \mu'(\pi^k)$ as we have seen earlier. Pick strictly positive values for λ^k and $\lambda^{\bar{\imath}}$ such that

$$\frac{\lambda^{\bar{\imath}}}{\lambda^k} = \frac{v(x^k) - v_{J_k}(x^k)}{v(x^{\bar{\imath}}) - v_{J_{\bar{\imath}}}(x^{\bar{\imath}})} \frac{\mu(\pi^k)}{1 - \mu(\pi^{\bar{\imath}})}. \tag{36}$$

Indeed, we can select such λ^k and $\lambda^{\bar{\imath}}$ because the right hand side of (36) is strictly positive, while the ratio $\lambda^{\bar{\imath}}/\lambda^k$ can be manipulated arbitrarily, at least for small values of λ^k and $\lambda^{\bar{\imath}}$.

It is easy to verify that the equations (32) and (36) jointly imply (25), with equality. That is,

$$\frac{v(x^k) - \hat{v}(x^k)}{v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}})} = \frac{1 - \mu(\pi^{\bar{\imath}})}{\mu(\pi^k)}. \tag{37}$$

Multiply (26) by $\frac{\mu(\pi^k)}{v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}})}$, and then utilize (37) to substitute for $\frac{v(x^k) - \hat{v}(x^k)}{v(x^{\bar{\imath}}) - \hat{v}(x^{\bar{\imath}})}$. This gives

$$(\mu(\pi^k) - \mu'(\pi^k))(1 - \mu(\pi^{\bar{\imath}})) + (\mu(\pi^{\bar{\imath}}) - \mu'(\pi^{\bar{\imath}}))\mu(\pi^k) \geq 0,$$

or equivalently, $\mu(\pi^k)(1 - \mu'(\pi^{\bar{\imath}})) - \mu'(\pi^k)(1 - \mu(\pi^{\bar{\imath}})) \geq 0$. So, $\frac{\mu(\pi^k)}{1-\mu(\pi^{\bar{\imath}})} \geq \frac{\mu'(\pi^k)}{1-\mu'(\pi^{\bar{\imath}})}$, which proves (35).

(35), the first inequality in (34) and the regularity of μ' jointly imply

$$1 \geq \sum_{i \neq \bar{\imath}}^{n} \frac{\mu(\pi^i)}{1 - \mu(\pi^{\bar{\imath}})} \geq \sum_{i \neq \bar{\imath}}^{n} \frac{\mu'(\pi^i)}{1 - \mu'(\pi^{\bar{\imath}})} = 1.$$

Clearly, both inequality signs above can be replaced with equality. That is,

$$1 = \sum_{i \neq \bar{\imath}}^{n} \frac{\mu(\pi^i)}{1 - \mu(\pi^{\bar{\imath}})} = \sum_{i \neq \bar{\imath}}^{n} \frac{\mu'(\pi^i)}{1 - \mu'(\pi^{\bar{\imath}})}. \tag{38}$$

By the latter equality in (38), we can strengthen (35) to read as

$$\frac{\mu(\pi^i)}{1 - \mu(\pi^{\bar{\imath}})} = \frac{\mu'(\pi^i)}{1 - \mu'(\pi^{\bar{\imath}})} \quad \forall i \in \{1, ..., n\} \backslash \{\bar{\imath}\}.$$

Set $\alpha := \frac{1-\mu'(\pi^{\bar{\imath}})}{1-\mu(\pi^{\bar{\imath}})}$, so that $\alpha\mu(\pi^i) = \mu'(\pi^i)$ for every $i \in \{1, ..., n\} \backslash \{\bar{\imath}\}$. Note that $0 < \alpha \leq 1$, where the latter inequality follows from (34). Moreover, by the first equality in (38), for any $i > n$ we have $\mu(\pi^i) = 0$. So, $\mu = \sum_{i=1}^{n} \mu(\pi^i) D_{\pi^i}$, and

$$\begin{aligned} (1 - \alpha)D_{\pi^{\bar{\imath}}} + \alpha\mu &= (1 - \alpha + \alpha\mu(\pi^{\bar{\imath}}))D_{\pi^{\bar{\imath}}} + \sum_{i \neq \bar{\imath}}^{n} \alpha\mu(\pi^i) D_{\pi^i} \\ &= (1 - \alpha + \alpha\mu(\pi^{\bar{\imath}}))D_{\pi^{\bar{\imath}}} + \sum_{i \neq \bar{\imath}}^{n} \mu'(\pi^i) D_{\pi^i}. \end{aligned} \tag{39}$$

