Making the Anscombe-Aumann Approach to Ambiguity Suited for Descriptive Applications

Stefan Trautmann[a] & Peter P. Wakker[b]

[a] Alfred-Weber-Institute for Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany, 49 6221 54 2952 (O), trautmann@uni-hd.de

[b] Corresponding author: Econometric Institute, Erasmus University, P.O. Box 1738, Rotterdam, 3000 DR, the Netherlands, +31-(0)10 - 408.12.65 (O); +31-(0)10 - 408.12.78/12.64; +31-(0)10 - 408.91.62 (F); Wakker@ese.eur.nl

September, 2015

Abstract

The Anscombe-Aumann (AA) model, originally introduced to give a normative basis to expected utility, is nowadays mostly used for another purpose: to analyze deviations from expected utility due to ambiguity (unknown probabilities). The AA model makes two ancillary assumptions that do not refer to ambiguity: expected utility for risk and backward induction. These assumptions, even if normatively appropriate, fail descriptively. We relax them while maintaining AA's convenient mixture operation, and thus make it possible to test and apply AA based ambiguity theories descriptively. We find three common assumptions violated: reference independence, universal ambiguity aversion, and weak certainty independence. We introduce and axiomatize a reference dependent generalization of Schmeidler's CEU theory that accommodates the violations found. That is, we extend the AA model to prospect theory.

Keywords: ambiguity, reference dependence, certainty independence, prospect theory, loss aversion

1 Introduction

Keynes (1921) and Knight (1921) emphasized the need to develop theories for decision making when probabilities are unknown. This led Savage (1954) and others to provide a behavioral foundation of (subjective) expected utility: if no objective probabilities are available, then subjective probabilities should be used instead. This result provided the basis of decision analysis. However, Ellsberg (1961) provided two paradoxes showing that Savage's theory fails descriptively, and according to some also normatively. It led to the development of modern ambiguity theories; i.e., decision theories for unknown probabilities that deviate from expected utility. The implementation of such theories in decision analysis is explained by Borgonovo & Marinacci (2015).

Anscombe & Aumann (1963; AA henceforth) presented a two-stage model of uncertainty so as to obtain a simpler foundation of expected utility than Savage's.[1] Gilboa & Schmeidler (1989) and Schmeidler (1989) showed that the AA two-stage model is in fact well suited for another purpose: to analyze ambiguity theoretically. Since then, the AA model has become the most-used model for this alternative purpose.

The AA model makes two ancillary assumptions, explained below, that do not concern ambiguity. These assumptions have been justified on normative grounds. They fail descriptively though, as many studies cited later have shown. We will show how these ancillary assumptions can be relaxed to become descriptively valid while maintaining the convenient mixture operation of AA. We thus make the AA model suited for descriptive purposes while maintaining its analytical convenience and power. Now the many AA decision models of ambiguity introduced in the literature can be applied and tested descriptively while avoiding confounds due to violated ancillary assumptions. We call our modification the reduced AA (rAA) model. It is implemented in a simple experiment, where we find virtually all AA-based ambiguity models violated due to reference dependence. We then provide a reference dependent generalization of Schmeidler's (1989) Choquet expected utility to accommodate these violations.

In the AA model, acts are defined as maps from a set of states of nature to lotteries (probability distributions over outcomes). The decision maker first chooses an act. Then, in a first stage, nature chooses a state of nature, and in the second stage the corresponding lottery is played, resulting in an outcome. We discuss the following four assumptions that are commonly made: (1) Lotteries, being unambiguous, are evaluated using expected utility (EU); (2) backward induction (certainty equivalent substitution) is used to evaluate the two stages; (3) there is no reference dependence, with gains and losses treated the same; (4) there is universal ambiguity aversion.

[1] In fact, AA used a three-stage model but they assumed that it can be reduced to two stages and this is how their model is commonly used today.

Assumptions 1 and 2 capture the two-stage structure typical of AA. These two assumptions are not substantive but only ancillary, generating convenient linearity through probabilistic mixtures of consequences. Assumptions 3 and 4 concern ambiguity, i.e., the uncertainty about the states of nature, which is the focus of interest in the literature on AA models. These two assumptions are called *substantive*.

The descriptive and possibly normative problems of Assumption 1 have been widely acknowledged since the 1950s (Starmer 2000). Whereas Assumption 2 is natural under classical expected utility, it becomes problematic, both normatively and descriptively, under nonexpected utility, as has been understood since the end of the 1980s (Dominiak & Lefort 2011; Machina 1989; Sayman & Onculer 2009). An extra problem for the study of ambiguity is that the complexity of multi stages by itself induces a perception of ambiguity (Abdellaoui, Klibanoff, & Placido 2014; Halevy 2007; Maafi 2011). Further discussion of these problems and further references are in Appendix E. Because of these problems, analyses of ambiguity based on the AA two-stage model up to today are, surely for descriptive purposes, confounded by violations of the ancillary assumptions.

We will not enter normative debates. The purpose of this paper is descriptive. Section 2 explains the rAA model informally, showing how we can test AA theories without being bothered by violations of Assumptions 1 and 2. In particular, no two-stage uncertainty occurs in the rAA approach. An additional advantage is that complex stimuli, that can only be used hypothetically for normative purposes (Kreps 1988 p. 101), are avoided, reducing the burden for the subjects and the noise in the data. A formal analysis, showing model-theoretic isomorphism of the rAA model with the full AA model, is in Appendix E. This shows that no information is lost when going from the full AA model to the rAA model: everything done in full AA models can formally be redone in rAA models.

With the ancillary Assumptions 1 and 2 taken care of, substantive assumptions in the AA model can now for the first time be tested empirically. Section 3 provides a simple experiment doing so. We test the substantive Assumptions 3 and 4. Unsurprisingly, reference dependence, demonstrated in many decision fields outside of ambiguity and for decision under ambiguity outside of the AA model (reviewed by Trautmann & van de Kuilen 2015), also holds for ambiguity within the AA model. Losses are treated differently than gains, with more ambiguity seeking.[2] This reflects sign dependence, and shows that we need to incorporate reference dependence. Thus Assumptions 3 and 4 are violated.

Our experiment finds direct violations of the weak certainty independence condition, a con-

[2]See Abdellaoui, Vossmann, & Weber (2005), Baillon & Bleichrodt 2015; Chakravarty & Roy (2009), de Lara Resende & Wu (2010), Dimmock et al. (2015), Du & Budescu (2005), Hogarth & Einhorn (1990), and Ho, Keller, & Keltyka (2002).

dition assumed in most AA ambiguity theories today. Hence these theories are falsified, and reference dependent generalizations are called for. We turn to that in the next, theoretical, part of the paper. Following §4 with basic definitions and axiomatizations, §5 introduces a reference dependent generalization for the first modern ambiguity theory: Schmeidler's (1989) Choquet expected utility. The theoretical analyses of this paper assume general mixure spaces of outcomes, which include the classical AA models as a special case. The isomorphism between the classical AA model and the rAA model demonstrated in Appendix E implies that the outcome space of the rAA model also is a mixture space, and all the theorems of §4 and §5 directly apply to our rAA model.

Our generalization of Schmeidler's model can accommodate loss aversion, and ambiguity aversion for gains combined with ambiguity seeking for losses. Put differently, we show how the AA model can be extended to cover Tversky & Kahneman's (1992) prospect theory. In many applications of ambiguity (asset markets, insurance, health) the gain-loss distinction is important, and descriptive modelings that assume reference-independent universal ambiguity aversion will be flawed. As regards our finding of violations of weak certainty independence, reference dependence is the only generalization needed to accommodate these violations. Weak certainty independence remains satisfied if we restrict our attention to gains for instance. Our generalization of Schmeidler (1989) leads to a new concept: ambiguity-loss aversion (§6). Here loss aversion can be stronger (or weaker) under ambiguity than under risk, providing an additional way to induce ambiguity aversion. A discussion, with implications for existing ambiguity models, is in §7. Section 8 concludes.

We next discuss related literature. The smooth model of ambiguity (Klibanoff, Marinacci, & Mukerji 2005) and other utility-driven theories of ambiguity (Chew et al. 2008; Ju & Miao 2012; Nau 2006; Neilson 2010) can also treat losses differently than gains. Dobbs (1991) also proposed a general recursive utility-driven theory of ambiguity and emphasized the importance of different attitudes for gains than for losses, which he demonstrated in an experiment. He thus is close to our approach. Generalizations of other AA ambiguity theories to incorporate reference dependence, ambiguity seeking for particular choices, and other descriptive generalizations are a topic for future research.

An important insight underlying the modern behavioral aproach to decision theory is that, also for prescriptive applications of decision analysis, we need insights into descriptive deviations from normative models. First, such insights show where human decisions can be improved. Second, many measurements of subjective probabilities, discounting, utility, quality of life, and other attitudinal parameters needed as inputs in prescriptive decision analyses, are descriptive. We need to know about deviations and biases to properly reckon with them. This insight in itself has been understood in decision analysis from the beginning. For instance, Raiffa (1961) wrote: "We do not have to teach what comes naturally." Many decision analysts emphasized the importance of cross-checks in preference measurements (Keeney & Raiffa 1976 §5.8.3). Yet, descriptive and prescriptive insights are formally integrated only in the recent behavioral models. It has led to fruitful new ways to use decision analysis to improve human decisions (Thaler & Sunstein 2008).

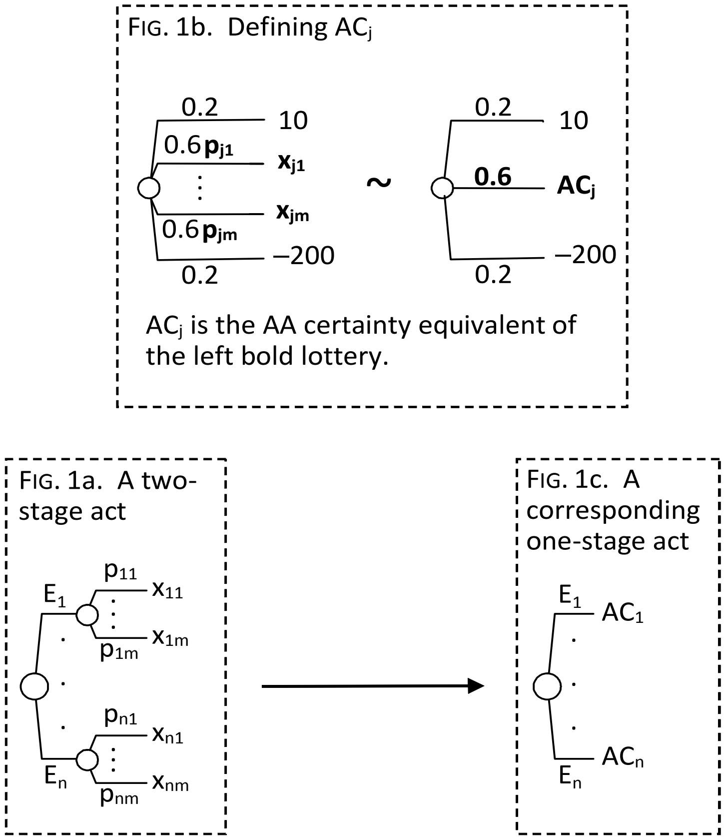
2 The reduced AA approach: constructing an AA-twin of the decision maker

This section explains the reduced AA (rAA) approach informally. Appendix E gives a formal presentation. We construct an AA-twin of the decision maker, who has the same substantive preferences over ambiguous acts but who, unlike the decision maker, satisfies the AA model. We infer preferences of the AA-twin from preferences of the decision maker using only choices where the two agree. These choices will not involve any two-stage acts. The inferred preferences of the twin will suffice to derive the whole preference relation of the twin. Inferences about the substantive (ambiguity) preferences of the twin that we then derive using AA techniques pertain to the decision maker herself because her substantive preferences agree with the twin's.

Fig. 1a depicts a typical two-stage act of the AA model. In the first stage, indicated by the left circle, one and exactly one of n events E_j obtains. Say a horse race takes place with n horses participating, and exactly one will win. Event E_j refers to horse j winning. In the second stage, a lottery over money is resolved. If E_j is true, then the lottery results that with probability p_{ji} gives outcome x_{ji}, with i ranging from 1 to m. To simplify notation, we take m independent of j, which can always be achieved by adding 0 probabilities.

We avoid the complexity of two stages for the decision maker by replacing every lottery in the second stage by a sure outcome, leading to the act in Fig. 1c. The new act has no more risk in the second stage and is one-stage. Every lottery has been replaced by an AA-certainty equivalent (CA) as defined in Fig. 1b. In that figure, we stayed away from degenerate lotteries and long shots by mixing in a 0.2 probability at a best outcome, 10 in our experiment, and a 0.2 probability at a worst outcome, −20 in our experiment. For the decision maker, the violations of the ancillary Assumptions 1 and 2 (where (2) plays no role for Fig. 1b) do not seriously affect the preferences over acts in Fig. 1b, and her preferences agree with those of the AA-twin. As regards the AA-twin, she satisfies EU and the mixing in of two other outcomes (10 and −20) does not affect her preferences. We obtain her regular certainty equivalent CA_j this way. If only few CAs

Figure 1: Relating a general two-stage act of the AA model to a one-stage act



Lotteries are substituted by their AC, as defined/measured in Fig. 1b. Lotteries are bold in Fig. 1b. (They do not involve the 0.6 probabilities.)

are to be measured, we can measure each separately as in Fig. 1b. If there are many, then we can carry out a few measurements as in Fig. 1b, derive the EU utility function from it, and use that to determine all CAs needed.

For the decision maker, possible violations of the ancillary AA Assumptions 1 and 2 do not affect preference between acts as in Fig. 1c. Her preferences agree with those of the twin there and, hence, we can find out the twin's preferences there. Because the twin satisfies the AA assumptions, the replacements by CAs does not affect her preferences. For her, the act in Fig. 1c is indifferent to the one in Fig. 1a. Preferences referring to acts in Fig. 1a can all be inferred from preferences between acts as in Fig. 1c. This way we obtain the twin's preferences between all two-stage AA acts. We have inferred preferences of the AA-twin from those of the decision maker using only Figs. 1b and 1c, where they agree, so that we have inferred genuine AA-twin preferences. They suffice to obtain the entire AA preference relation in the complete AA model. Hence we can use all AA techniques to analyze the substantive preferences of the AA-twin. Those substantive preferences relate to Fig. 1c. They all agree with the substantive preferences of the decision maker, so that we have learned about those.

3 Experimental illustration of the reduced AA model and reference dependence

This section presents a small experiment implementing the rAA approach. First, to prepare, we present a common example. The unit of payment in the example can be taken to be money or utility. In the experiment that follows after, the unit of payment will be utility and not money, so that the violations found there directly pertain to the general AA model. In general, the rAA model is a submodel of the entire AA model (large enough to recover the latter entirely). Hence, any violation of a preference condition found in the rAA model immediately gives a violation of that preference condition in the whole AA model.

EXAMPLE 1 [Reflection of ambiguity attitudes]. A known urn K contains 50 red (R) and 50 black (B) balls. An unknown (ambiguous) urn A contains 100 black and red balls in unknown proportion. One ball will be drawn at random from each urn, and its color will be inspected. R_k denotes the event of a red ball drawn from the known urn, and B_k, R_a, and B_a are analogous. People usually prefer to receive 100 under B_k (and 0 otherwise) rather than under B_a and they also prefer to receive 100 under R_k rather than under R_a. These choices reveal ambiguity aversion for gains.

We next multiply all outcomes by -1, turning them into losses. This change of sign can affect decision attitudes. Many people now prefer to lose 100 under B_a rather than under B_k and also to lose 100 under R_a rather than under R_k. That is, many people exhibit ambiguity seeking for losses. $\square$

The above example illustrates that ambiguity attitudes are different for gains than for losses, making it desirable to separate these, similarly as this has been found for risk (Baucells & Villasis 2010; Tversky & Kahneman 1992). This separation is impossible in most ambiguity models existing today. We tested the above choices in our experiment. Subjects were $N = 45$ undergraduate students from Tilburg University. We asked both for preferences with red the winning color and for preferences with black the winning color. This way we avoided suspicion about the experimenter rigging the composition of the unknown urn (Pulford 2009).

We scaled utility to be 0 at 0 and 100 at €10. That is, the winning amount was always €10. We wanted the loss outcome to be -100 in utility units for each subject, which required a different monetary outcome α for each subject. Thus, under EU as assumed in the AA model and as holding for the AA-twins of the subjects, we must have, with the usual notation for lotteries (probability distributions over money),

$$€0 \sim (0.5{:}€10, 0.5{:}€\alpha). \tag{1}$$

One simplifying notation for lotteries: we often rewrite $(p : \alpha, 1 - p : \beta)$ as $\alpha_p\beta$. The indifference displayed involves a degenerate (nonrisky) prospect (€0), and those are known to cause many violations of the assumed EU.[3] Hence we use the modification in Fig. 1b. We write $R = (€10_{0.5} - €20)$, and rather elicit the following indifference from our subjects, as in Fig. 1b, using the common probabilistic mixtures of lotteries, and mixing in $0.4{:}R$:

$$(R_{0.4}€0) \sim (R_{0.4}(€10_{0.5}€\alpha)) \tag{2}$$

Under EU as holding for the AA-twin, the latter indifference is equivalent to the former, but the latter indifference is less prone to violations of EU so here our subjects agree with their AA-twins.

To elicit the latter indifference from each subject, we asked each subject to choose between lotteries (replacing α in Eq. 2 by $-j$),

$$(0.2{:}€10, 0.6{:}€0, 0.2{:} - €20)\ (\text{“}safe\text{”})\ \mathit{and}\ (0.2{:}€10, 0.3{:}€10, 0.3{:} - €j, 0.2{:} - €20)\ (\text{“}risky\text{”})$$

for each $j = 0, 2, 4, \ldots, 18, 20$. If the subject switched from risky to safe between $-j$ and $-j-2$, then we defined α to be the midpoint between these two values, i.e., $\alpha = -j-1$. We then assumed

[3] See Bruhin, Fehr-Duda, & Epper (2010), Chateauneuf, Eichberger, & Grant (2007), Carthy et al. (1999 §2), McCord & de Neufville (1986).

indifference between the safe and risky prospect with that outcome α instead of $-j$ in the risky prospect. We used the monetary outcome α, depending on the subject, as the loss outcome for this subject. This way the loss outcome was -100 in utility units for each subject (as for their AA-twin).[4] Details of the experiment are in the web appendix.

We elicited the preferences of Example 1 from our subjects using utility units, with the gain outcome €10 giving utility $+100$, and the loss outcome α giving utility -100. Combining the bets on the two colors, the number of ambiguity averse choices was larger for gains than for losses (1.49 vs. 1.20, $z = 2.015$, $p < .05$, Wilcoxon test, two-sided), showing that ambiguity attitudes are different for gains than for losses. For gains we replicate strong ambiguity aversion ($z = 3.773$, $p < .01$, Wilcoxon test, two-sided), but for losses we cannot reject the null of ambiguity neutrality ($z = 1.567$, $p > .10$, Wilcoxon test, two-sided).[5] Our experiment confirms that attitudes towards ambiguity are different for gains than for losses, suggesting violations of most ambiguity models used today. The following sections will formalize this claim. The experiment has demonstrated how we can investigate the substantive (ambiguity) preferences of our subjects. We transfer to preferences of AA-twins in places where this can be done reliably, use AA techniques for the AA-twins to learn about their ambiguity attitudes, and transfer those back to our subjects.

4 Definitions, notation, classical expected utility, and Choquet expected utility for mixture spaces

This section provides definitions and surveys some well-known results. Proofs can be found in the source papers, or in the didactical Ryan (2009). We present our main theorems for general mixture spaces, which covers the traditional two-stage AA model but also some other models. By Observation 16 in the appendix, all results proved in the literature for the traditional two-stage AA model also hold for general mixture spaces.

M denotes a set of *consequences*, with generic elements x, y. M is a *mixture space*. That is, M is endowed with a mixture operation, assigning to all $x, y \in M$ and $p \in [0, 1]$ an element of M denoted $px + (1-p)y$ or $x_p y$. The following conditions define a *mixture operation*.

(i) $x_1 y = x$ [identity];

(ii) $x_p y = y_{1-p} x$ [commutativity];

(iii) $(x_p y)_q y = x_{pq} y$ [associativity].

[4] How our measurement of utility incorporates loss aversion under risk is discussed in §6.

[5] Testing is against the null of one ambiguity averse choice in two choice situations. The exact distribution of subjects choosing ambiguous never, once, or twice is $(28, 11, 6)$ for gains, and $(21, 12, 12)$ for losses.