Given that the last expression in (39) is a probability measure on $\Delta(S)$, we must have $1-\alpha+\alpha\mu(\pi^{\bar{\imath}}) = 1-\sum_{i\neq\bar{\imath}}^{n}\mu'(\pi^i)$, which also equals $\mu'(\pi^{\bar{\imath}})$. Hence, by (39),

$$(1-\alpha)D_{\pi^{\bar{\imath}}}+\alpha\mu = \mu'(\pi^{\bar{\imath}})D_{\pi^{\bar{\imath}}}+\sum_{i\neq\bar{\imath}}^{n}\mu'(\pi^i)D_{\pi^i}. \tag{40}$$

Finally, note that the right hand side of (40) equals μ', while $\pi^{\bar{\imath}} := \bar{\mu}$. Thus, (40) delivers the desired conclusion: $(1-\alpha)D_{\bar{\mu}}+\alpha\mu = \mu'$.

As for risk loving preferences, observe that if v is a continuous, convex and increasing function on X, so is v_J for any open interval $J \subseteq (x_*, x^*)$. Moreover, $v_J \geq v$, with strict inequality for $x \in J$ whenever v is strictly convex; which is a key difference compared to the case of concave functions.

The proof of Claim 6 can be modified in a straightforward way to obtain a strictly convex function v such that $E_{\pi_f}(v) \neq E_{\hat{\pi}_f}(v)$ for any distinct $\pi, \hat{\pi} \in \Pi$. After defining $\hat{v}$ and c just as before, set $z^i := \hat{v}^{-1}(E_{\pi_f^i}(\hat{v}))$ for $i = 1, ..., n+m$. Note that $\lim_{\lambda^k+\lambda^{\bar{\imath}}\to 0} \hat{v} = v$, and hence, $\lim_{\lambda^k+\lambda^{\bar{\imath}}\to 0} z^i = x^i$ for $i = 1, ..., n+m$. In particular, we can assume $z^i \in J_i$ for every i.

In the remainder of the proof, the points $z^1, ..., z^{n+m}$ take the role of $x^1, ..., x^{n+m}$ in the earlier proof because $\hat{v} \geq v$ implies $c(\pi_f^i) = z^i$ for every i. Consequently, in (25) and (26) the points x^k and $x^{\bar{\imath}}$ should be replaced with z^k and $z^{\bar{\imath}}$, respectively. Following the logic of Claim 7, it can be shown that the direction of these inequalities also change. That is,

$$\begin{gathered}\mu(\pi^k)(v(z^k)-\hat{v}(z^k)) \leq (1-\mu(\pi^{\bar{\imath}}))(v(z^{\bar{\imath}})-\hat{v}(z^{\bar{\imath}})) \quad \Rightarrow \\ (\mu(\pi^k)-\mu'(\pi^k))(v(z^k)-\hat{v}(z^k)) + (\mu(\pi^{\bar{\imath}})-\mu'(\pi^{\bar{\imath}}))(v(z^{\bar{\imath}})-\hat{v}(z^{\bar{\imath}})) \leq 0.\end{gathered} \tag{41}$$

Also recall that $v(z^i)-\hat{v}(z^i) \leq 0$ for every i. Thus, in the remaining arguments, (41) acts as a perfect analogue of Claim 7. Regarding this part of the proof, the only notable issue is how λ^k and $\lambda^{\bar{\imath}}$ should be selected so that the first inequality in (41) holds with equality. (This is not so obvious because z^k and $z^{\bar{\imath}}$ depend on λ^k and $\lambda^{\bar{\imath}}$.) As in the case of concave functions, without loss of generality assume $\mu(\pi^k)$ and $1-\mu(\pi^{\bar{\imath}})$ are strictly positive. Given a small number $\varepsilon > 0$, focus on $\lambda^k \leq \varepsilon$ and set $\lambda^{\bar{\imath}} := \varepsilon - \lambda^k$. Then, $\lambda^k = \varepsilon$ implies $\mu(\pi^k)(v(z^k)-\hat{v}(z^k)) < 0 = (1-\mu(\pi^{\bar{\imath}}))(v(z^{\bar{\imath}})-\hat{v}(z^{\bar{\imath}}))$, while the converse inequality holds for $\lambda^k = 0$. Hence, by a standard argument for continuous functions, there exists a $\lambda^k \in (0, \varepsilon)$ such that $\mu(\pi^k)(v(z^k)-\hat{v}(z^k)) = (1-\mu(\pi^{\bar{\imath}}))(v(z^{\bar{\imath}})-\hat{v}(z^{\bar{\imath}}))$. $\square$