The following two examples give the two commonly considered special cases of mixture spaces.

EXAMPLE 2 [*Two-stage AA framework*] D denotes a set of (deterministic) outcomes, and M consists of all *(roulette) lotteries*, which are probability distributions over D taking finitely many values. The mixture operation concerns probabilistic mixing. □

This example was popularized by Gilboa & Schmeidler (1989) and Schmeidler (1989), and is commonly used in the modern literature on ambiguity. We use the terminology of those two papers as much as possible.

EXAMPLE 3 $M = I\!R$ and mixing is the natural mixing of real numbers. □

Our rAA model will provide a third example (Appendix E). S denotes the state space. It is endowed with an algebra of subsets, called *events*. An *algebra* contains S and $\emptyset$ and is closed under complementation and finite unions and intersections. An *act* $f = (E_1{:}f_1, ..., E_n{:}f_n)$ takes values in M and the E_j's are events partitioning the state space. The set of acts, denoted $\mathcal{A}$, is endowed with pointwise mixing, which satisfies all conditions for mixture operations. Hence, $\mathcal{A}$ itself is also a mixture space. A *constant* act f assigns the same consequence $f(s) = x$ to all s. It is identified with this consequence.

Preferences are over the set of acts $\mathcal{A}$ and are denoted $\succcurlyeq$, inducing preferences $\succcurlyeq$ over consequences through constant acts. Strict preference $\succ$ and indifference $\sim$ are defined as usual. A function V *represents* $\succcurlyeq$ if $V : \mathcal{A} \to I\!R$ and $f \succcurlyeq g \Leftrightarrow V(f) \geq V(g)$. If a representing function exists then $\succcurlyeq$ is a *weak order*, i.e. $\succcurlyeq$ is *complete* (for all acts f and g, $f \succcurlyeq g$ or $g \succcurlyeq f$) and transitive. $\succcurlyeq$ is *nontrivial* if (not $f \sim g$) for some f and g in $\mathcal{A}$.

Continuity holds if, whenever $f \succ g$ and $g \succ h$, there are p and q in $(0, 1)$ such that $f_p h \succ g$ and $f_q h \prec g$. Hence continuity concerns the mixing of consequences. In the two-stage AA framework, continuity concerns probability. An *affine* function u on M satisfies $u(x_p y) = pu(x) + (1-p)u(y)$. In the two-stage AA framework, a function is affine if and only if it is EU (defined in Appendix E; it follows from substitution and induction).

Monotonicity holds if $f \succcurlyeq g$ whenever $f(s) \succcurlyeq g(s)$ for all s in S. Whereas it is uncontroversial for one-stage acts with degenerate monetary outcomes $f(s)$, it becomes nontrivial when the $f(s)$'s are nondegenerate lotteries as in Example 2. Monotonicity then implies that the decision maker's evaluation of $f(s)$, i.e. of f conditional on state s, is independent of what happens outside of s. This kind of separability may be undesirable for ambiguous states s, as pointed out by Jaffray. References and examples are in Wakker (2010 §10.7.3). Schneider & Schonger (2015) discuss the condition in detail and find it violated in an experiment. Monotonicity implies backward induction in Example 2; i.e., every $f(s)$ can be replaced by its certainty equivalent without affecting the preference of the decision maker. It is the ancillary Assumption 2 in the introduction. Several authors have argued against backward induction for nonexpected utility on normative grounds.[6] Many empirical studies have found it violated.[7] Dominiak & Schnedler (2011) tested Schmeidler's (1989) uncertainty aversion for two-stage acts, and found no clear results, or relations with Ellsberg-type ambiguity aversion. This can be taken as evidence against the descriptive usefulness of two-stage acts. A topic for future research is how their tests of uncertainty aversion will work out in the rAA model.

The following condition is the most important one in the axiomatization of affine representations and, hence, of EU.

DEFINITION 4 *Independence* holds on M if

$$x \succ y \Rightarrow x_p c \succ y_p c$$

for all $0 < p < 1$ and consequences x, y, and c. □

THEOREM 5 *[von Neumann-Morgenstern]. The following two statements are equivalent:*

(i) There exists an affine representation u on the consequence space M.

(ii) The preference relation $\succcurlyeq$ when restricted to M satisfies the following three conditions: (a) weak ordering; (b) continuity; (c) independence.

u is unique up to level and unit. □

Uniqueness of *u up to level and unit* means that another function u^* satisfies the same conditions as u if and only if $u^* = \tau + \sigma u$ for some real τ and positive σ. Affinity, independence, and Theorem 5 can be applied to any mixture set other than M, such as the set of acts $\mathcal{A}$. We next turn to two classic results.

Anscombe & Aumann's subjective expected utility. A *probability measure* P on S maps the events to $[0, 1]$ such that $P(\emptyset) = 0$, $P(S) = 1$, and P is *additive* ($P(E \cup F) = P(E) + P(F)$ for all disjoint events E and F). *Subjective expected utility* (*SEU*) holds if there exists a probability

[6] See Dominiak & Lefort (2011), Eichberger & Kelsey (1996), Karni & Safra (1990), Karni & Schmeidler (1991), Machina (1989), Ozdenoren & Peck (2008), Siniscalchi (2011). Wakker (2010, §10.7.3) criticized backward induction in the context of the AA model. Machina (2014 Example 3) also provides counter-evidence.

[7] See Barkan & Busemeyer (1999), Cubitt, Starmer, & Sugden (1998), Dominiak, Duersch, & Lefort (2012), and Yechiam et al. (2005).

measure P on S and a function u on M, such that $\succcurlyeq$ is represented by the function SEU defined as follows:

$$SEU : f \mapsto \int_S u(f(s))dP. \tag{3}$$

In the following theorem, independence on $\mathcal{A}$ implies independence on M through constant acts.

THEOREM 6 *[Anscombe & Aumann (1963)]. The following two statements are equivalent:*

(i) Subjective expected utility holds with a nonconstant affine u on M.

(ii) The preference relation $\succcurlyeq$ satisfies the following conditions: (a) nontrivial weak ordering; (b) continuity; (c) monotonicity; (d) independence.

The probabilities P on S are uniquely determined and u on M is unique up to level and unit. □

If we apply the above theorem to Example 3, then we obtain subjective expected value as in de Finetti (1937; Wakker 2010 Theorem 1.6.1). Thus two classical derivations of subjective probabilities, by Anscombe & Aumann (1963) and by de Finetti (1937), are based on the same underlying mathematics.

Schmeidler's Choquet Expected Utility. A *capacity* v on S maps events to $[0, 1]$, such that $v(\emptyset) = 0$, $v(S) = 1$, and $E \supset F \Rightarrow v(E) \geq v(F)$ (*set-monotonicity*). Unless stated otherwise we use a rank-ordered notation for acts $f = (E_1{:}x_1, \cdots, E_n{:}x_n)$, that is, $x_1 \succcurlyeq \cdots \succcurlyeq x_n$ is implicitly understood.

Let v be a capacity on S. Then, for any function w: $S \to \mathbb{R}$, the *Choquet integral* of w with respect to v, denoted $\int wdv$, is

$$\int_0^\infty v(\{s \in S : w(s) \geq \tau\})d\tau + \int_{-\infty}^0 [v(\{s \in S : w(s) \geq \tau\}) - 1]d\tau. \tag{4}$$

Choquet expect utility holds if there exist a capacity v and a function u on M such that preferences are represented by

$$CEU : f \mapsto \int_S u(f(s))dv. \tag{5}$$

Two acts f and g in $\mathcal{A}$ are *comonotonic* if for no s and t in S, $f(s) \succ f(t)$ and $g(s) \prec g(t)$. Thus any constant act is comonotonic with any other act. A set of acts is *comonotonic* if every pair of its elements is comonotonic.

DEFINITION 7 *Comonotonic independence* holds if

$$f \succ g \Rightarrow f_p c \succ g_p c$$

for all $0 < p < 1$ and comonotonic acts f, g, and c. □

Under comonotonic independence, preference is not affected by mixing with constant acts (consequences) (with some technical details added in Lemma 19). Because constant acts are comonotonic with each other, comonotonic independence on $\mathcal{A}$ still implies independence on M.

THEOREM 8 *[Schmeidler]. The following two statements are equivalent:*

(i) Choquet expected utility holds with nonconstant affine u on M;

(ii) The preference relation $\succcurlyeq$ satisfies the following conditions: (a) nontrivial weak ordering; continuity; (c) monotonicity; (d) comonotonic independence.

The capacity v on S is uniquely determined and u on M is unique up to level and unit. □

If we apply the above theorem to Example 3, then we obtain a derivation of Choquet expected utility with linear utility that is alternative to Chateauneuf (1991, Theorem 1). Cerreia-Vioglio, Maccheroni, & Marinacci (2015) provide a recent survey of applications.

Comonotonic independence implies a condition assumed by most models for ambiguity proposed in the literature.

DEFINITION 9 *Weak certainty independence* holds if

$$f_q x \succcurlyeq g_q x \Rightarrow f_q y \succcurlyeq g_q y$$

for all $0 < q < 1$, acts f, g, and all consequences x, y. □

That is, preference between two mixtures involving the same constant act x with the same weight $1 - q$ is not affected if x is replaced by another constant act y. This condition follows from comonotonic independence because both preferences between the mixtures should agree with the unmixed preference between f and g (again, with some technical details added in Lemma 19). Grant & Polak (2013) demonstrated that the condition can be interpreted as constant absolute uncertainty aversion: adding a constant to all utility levels will not affect preference. For a detailed analysis see Skiadas (2013).

5 Reference dependence in the AA model

Example 1 violates CEU: In the gain preference $100_{B_k}0 \succ 100_{B_a}0$, the best outcome (= consequence) 100 is preferred under B_k, implying $v(B_k) > v(B_a)$. In the loss preference $0_{B_a}(-100) \succcurlyeq 0_{B_k}(-100)$, the best outcome 0 is preferred under B_a, implying $v(B_a) \geq v(B_k)$. A contradiction has resulted. This reasoning does not use any assumption about (utility of) the outcomes 100,

-100, other than that they are of different signs (with $u(0) = 0$). For later purposes we show that even weak certainty independence is violated. In the proof of the following observation we essentially use the linear (probabilistic) mixing of outcomes typical of the AA models.

OBSERVATION 10 *Example 1 violates comonotonic independence, and even weak certainty independence.* □

Example 1 has confirmed for the AA model what many empirical studies have demonstrated for other models: ambiguity attitudes are different for gains than for losses, violating CEU and most other ambiguity models (reviewed by Trautmann & van de Kuilen 2015). Hence generalizations incorporating reference dependence are warranted. This section presents such a generalization. As in all main results, the analysis will be analogous to Schmeidler's analysis of rank dependence in Choquet expected utility as much as possible. Given this restriction, we stay as close as possible to the analysis of Tversky & Kahneman (1992).

In prospect theory there is a special role for a *reference point*, denoted θ. In our model it is a consequence that indicates a neutral level of preference. It often is the status quo of the decision maker. In Example 1, the deterministic outcome 0 was the reference point. Under the certainty equivalent condition in the AA model, we can always take a deterministic outcome as reference point. Sugden (2003) emphasized the interest of nondegenerate reference points. Many modern studies consider endogenous reference points that can vary (Baucells, Weber, & Welfens 2011; Köszegi & Rabin 2006). Our axiomatization will concern one fixed reference point. Extensions to variable reference points can be obtained by techniques as in Schmidt (2003).

Other consequences are evaluated relative to the reference point. A consequence $f(s)$ is a *gain* if $f(s) \succ \theta$, a *loss* if $f(s) \prec \theta$, and it is *neutral* if $f(s) \sim \theta$. An act f is *mixed* if there exist s and t in S such that $f(s) \succ \theta$ and $f(t) \prec \theta$. For an act f, the *gain part* f^+ has $f^+(s) = f(s)$ if $f(s) \succcurlyeq \theta$ and $f^+(s) = \theta$ if $f(s) \prec \theta$. The *loss part* f^- is defined similarly, where now all gains are replaced by the reference point. Prospect theory allows different ambiguity attitudes towards gains than towards losses. Hence we use two capacities, v^+ for gains and v^- for losses. For losses it turns out to be more natural to use a dual way of integration. We thus define the *dual* of v^-, denoted $\hat{v}^-$, by $\hat{v}^-(A) = 1 - v^-(A^c)$.

Prospect theory holds if there exist two capacities v^+ and v^- and a function U on consequences with $U(\theta) = 0$ such that $\succcurlyeq$ is represented by

$$PT : f \mapsto \int_S U(f^+(s))dv^+ + \int_S U(f^-(s))d\hat{v}^-. \tag{6}$$

We call U in Eq. 6 the *(overall) utility function.* There is a *basic utility* u, and a *loss aversion parameter* $\lambda > 0$, such that

$$U(x) = u(x) \text{ if } x \succ \theta \tag{7}$$

$$U(x) = u(x) = 0 \text{ if } x \sim \theta \tag{8}$$

$$U(x) = \lambda u(x) \text{ if } x \prec \theta. \tag{9}$$

For reasons explained later, we call λ the *ambiguity-loss aversion* parameter. Because $U(\theta) = 0$ we now add the scaling convention that also $u(\theta) = 0$. For identifying the separation of U into u and λ, further assumptions are needed. We will consider a new kind of separation based on the AA model and the mixture space setup of this paper. Wakker (2010 Chs. 8 and 12) discusses other separations in other models. The parameter λ is immaterial for preferences over consequences M, affecting neither preferences between gains or losses, nor within. Thus loss aversion in our model does not affect preferences over M (consequences), that is, over lotteries (risk) in the AA models. It only concerns ambiguity.

For later purposes we rewrite Eq. 6 as

$$PT = \sum_{j=1}^{n} \pi_j U(f(j)) \tag{10}$$

with decision weights π_j defined as follows. Assuming, for act $(E_1{:}x_1, ..., E_n{:}x_n)$, the rank-ordering $x_1 \succcurlyeq \cdots \succcurlyeq x_k \succcurlyeq \theta \succcurlyeq x_{k+1} \succcurlyeq \cdots \succcurlyeq x_n$. We define

$$\text{for } j \leq k: \quad \pi_j = \pi_j^+ \quad = \quad v^+\Big(\cup_{i=1}^{j} E_i\Big) - v^+\Big(\cup_{i=1}^{j-1} E_i\Big); \tag{11}$$

$$\text{for } j > k: \quad \pi_j = \pi_j^- \quad = \quad v^-\Big(\cup_{i=j}^{n} E_i\Big) - v^-\Big(\cup_{i=j+1}^{n} E_i\Big). \tag{12}$$

For gain events, the decision weight depends on cumulative events that yield better consequences. For loss events, the decision weight similarly depends on decumulative events that yield worse consequences. CEU analyzed in the preceding section is the special case of PT where v^- is the dual of v^+ and λ in Eq. 9 is 1.

We next turn to preference conditions that characterize prospect theory. We generalize comonotonicity by adapting a concept of Tversky & Kahneman (1992) to the present context. Two acts f and g are *cosigned* if they are comonotonic and if there exists no s in S such that $f(s) \succ \theta$ and $g(s) \prec \theta$. Note that, whereas for any act g and any constant act f, f is comonotonic with g, an analogous result need not hold for cosignedness. Only if the constant act is neutral, is it cosigned with every other act. This point complicates the proofs in the appendix. A set of acts is *cosigned* if every pair is cosigned. We next generalize comonotonic independence to allow reference dependence.

DEFINITION 11 *Cosigned independence* holds if

$$f \succ g \Rightarrow f_p c \succ g_p c$$

for all $0 < p < 1$ and cosigned acts f, g, and c. $\square$

$\succcurlyeq$ is *truly mixed* if there exists an act f with $f^+ \succ \theta$ and $\theta \succ f^-$. *Double matching* holds if, for all acts f and g, $f^+ \sim g^+$ and $f^- \sim g^-$ implies $f \sim g$. Now we are ready for the main theorem of this paper.

THEOREM 12 *Assume true mixedness. The following two statements are equivalent:*

(i) Prospect theory holds with U as in Eqs. 7-9.

(ii) The preference relation $\succcurlyeq$ satisfies the following conditions: (a) nontrivial weak ordering; (b) continuity; (c) monotonicity; (d) cosigned independence; (e) double matching.

The capacities are uniquely determined and the global utility function U is unique up to its unit. $\square$

We give the proof of the following observation in the main text because it is clarifying.

OBSERVATION 13 *Example 1 can be accommodated by prospect theory.*

PROOF. In Example 1, choose $v^+(B_k) > v^+(B_a)$, $v^+(R_k) > v^+(R_a)$, $v^-(B_k) > v^-(B_a)$, and $v^-(R_k) > v^-(R_a)$. Remember here that large values of v^- correspond with low values of its dual capacity as used in the Choquet integral. $\square$

We can take v^- different than v^+, letting the former accommodate ambiguity seeking in agreement with empirical evidence.

A number of new problems have to be resolved in the proof of Theorem 12. In the proof of Schmeidler's Theorem 8, constant acts are comonotonic with all acts, and serve to compare preferences across different comonotonic sets. In the proof of Theorem 12, however, gains are not cosigned with losses, and there is no direct way to compare preferences across different sign-comonotonic sets. We similarly lose the possibility to substitute comonotonic conditional certainty equivalents. A third problem is that the global utility function U is only piecewise linear in risky utility u, with a nonlinearity ("kink") at 0, under prospect theory.

Another, fourth, problem in the proof is that we do not get full-force independence on the mixture set, but we get it only separately for gains and losses. We show that this weakened condition still implies an affine representation (EU for risk). We display this generalization of the von Neumann-Morgenstern EU theorem because it is of interest on its own. It shows that independence can be weakened to cosigned independence in Theorem 5.

PROPOSITION 14 *For the preference relation $\succcurlyeq$ restricted to consequences, there exists an affine representation u if and only if $\succcurlyeq$ satisfies nontrivial weak ordering, continuity, and cosigned independence.*[8]

6 Measurements and interpretations of ambiguity loss aversion

This section considers a number of interpretations of the ambiguity-loss aversion parameter λ in Theorem 12 and Eqs. 7-9. Before turning to interpretations, it is useful to demonstrate how λ can be directly revealed from preference. This direct measurement is typical of the mixture (and AA) models, and cannot be used in other models.

OBSERVATION 15 *For all f in $\mathcal{A}$, $x^+, x^- \in M$, and $\lambda \in \mathbb{R}$, if $f \sim \theta$, $f^+ \sim x^+ \succ \theta$, and $f^- \sim x^- \prec \theta$, then $x^+_{\frac{1}{1+\lambda}} x^- \sim \theta$.* $\square$

In other words, with f, x^+, and x^- as in the observation, we find p such that $x^+_p x^- \sim \theta$, and then solve λ from $\frac{1}{1+\lambda} = p$ ($\lambda = \frac{1-p}{p}$). The condition in the theorem is intuitive: The indifference $x^+_{\frac{1}{1+\lambda}} x^- \sim \theta$ shows that, when mixing consequences (lotteries in the AA model), the loss must be weighted λ times more than the gain to obtain neutrality. Under ambiguity, however, f combines the preference values of x^+ and x^- in an "unweighted" manner (see the unweighted sum of the gain- and loss-part in Eq. 6), leading to the same neutrality level. Apparently, under ambiguity, losses are weighted λ times more than when mixing consequences (risk in the AA model). In the AA model, with consequences referring to lotteries and decision under risk, λ indicates how much *more* losses are overweighted under ambiguity than they are under risk. Thus λ purely reflects ambiguity attitude.

As mentioned before, the smooth ambiguity model (Klibanoff, Marinacci, & Mukerji 2005) can accommodate sign dependence of ambiguity attitudes. It can accommodate extra loss aversion due to ambiguity in the same way as our parameter λ does: through a kink of their second-order ambiguity-utility transformation function φ at 0.

For a first prediction on values of λ, we consider an extreme view on loss aversion for the AA models. It entails that all loss aversion will show up under risk, and that there can be expected to be no additional loss aversion due to ambiguity. This interpretation is most natural if loss aversion only reflects extra suffering experienced under losses, rather than an overweighting of losses without them bringing disproportional suffering when experienced. That is, this extreme interpretation ascribes loss aversion entirely to the (utility of) consequences. Then it is natural

[8] For consequences, this means that the independence condition in Definition 4) is restricted to cases where the consequences x, c, y are all better than the reference point or are all worse.

to predict that $\lambda = 1$, with no special role for ambiguity. We display the preference condition axiomatizating this prediction and showing how it can be tested:

Neutral ambiguity-loss aversion holds if $\lambda = 1$ in Observation 15.