References

M. Abdellaoui, P. Klibanoff, L. Placido, Experiments on compound risk in relation to simple risk and to ambiguity, *Management Science* **61** (2015), 1306–1322.

M. Allais, Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine, *Econometrica* **21** (1953), 503-546.

F.J. Anscombe, R.J. Aumann, A definition of subjective probability, *Annals of Mathematical Statistics* **34** (1963), 199–205.

S. Artstein-Avidan, D. Dillenberger, Dynamic disappointment aversion, University of Pennsylvania, 2011, mimeo.

S. Cerreia-Vioglio, D. Dillenberger, P. Ortoleva, Cautious expected utility and the certainty effect, *Econometrica* **83** (2015), 693–728.

M. Dean, P. Ortoleva, Allais, Ellsberg, and preferences for hedging, *Theoretical Economics,* in press.

M. Dean, P. Ortoleva, Is it all connected? A testing ground for unified theories of behavioral economics phenomena, Columbia University, 2015, mimeo.

D. Dillenberger, Preferences for one-shot resolution of uncertainty and Allais-type behavior, *Econometrica* **78** (2010), 1973–2004.

D. Dillenberger, U. Segal, Recursive ambiguity and Machina's examples, *International Economic Review* **56** (2015a), 55–61.

D. Dillenberger, U. Segal, Skewed noise, University of Pennsylvania, 2015b, mimeo.

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

D. Ellsberg, Risk, ambiguity, and the Savage axioms, *Quarterly Journal of Economics* **75** (1961), 643–669.

L.G. Epstein, A definition of uncertainty aversion, *Review of Economic Studies* **66** (1999), 579–608.

H. Ergin, F. Gul, A theory of subjective compound lotteries, *Journal of Economic Theory* **144** (2009), 899–929.

Ö. Evren, On the existence of expected multi-utility representations, *Economic Theory* **35** (2008), 575–592.

P. Ghirardato, M. Marinacci, Ambiguity made precise: A comparative foundation, *Journal of Economic Theory* **102** (2002), 251–289.

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

Y. Halevy, Ellsberg revisited: An experimental study, *Econometrica* **75** (2007), 503–536.

P. Klibanoff, M. Marinacci, S. Mukerji, A smooth model of decision making under ambiguity, *Econometrica* **73** (2005), 1849–1892.

M.G. Kocher, A.M. Lahno, S.T. Trautmann, Ambiguity aversion is the exception, University of Heidelberg, 2016, mimeo.

I. Kopylov, Canonical utility functions and continuous preference extensions, *Journal of Mathematical Economics* **67** (2016), 32–37.

M.J. Machina, Risk, ambiguity, and the rank-dependence axioms, *American Economic Review* **99** (2009), 385–392.

M.J. Machina, Ambiguity aversion with three or more outcomes, *American Economic Review* **104** (2014), 3814–3840.

M. Machina, K. Ritzberger, N. Yannelis, D. Ellsberg, Introduction to the symposium issue, *Economic Theory* **48** (2011), 219–227.

M.J. Machina, D. Schmeidler, A more robust definition of subjective probability, *Econometrica* **60** (1992), 745–780.

M.J. Machina, D. Schmeidler, Bayes without Bernoulli: Simple conditions for probabilistically sophisticated choice, *Journal of Economic Theory* **67** (1995), 106–128.

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

M. Rothschild, J.E. Stiglitz, Increasing risk (i): A definition, *Journal of Economic Theory* **2** (1970), 225–243.

U. Segal, The Ellsberg paradox and risk aversion: An anticipated utility approach, *International Economic Review* **28** (1987), 175–202.

U. Segal, Two stage lotteries without the reduction axiom, *Econometrica* **58** (1990), 349–377.

K. Seo, Ambiguity and second-order belief, *Econometrica* **77** (2009), 1575–1605.