A less extreme interpretation of ambiguity-loss aversion is as follows: There is loss aversion under risk, which can be measured in whatever is the best way provided in the literature.[9] For monetary prizes with a fixed reference point as considered in this paper, loss aversion will generate a kink of risky utility at that reference point. As an aside, in our model loss aversion under risk does not imply violations of expected utility and is fully compatible with our AA model, simply giving a kinked function u. Ambiguity can give extra loss aversion, and it can amplify ($\lambda > 1$) or moderate ($\lambda < 1$) it. The following preference condition characterizes λ:

Nonneutral ambiguity-loss aversion. For all f in $\mathcal{A}$, $x^+, x^- \in M$, and $\lambda \in \mathbb{R}$, if $f \sim \theta$, $f^+ \sim x^+ \succ \theta$, and $f^- \sim x^- \prec \theta$, then $x^+_{0.5}x^- \succ \theta$ if and only if $\lambda > 1$, and $x^+_{0.5}x^- \prec \theta$ if and only if $\lambda < 1$.

In the two-stage AA model, some consequences are outcomes and other consequences are lotteries. Reference dependence as analyzed in this paper takes lotteries as a whole, and their indifference class determines if they are gains or losses. This is analogous to the way in which Schmeidler (1989) modeled rank dependence in his model, which also concerned lotteries as a whole. Another approach can be considered, both for reference dependence and rank dependence, where outcomes within a lottery are perceived as gains or losses, and are weighted in a rank dependent manner. Here, as elsewhere, we followed Schmeidler's approach. For reference dependence, it was recommended by Tversky & Kahneman (1981, p. 456 penultimate para). In the rAA model, subjects are never required to perceive whole lotteries in a reference or rank dependent manner, but we implement it ourselves, and subjects only see the CAs that we inserted. Hence the above issue is no problem for us.

7 Discussion

Kreps (1988 p. 101) wrote about the non-descriptive nature of two-stage acts in the AA model: "imaginary objects. . . . makes perfectly good sense in normative applications . . . But this is a very

[9] There have been many discussions of ways to measure loss aversion under risk (Abdellaoui et al. 2015; Abdellaoui, Bleichrodt, & Paraschiv 2007). This debate is outside the scope of this paper.

dicy and perhaps completely useless procedure in descriptive applications. ... what sense does it make ... because the items concerned don't exist? I think we have to view the theory to follow [the traditional two-stage AA model] as being as close to purely normative as anything that we do in this book." We have shown how to increase the realism of the AA stimuli.

We next analyze to what extent we have succeeded in avoiding violations of EU. Because we always assign a nonnegligeable probability (0.2 in our experiment) to the best outcome and also to the worst outcome, for the preferences that we consider the nonlinear processing of probability typical of nonEU is only relevant in the middle of the domain, bounded away from $p = 0$ and $p = 1$. The common empirical finding is that deviates from linearity mostly occur at the boundaries (Starmer 2000; Tversky & Kahneman 1992).[10] Hence the deviations from EU will be weak for our stimuli. As explained before, loss aversion is incorporated in u, as a kink at zero.

Some other theoretical papers derived mixture operations on outcomes endogenously in a Savagean setup (with acts mapping states to outcomes) using bisymmetry axioms (surveyed by Baillon, Driesen, & Wakker 2012). As we did, they avoided two-stage acts and expected utility for risk, thus also avoiding AA's ancillary assumptions. Unlike us, they also avoided using objective probabilities. One drawback was that they could only observe mixtures for some fixed mixture weights such as 1/2. Thus, unlike us (Observation 23), they could not use the full richness of the AA model. A second drawback was that they needed to observe several indifferences to obtain one mixture (e.g. by involving many certainty equivalents) so that they do not have the tractability of the AA approach. Hence such techniques have not been applied in empirical studies. A third drawback is that these approaches cannot be extended to reference dependence because the use of certainty equivalents in bisymmetry axioms cannot be reconciled with sign dependence. Ghirardato et al. (2003) did obtain endogenous mixtures for all decision weights, but needed infinitely many indifferences to observe nondyadic mixture weights such as 1/3, which is again problematic for empirical applications.

Wakker (2010) used a Savagean setup avoiding the ancillary AA assumptions as did the aforementioned references, but used an endogenous difference rather than mixture operation. This approach can be reconciled with cosignedness, leading to an axiomatization of prospect theory (his Theorem 12.3.5). It however shares the first two aforementioned drawbacks.

The violation of CEU in Example 1 (see beginning of §5) concerned a violation of the elementary condition of weak certainty independence. Hence, every model implying this condition is violated the same way as CEU is. This concerns most ambiguity models considered in the

[10] As a technical point, if probability weighting is more or less steep in the interior for losses than for gains, then this can be captured by ambiguity-loss aversion.

literature.[11]

Further, the above violation involved only binary acts, implying that every model agreeing with CEU on this subdomain is violated too (Ghirardato & Marinacci 2001: biseparable preference; Luce 2000 Ch. 3: binary rank-dependent utility; tested by Choi et al. 2007). For all these models, it is desirable to develop reference-dependent generalizations.

8 Conclusion

Up to now, the AA ambiguity model could only be used for normative purposes (Kreps 1988 p. 101). We have made it suited for descriptive and, hence, prescriptive purposes. We demonstrated how the two major descriptive problems (violations of EU for risk and of backward induction) can be resolved, through a reduced AA model. An informal presentation and experiment showed how the reduced AA model can be implemented. In the experiment we benefitted from an additional advantage of the rAA: it only needs simple one-stage stimuli and those are easy to understand for subjects. A formal model-theoretic isomorphism shows that the reduced AA model maintains the full analytical power of the AA model. Our implementation, the first empirical test of the AA model not confounded by the violations of its ancillary assumptions, falsified an assumption of virtually all AA ambiguity models today: reference independence.

To accommodate the violation found, we introduced a reference dependent generalization of the first decision model of ambiguity that received a preference foundation: Schmeidler´s (1989) Choquet expected utility. Our generalization amounts to extending the AA model to prospect theory. We provided a preference foundation of our generalization. Topics for future research include the development of reference dependent generalizations of the many other ambiguity theories in the literature, and empirical tests of such models, as of our generalization of Schmeidler´s model. We hope that our paper will advance descriptive and prescriptive applications of ambiguity AA theories, having removed the major obstacles, and thus potentially doubling the impact of the AA model.

[11] See Chambers et al. (2014): dispersion aversion; Maccheroni, Marinacci, & Rustichini (2006): variational model; Siniscalchi (2009): vector theory; several multiple priors theories (Gajdos et al. 2008): contraction model; Chateauneuf 1991 and Gilboa & Schmeidler 1989: maxmin expected utility); Eichberger et al. (2011), Ghirardato, Maccheroni, & Marinacci (2004, also $\alpha(f)$ model), Grant & Polak (2013), Jaffray (1994): α-maxmin theory, Kopylov (2009; choice deferral), Strzalecki (2011; multiplier preferences). Exceptions are Cerreia-Vioglio et al. (2011a, b), Hayashi & Miao (2011), Ju & Miao (2012), Klibanoff, Marinacci, & Mukerji (2005), and Chateauneuf & Faro (2009). Hill (2013) also allows for violations of weak independence, but assumes that decreasing the worst outcome will increase ambiguity aversion, whereas in our experiment it was decreased.

Appendix. Proofs and an isomorphism

A Preparation

Several results in the ambiguity literature (e.g., Schmeidler 1989), were formulated for the two-stage AA model, and not for general mixture spaces as we use them. These results can routinely be transferred to acts for general mixture spaces. Rather than spell out details for specific theorems, do we explain the procedure in general and somewhat informally. In all our results, Theorem 5 (or Proposition 14) gives an affine representation u on M. In a general mixture space M, we can take two consequences $B \succ W$, and first focus on the consequences that are between B and W in preference. We replace these consequences by their u values (effectively, collapsing equivalence classes), endowing those with the natural mixture on real numbers. By monotonicity, this reduction maintains all preferentially relevant information. We now have the AA model with two outcomes B, W, can apply the results from the literature here, leading back to results for the original mixture space for consequences between B and W. We then extend the results to more and more extreme values $u(B)$ and $u(W)$, inductively covering the whole mixture space. We, again informally, summarize the reasoning.

OBSERVATION 16 *All cited preference foundations for AA models hold for general mixture spaces.* $\square$

B Proof of proposition 14: cosigned expected utility

A *nonloss* is a consequence that is a gain or is neutral, and a *nongain* is a consequence that is a loss or is neutral. We first derive a preparatory lemma.

LEMMA 17 *Assume that the preference relation $\succcurlyeq$, restricted to consequences, satisfies weak ordering, continuity, and cosigned independence. If x and y are nonlosses, then so are all $x_p y$ for $0 \le p \le 1$. If x and y are nongains, then so are all $x_p y$ for $0 \le p \le 1$.*

PROOF. Assume the conditions in the lemma. We consider the case of nonlosses x, y. Assume, for contradiction, $x_q y \prec \theta$ for some q. Continuity readily implies existence of a largest $p < q$ such that $x_p y \sim \theta$ and a smallest $r > q$ such $x_r y \sim \theta$. Define $x' = x_p y$ and $y' = x_r y$. Then x' and y' are neutral but, by continuity, every $x'_p y'$ must be a loss. The set of $x'_p y'$ ($0 \le' p \le 1$) is cosigned, implying that von Neumann-Morgenstern independence holds here without a cosignedness restriction. $x' \succ x'_{1/3} y'$ and independence imply that their $0.5-0.5$ mixture is strictly preferred to $x'_{1/3} y'$ (take $c = x'_{1/3} y'$ in the definition of cosigned independence), implyingthat $x'_{2/3}y' \succ x'_{1/3}y'$. In contradiction with this, $y' \succ x'_{2/3}y'$ and independence imply that their $0.5 - 0.5$ mixture is strictly preferred to $x'_{2/3}y'$, implying $x'_{1/3}y' \succ x'_{2/3}y'$. A contradiction has resulted. $\square$

We next turn to the proof of Proposition 14. Necessity of the preference conditions is obvious. We, hence, assume these preference conditions and derive an affine representation. We assume the *vNM axioms* (the axioms in Theorem 5) for $\succcurlyeq$ over consequences with, however, independence weakened to sign-independence: $x \succ y \Leftrightarrow x_p z \succ y_p z$ only if either all consequences are nonlosses or they all are nongains. By true mixedness, there exist consequences α and β with $\alpha \succ \theta \succ \beta$, and we will use these consequences in the following derivation.

Lemma 17 implies that the set of nonlosses is a mixture set (closed under mixing). On this set, all vNM axioms are satisfied, and an affine representing functional u^+ is obtained. We normalize $u^+(\theta) = 0, u^+(\alpha) = 1$. We will similarly obtain an affine u^- on nongains. To extend the representation and its affinity to mixed consequences, we define an as-if gain preference relation $\succcurlyeq^+$ over consequences, including losses, as follows. It agrees with $\succcurlyeq$ for gains, as we will see, and affinity extends it to losses: $x \succcurlyeq^+ y$ if there exists $p > 0$ such that $\alpha_p x \succcurlyeq \alpha_p y \succcurlyeq \theta$. We first show that the choice of p in the definition of $\succcurlyeq^+$ is immaterial.

LEMMA 18 *If $x \succcurlyeq^+ y$ then $\alpha_p x \succcurlyeq \alpha_p y$ for all $p > 0$ for which both mixtures are nonlosses.*

PROOF. Consider $\alpha_p x$, $\alpha_p y$, $\alpha_r x$, and $\alpha_r y$, and assume that all are nonlosses. Assume $p > r$. Then $\alpha_p x$ is a mixture of $\alpha_r x$ and α, and $\alpha_p y$ is a mixture of $\alpha_r y$ and α, where both mixtures use the same weights $((1-p)/(1-r)$ and $(p-r)/(1-r))$. All consequences involved in these mixtures are nonlosses by Lemma 17. By the affine representation for nonlosses, the preference between $\alpha_p x$ and $\alpha_p y$ is the same as between $\alpha_r x$ and $\alpha_r y$. $\square$

The above lemma shows that $\succcurlyeq^+$ indeed agrees with $\succcurlyeq$ for nonlosses (take $p = 1$). To see that it establishes an affine extension for losses, we briefly show that $\succcurlyeq^+$ satisfies all usual vNM axioms, also on losses. Completeness, transitivity, nontriviality, and independence all readily follow from the definition of $\succcurlyeq^+$ by taking a mixture weight p in its definition so close to 1 that this same mixture weight p can be used for all consequences concerned in the axioms. This also holds for continuity, where, applying it to $\succcurlyeq$ and $\alpha_p f$, $\alpha_p g$, and $\alpha_p h$ with p sufficiently close to 1, implies it for $\succcurlyeq^+$, f, g, and h. All vNM axioms are satisfied for $\succcurlyeq^+$, giving an affine representation, denoted u^+ of $\succcurlyeq^+$ and, hence, also of $\succcurlyeq$ on all nonlosses.

We similarly define an as-if loss preference relation: $x \succcurlyeq^- y$ if there exists $p > 0$ such that $\theta \succcurlyeq \alpha_p x \succcurlyeq \alpha_p y$. We similarly obtain an affine representation, denoted u^-, of $\succcurlyeq^-$ that agrees with

$\succcurlyeq$ for all nongains. u^+ and u^- both represent $\succcurlyeq$ on the set of neutral consequences. We show that this overlap is big enough to ensure that the two representations are identical.

We can set $u^+(\theta) = 0 = u^-(\theta)$. By continuity we can take $0 < p < 1$ such that $\alpha_p\beta \sim \theta$. Because u^- represents $\succcurlyeq$ for losses, $u^-(\beta) < u^-(\theta) = 0$, and hence $u^-(\alpha) > 0$. We normalize $u^-(\alpha) = u^+(\alpha) = 1$. Indifferences $\alpha_q\gamma \sim \theta$ for losses γ, and the affine representations, imply that $u^+ = u^-$ for losses γ. Thus, $u^+(\beta) = u^-(\beta)$. This and indifferences $\delta_r\beta \sim \theta$ imply that $u^+ = u^-$ for gains γ too. Hence $u^+ = u^-$ everywhere, and $u^+ = u^-$. Thus both these functions represent $\succcurlyeq$ on nonlosses and on nongains. They also represent preferences between gains and losses properly, assigning positive values to the former and negative values to the latter. Thus we have obtained an affine representation $u^+ = u^-$ of $\succcurlyeq$, implying all the vNM conditions for consequences without sign restrictions. We denote $u = u^+ = u^-$. This completes the proof of Proposition 14. □

C Proof of Theorem 12

We first show that the implications in the definitions of independence can be reversed. We use the term *strong (comonotonic/cosigned) independence*, to refer to these reversed versions.

LEMMA 19 *Assume that $\succcurlyeq$ is a continuous weak order. Then the reversed implications also hold.*

PROOF. Assume the conditions in the lemma, and the implication of the definition considered. Consider three acts f, g, h. If f, g, h are comonotonic (or cosigned), then so is the mixture set of all their mixtures, by Proposition 14. Hence, in each case, independence holds on the mixture set considered without a comonotonicity/cosignedness restriction, and we have the usual axioms that imply expected utility and the reversed implications of Lemma 19. □

NECESSITY OF THE PREFERENCE CONDITIONS IN THEOREM 12 ((i) implies (ii)). We assume (i), PT, and briefly indicate how cosigned independence is implied. The other conditions are routine. Consider cosigned f, g, c. We may assume a common partition $E_1, \ldots, E_n$ such that the consequences of the acts depend on these events. Because of cosignedness we can have

$$h_1 \succcurlyeq \ldots \succcurlyeq h_k \succcurlyeq \theta \succcurlyeq h_{k+1} \succcurlyeq \cdots \succcurlyeq h_n \tag{13}$$

for all h equal to f, g, or c, or a mixture of these acts. For example, if for j there exists a h' from $\{f, g, c\}$ with h'_j a gain, then all h_js are nonlosses and $j \leq k$. If $h_i \succ h_j$ for a h' from $\{f, g, c\}$, then $h_i \succcurlyeq h_j$ for all three acts, and $i < j$. Thus, we can use the same decision weights (Eqs. 11 and 12) for all three acts and for all their mixtures. It implies that $PT(f_pc) = pPT(f) + (1-p)PT(c)$, with the same equality for g instead of f. This implies cosigned independence. □

SUFFICIENCY OF THE PREFERENCE CONDITIONS IN THEOREM 12 ((ii) implies (i)). In Proposition 14 we derived expected utility for consequences if only cosigned independence is assumed. In agreement with the definition of prospect theory, we normalize expected utility for consequence θ such that $u(\theta) = 0$ and for some consequence (existing because of true mixedness) $\breve{\alpha} \succ \theta$ such that $u(\breve{\alpha}) = 1$. Let a *nonloss act* be an act g such that $g(s)$ is a nonloss for all s. A *nongain act* is defined similarly. By Lemma 17 the set of nonloss acts is closed under mixing, and so is the set of nongain acts. By Schmeidler's Theorem 8 there exists a CEU functional $CEU^+ = \int_S u(g^+(s))dv^+$ on the nonloss acts g^+ that represents $\succcurlyeq$ there.

By true mixedness, there exists a truly mixed act. By monotonicity, we can replace all nonloss consequences of the act by its maximal consequence, and all loss consequences by its minimal consequence, without affecting its true mixedness. The act now only has two consequences and can be written as $\gamma_F\beta$ with $\gamma \succ \theta \succ \beta$. ($\gamma$ abbreviates good (or gain) and β abbreviates bad.) By continuity, we may, and will, assume that $\gamma_F\beta \sim \theta$, by either improving (by mixing with θ) β or worsening (by mixing with θ) γ. $\gamma_F\beta$ will be used for calibrating the PT functional, and is called the *calibration act*.

We now define a functional PT^+ on nonloss acts and a functional PT^- on nongain acts, and a prospect theory functional PT that is the sum of those two. Next we show that PT represents preference. More precisely, we define

$$PT(f) = PT^+(f^+) + PT^-(f^-) = CEU^+(f^+) + \lambda CEU^-(f^-), \tag{14}$$

where $\lambda > 0$ is such that $PT(\gamma_F\beta) = 0$. Thus, $PT(\gamma_F\beta) = PT^+(\gamma_F\theta) + PT^-(\theta_F\beta)$, and $\lambda = -CEU^+(\gamma_F\theta)/CEU^-(\theta_F\beta)$. We define c as the PT value of the gain part of $\gamma_F\beta$; i.e.,

$$c = PT^+(\gamma_F\theta) > 0. \tag{15}$$

This c is minus the PT value of the loss part of $\gamma_F\beta$; i.e., $PT^-(\theta_F\beta) = -c$.

PT represents preference on all nonloss acts, and also on all nongain acts. Because it also compares nonloss acts properly with nongain acts (this holds in fact for every $\lambda > 0$), it is representing on the union of these, which is the set of all nonmixed acts. We call an act f *proper* if $PT(f) = PT(g)$ for some nonmixed act g with $f \sim g$. To prove that PT is representing it suffices, by transitivity, to show that all acts are proper, and this is what we will do. That is, we will use the nonmixed acts for calibrating PT relative to preferences. We start with a set of binary acts cosigned with the calibration act: $\mathcal{A}_F$ is defined as the set of all acts $\delta_F\alpha$ with $\delta \succcurlyeq \theta \succcurlyeq \alpha$.

LEMMA 20 *All acts in $\mathcal{A}_F$ are proper.*

PROOF. In this proof we only consider acts from $\mathcal{A}_F$. All these acts are cosigned, implying that we can use cosigned independence for all mixtures. We choose particular nonmixed acts. For any act f we find a *nonmixed equivalent* g defined as follows. Let x be a consequence such that with $g = x_F\theta$ we have $PT(g) = PT(f)$. By continuity of PT, such an x always exists. Thus g is a nonmixed binary act with the same PT value as f, but it is in $\mathcal{A}_F$ and is cosigned with f and θ. We will demonstrate properness on $\mathcal{A}_F$ by showing that each act is equivalent to a nonmixed equivalent.

CASE 1 [acts with PT value zero]: Let $PT(f) = 0$. Define $a = PT^+(f^+) = -PT^-(f^-) \geq 0$. θ is a nonmixed equivalent of f. We show that $f \sim \theta$.

CASE 1.1: $a \leq c$ (c as in Eq. 15). $PT^+(f^+) = \frac{a}{c}PT^+(\gamma_F\theta)$. By CEU for nonlosses, $f^+ \sim (\gamma_F\theta)_{a/c}\theta$. Similarly, $f^- \sim (\theta_F\beta)_{a/c}\theta$. By double matching, $f \sim (\gamma_{a/c}\theta)_F(\beta_{a/c}\theta) = (\gamma_F\beta)_{a/c}\theta \sim \theta$ (the last indifference by cosigned independence). By transitivity, $f \sim \theta$ and f is proper.

CASE 1.2: $a > c$. We consider a mix of f with θ, $f_p\theta$. From the definition of the PT functional we have $PT(f_p\theta) = pPT(f) = 0$ and $PT^+((f_p\theta)^+) = -PT^-((f_p\theta)^-) = pa$. We choose p so small that $0 < pa < c$. From case 1.1 we have $f_p\theta \sim \theta$. By strong cosigned independence, this implies $f \sim \theta$. f is proper.

CASE 2 [acts with positive PT value]: Let $PT(f) > 0$. By continuity and the definition of PT, there exists a consequence δ between θ and the maximal consequence in f such that $PT(\delta_F\theta) = PT(f) > 0$. $\delta_F\theta$ is a nonmixed equivalent of f. Define $a^+ = PT^+(f^+)$ and $a^- = -PT^-(f^-)$. Then $PT(\delta_F\theta) = a^+ - a^-$.

CASE 2.1: $a^+ \leq c$ (hence $a^- < c$). Write $b^+ = a^+/c$ and $b^- = a^-/c$.

$PT^+(f^+) = b^+PT^+(\gamma_F\theta)$ and $PT^-(f^-) = b^-PT^-(\theta_F\beta)$. Then it follows from CEU for gains that $f^+ \sim (\gamma_F\theta)_{b^+}\theta$. For the loss part of f we similarly have $f^- \sim (\theta_F\beta)_{b^-}\theta$. By double matching, $f \sim (\gamma_{b^+}\theta)_F(\beta_{b^-}\theta)$. We now isolate a symmetric component with absolute prospect theory value a^- for the gain part and the loss part (this was the hardest step to find in this paper): $(\gamma_{b^+}\theta)_F(\beta_{b^-}\theta) = (\gamma_F\beta)_{b^-}\left[(\gamma_{\frac{b^+-b^-}{1-b^-}}\theta)_F\theta\right] \sim (\theta)_{b^-}\left[(\gamma_{\frac{b^+-b^-}{1-b^-}}\theta)_F\theta\right] \sim (\gamma_{(b^+-b^-)}\theta)_F\theta = f^*$. From $PT(f^*) = c(b^+-b^-) = a^+-a^- = PT(\delta_F\theta)$ and CEU for nonlosses it follows that $f^* \sim \delta_F\theta$. By transitivity, $f \sim \delta_F\theta$. f is proper.

CASE 2.2: $a^+ > c$. We mix f and $\delta_F\theta$ with θ to obtain $f_\downarrow = f_p\theta$ and $(\delta_F\theta)_\downarrow = (\delta_F\theta)_p\theta$. We define $a^+_\downarrow = PT((f_p\theta)^+)$, which is pa^+, and $a^-_\downarrow = PT((f_p\theta)^-)$, which is pa^-. We choose p so small that $a^-_\downarrow < a^+_\downarrow < c$. From prospect theory we have $PT(f_\downarrow) = PT((\delta_F\theta)_\downarrow)$, which, by Case 2.1, implies $f_\downarrow \sim (\delta_F\theta)_\downarrow$. Because f, $\delta_F\theta$, and θ are cosigned, $f_p\theta \sim (\delta_F\theta)_p\theta$ implies $f \sim \delta_F\theta$. Again, f is proper.

Case 3 [Acts with negative PT value]: Let $PT(f) < 0$. This case is similar to Case 2.

We have demonstrated that all acts in $\mathcal{A}_F$ are proper. $\square$

We next show that all acts are proper. Consider a general act g, and event E such that g yields nonlosses on E and losses on E^c.

Case 1: There exists a *matching act* $f \in \mathcal{A}_F$ such that $PT^+(f^+) = PT^+(g^+)$ and $PT^-(f^-) = PT^-(g^-)$. Thus $PT(f) = PT(g)$, and from CEU for nonlosses and for nongains we have $f^+ \sim g^+$ and $f^- \sim g^-$. From double matching, $f \sim g$. Because f and g have the same PT value and are equivalent, and f is proper, it follows that g is also proper.

Case 2. There exists no matching act $f \in \mathcal{A}_F$ for g as in Case 1. We mix act g with θ to obtain an act $g_\downarrow = g_p\theta$. We choose p so small that we find a matching act $f_\downarrow \in \mathcal{A}_F$, i.e. $PT^+(f_\downarrow^+) = PT^+(g_\downarrow^+)$ and $PT^-(f_\downarrow^-) = PT^-(g_\downarrow^-)$. Thus $PT(f_\downarrow) = PT(g_\downarrow)$, and case 1 implies $f_\downarrow \sim g_\downarrow$.

Let $\tilde{g}$ be the nonmixed equivalent of g. Let $\tilde{g}_\downarrow$ similarly be the nonmixed equivalent of $g_\downarrow$. We have $PT(g_\downarrow) = PT(\tilde{g}_\downarrow)$, and because of Case 1 this implies $\tilde{g}_\downarrow \sim g_\downarrow$. Because g, $\tilde{g}$, and θ are cosigned, $g_p\theta \sim \tilde{g}_p\theta$ implies $g \sim \tilde{g}$. Thus, g is proper. We have proved sufficiency of the preference conditions.

Uniqueness Results Uniqueness of v^+ (v^-) follows from Schmeidler's Theorem 8 applied to nonloss (nongain) acts. It is obvious that the unit of utility can be multiplied by any positive constant. We show that no other change is possible. Restricting attention to nonloss consequences shows, by Schmeidler's theorem, that u when restricted to nonlosses is unique up to a unit, given that the scale $u(\theta) = 0$ is fixed. Similarly, restricting attention to nongain consequences shows that u when restricted to nongains is unique up to a unit that, a priori, might be different than for gains. However, the equivalence $\gamma_F\beta \sim \theta$ shows that the unit of losses is joined with that of gains, and a change of one implies the same change of the other. Hence only one unit of utility is free to choose. $\square$

D Remaining proofs

Proof of Observation 10.

$$(200_{B_k}0)_{\frac{1}{2}}(0) \succ (200_{B_a}0)_{\frac{1}{2}}(0) \tag{16}$$

$$\text{and} \tag{17}$$

$$(200_{B_k}0)_{\frac{1}{2}}(-200) \preccurlyeq (200_{B_a}0)_{\frac{1}{2}}(-200) \tag{18}$$

violate weak certainty independence, but are exactly the preferences in Example 1 under the AA model. To see the latter point, the left mixture in Eq. 16, for example, yields $200_{\frac{1}{2}}0 = 100$ under event B_k. $\square$

PROOF OF OBSERVATION 15. $PT(f) = CEU^+(f^+) + \lambda CEU^-(f^-) = u(\alpha^+) + \lambda u(\alpha^-) = PT(\theta) = 0$ implies $u(\alpha^+) = -\lambda u(\alpha^-)$. Then for $g = \alpha^+_p\alpha^- \sim \theta$, $PT(g) = pu(\alpha^+) + (1-p)u(\alpha^-) = p(-\lambda)u(\alpha^-) + (1-p)u(\alpha^-) = 0$, implying $-\lambda p + 1 - p = 0 \Leftrightarrow p = \frac{1}{1+\lambda}$. $\square$

E An isomorphism between our reduced and the full AA model

This appendix formally shows that the procedure described in §2, considering only a subdomain of the preferences (later formalized through an rAA model), is model-theoretically isomorphic to the complete AA preference model. It thus contains all information in the complete model, and we can indeed use all techniques of the full AA model despite our restricted domain. We first present the usual AA model. We assume Example 2. Generic notation: α, β, x_i, y_i for outcomes; $x = (p_1{:}x_1, \ldots, p_m{:}x_m)$, with the obvious interpretation, for lotteries. General acts are also called *(two-stage) acts* because there now are two stages of uncertainty and risk.

We now turn to the definition of the (two-stage) Anscombe-Aumann (AA) model. Preferences over lotteries induce preferences over outcomes through degenerate lotteries. We assume that there exist a *best* outcome B and a *worst* outcome W, with $B \succcurlyeq \alpha \succcurlyeq W$ for all outcomes α. These best and worst outcomes will simplify utility scalings and relations between different models. A *certainty equivalent* (*CE*) of a lottery is an outcome that is equivalent to that lottery. Under EU, it agrees with the CA of the lottery. The *certainty equivalent condition* means that there exists a unique certainty equivalent for each lottery. Uniqueness can always be achieved by collapsing indifference classes of outcomes. A function u on L is *expected utility (EU)* if $u((p_1{:}x_1, \ldots, p_m{:}x_m)) = \sum_{j=1}^{m} p_j u(x_j)$ and it represents $\succcurlyeq$ on L. This is equivalent fo affinity of u. We use the same symbol u for the function defined on X and its expectation defined on L. We sometimes call u on L the *risky utility function*.

DEFINITION 21 The *(two-stage) AA model* holds if a nontrivial and monotonic weak order $\succcurlyeq$ is given on the set $\mathcal{A}$ of acts, with a best outcome B and a worst outcome W, the CE condition satisfied, and with expected utility (u) holding on L. $\square$

An act f is *one-stage* if all lotteries $f(s)$ are degenerate; i.e., f assigns outcomes rather than nondegenerate lotteries to all states (upper panel in Figure 2). Then all relevant uncertainty has been resolved in the first stage. A lottery, identified with the corresponding constant act, is sometimes also called a *one-stage lottery* (left panel in Figure 2). Now all relevant uncertainty is resolved in the second stage.

The AA assumptions of EU on L and of monotonicity are called *ancillary assumptions*. They imply that a function representing preferences over acts must be of the form

$$V(EU \circ f) \tag{19}$$

with V nondecreasing.[12] The ancillary assumptions serve to simplify the analysis and are not of central interest in applications of the AA model today. Many further assumptions are studied in addition to the ancillary assumptions and the common assumptions of weak ordering and non-triviality. These further assumptions concern the function V, i.e. the aggregation of uncertainty over S, and they deal with ambiguity. They are of central interest, and they are what we call *substantive assumptions* in the AA model.

The experiment in §3 concerns the two-stage AA model with: (a) $S = \{R_a, B_a\}$; (b) bets on the ambiguous urn are acts; (c) bets on the known urn are fifty-fifty lotteries; (d) also other lotteries were used; (e) $B = 10$, $W = -20$. We used EU to analyze risky choices. We only used a subpart of the two-stage AA model, the rAA model, in two respects. First, all acts and lotteries presented to subjects were one-stage (upper and left panel in Figure 2). We obtained all desired utility levels at the second stage using consequences that are outcomes, i.e. degenerate lotteries. Yet we could indirectly infer about mixtures of consequences under the two-stage AA model if we wanted. For example, we could derive, for the AA twin,

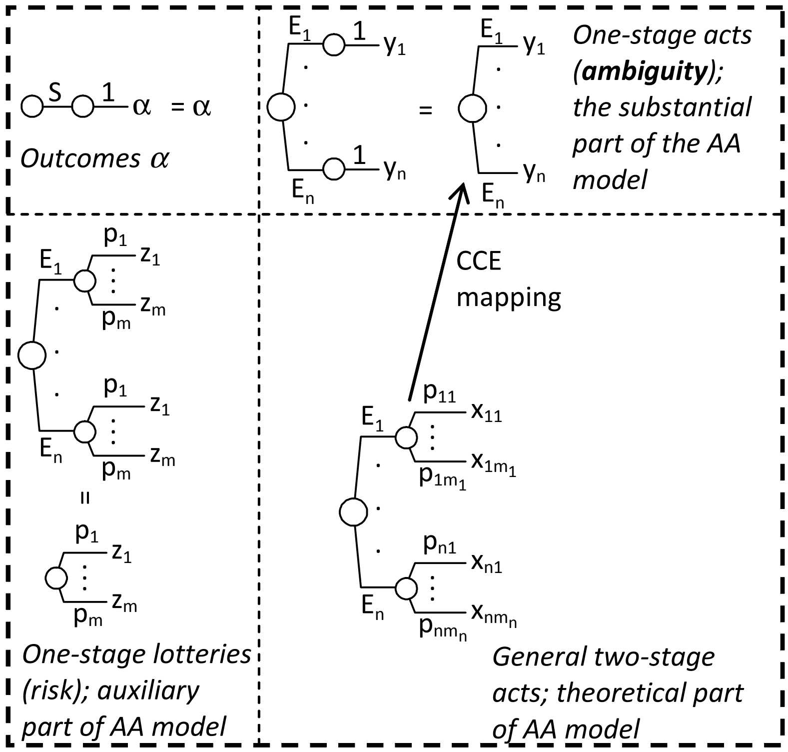
$$(R_a{:}0, B_a{:}0) \sim (R_a{:}10, B_a{:}0)_{0.5}(R_a{:}\alpha, B_a{:}0) = (R_a{:}(10_{0.5}\alpha), B_a{:}0), \tag{20}$$

because we elicited $CE(10_{0.5}\alpha) = 0$. The second respect in which we only used a subpart of the two-stage AA model is that we never used degenerate lotteries when eliciting risky preferences, avoiding the upper left box in Figure 2 here. We always incorporated the lottery $R = 10_{0.5}(-20)$ with probability 0.4 when eliciting risky preferences (Eq. 2).

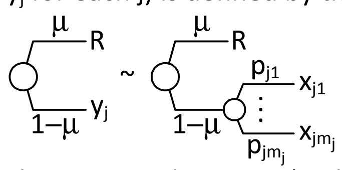
We next formalize our reduced form of the AA model, called the rAA model. The experiment in §3 suggested violations of the common ambiguity models. We in fact tested and falsified Schmeidler's (1989) comonotonic independence (in its weaker version of weak certainty independence). The rAA model sufficed for our test. This section formalizes the rAA model. The rAA model results from two modifications of the two-stage AA model.

[12] We first apply monotonicity with indifferences, to show that the representing function is a function V of $EU \circ f$, and then monotonicity in full force to show that V is nondecreasing.

Figure 2: The reduced AA model



The CCE mapping (value y_j for each j) is defined by the left-panel indifferences



The reduced AA model only concerns the upper (ambiguity) and left (risk) panel, and not the theoretical part. Of the risk panel we only use the subpart with a μ probability of lottery R, to stay away from the upper corner. Only this reduced model is used for empirical work, avoiding the extensive empirical violations in the theoretical part. Through the CCE mapping, the reduced AA model uniquely determines the theoretical part. Thus all theoretical features of the AA model, including its mixture operation, can be used, and all AA decision theories can be tested and applied empirically.

The first modification [focusing on one-stage acts and lotteries]. For each two-stage act f, each lottery $f(s)$ is replaced by $CE(f(s))$ (arrow in Figure 2). This replacement is not done by the decision maker, but by the researcher. All the decision maker does is express preferences over one-stage acts and one-stage lotteries, as in the experiment in §3. As this section will explain, the researcher then, for the purpose of using theories for the two-stage AA model, derives a corresponding two-stage AA model by implementing the ancillary assumptions by himself rather than assuming them on the part of the subjects. We thus need not assume that subjects behave as if replacing lotteries by CEs for risk, because we implement the replacement ourselves. We inferred Eq. 20 this way.

The rAA model preserves the underlying mixing of outcomes of the two-stage AA model, by temporarily returning, for a mix $\beta = \alpha_{'p}\gamma = p\alpha +' (1-p)\gamma$, to the underlying lotteries. We have added the primes before the mixing probability and after the plus symbol to indicate that this mixture operation is formally different from the (probabilistic) mixture operation in the two-stage AA model, although it will be isomorphic and is from the same indifference class. Informally, we take any lotteries x, z with $\alpha = CE(x), \gamma = CE(z)$, we take $y = x_p z$, and then get β as $CE(y)$. Because of EU on L, it does not matter which x and z we take in this process, and the operation is well defined. Always

$$\beta = \alpha_{'p}\gamma = p\alpha +'(1-p)\gamma = u^{-1}(pu(\alpha) + (1-p)u(\gamma)). \tag{21}$$

This holds irrespective of the particular choices x and z. We use Eq. 21 as the formal definition of the new mixture operation. The mixture operation is most easily observable from:

$$CE(\alpha_p\gamma) = \alpha_{'p}\gamma. \tag{22}$$

That is, we take $x = \alpha$ and $z = \gamma$. An example is Eq. 1 which showed that $0 = 10_{'0.5}\alpha$.

The second modification [avoiding degenerate lotteries for risky preferences]. In the first modification, we put deterministic outcomes central for the analysis of ambiguity by focusing on one-stage acts. Violations of EU for such acts, due to ambiguity, are our substantive interest. In the second modification considered now, concerning the analysis of risk through one-stage lotteries, we avoid degenerate lotteries, staying away from the upper left box in Figure 2. For such lotteries there are many violations of EU that distort our ancillary assumptions.

We define $R = B_{0.5}W$, and take some fixed $0 < \mu < 1$ (0.4 in the experiment in §3). For each lottery x, we define $x' = R_\mu x$. Under the ancillary AA assumptions, EU holds on L, and then a CE-indifference $\beta \sim y$ is not affected if we bring in μR, as in

$$R_\mu\beta \sim R_\mu y, \ i.e., \ \beta' \sim y'. \tag{23}$$

In general, indifferences are not affected under EU if we add or remove primes from all the lotteries. We call β in Eq. 23 the *conditional CE* of y, denoted $\beta = CA(y)$. We used this procedure in Eqs. 1 and 2. The CA condition means that there exists a unique CA for each lottery $x \in L$. Given existence, uniqueness can always be achieved by collapsing indifference classes.[13]

The two modifications combined. The rAA model results from combining the two modifications. Every two-stage act f in the two-stage AA model is replaced by $CA(f)$, defined by replacing every $f(s)$ by $CA(f(s))$, and turning every two-stage act into an equivalent one-stage act (arrow in Figure 2). We carried out the first modification, but with primes added to Eq. 22 because of the second modification.

All preferences between two-stage acts in the two-stage AA model can be recovered from their CA versions; we used this in Eq. 20. Thus, given the ancillary assumptions of AA, all substantive assumptions (concerning the upper panel in Figure 2) can be tested and analyzed using the rAA model.

We call the rAA model derived from the two-stage AA model as just described, the *corresponding* rAA model. Conversely, from every rAA model the uniquely determined corresponding two-stage AA model can be recovered, mostly by deriving preferences between two-stage acts from their CA images. We summarize the rAA model formally.

DEFINITION 22 The *reduced AA (rAA) model* holds if the following definitions and conditions are satisfied. Assume S, D, L as before. B and W are the best and worst outcomes, $R = W_{0.5}B$, and $0 < \mu < 1$ is fixed. For each lottery x, we have $x' = R_\mu x$, and $L' \subset L$ is the set of all lotteries x'. $\mathcal{OA}$ contains (a) all one-stage acts, and (b) L'. Thus all elements of $\mathcal{OA}$ are one-stage. Preferences $\succcurlyeq$ are defined only over $\mathcal{OA}$. Preferences over L' are represented by EU, the expectation u of a function on D also denoted u. Preferences over outcomes agree with those over constant acts, and are represented by u on D. Monotonicity holds. Conditional certainty equivalents, denoted CA, are defined as in Eq. 23, and are assumed to uniquely exist for every $x \in L$ (the CA condition).[14]

The mixture operation on outcomes is defined through Eq. 21, and can for instance be revealed from indifferences through the following analog of Eq. 22:

$$CA(\alpha_p\gamma) = \alpha'_p\gamma. \tag{24}$$

□

[13]Skiadas (2013) restricted the AA model to a fixed n-tuple of probabilities. If these all exceed gn and if constant roulette lotteries are excluded, then our second modification is satisfied.

[14]We can and do include all lotteries from L in this definition, also those not contained in L', because all preferences only involve elements of L'.

We summarize some useful relations between the corresponding reduced and two-stage AA models.

OBSERVATION 23 *There is a one-to-one correspondence between two-stage and reduced AA models (based on the maps $f \to CA(f)$ and $x \to x'$), and the preferences of one model uniquely determine those of the other. The rAA model is a substructure of the corresponding two-stage AA model, and its preferences agree with the restriction of the two-stage AA model preferences.* □

The agreement of the rAA model with the restrictions of the two-stage AA model makes it possible to present the experiment in §3 both as an application of the rAA model and of the two-stage AA model. Comparing the explanation of our experiment at the end of Appendix E with Definition 22 shows that this experiment in fact concerned the rAA model. Eq. 20 derived a preference in the corresponding two-stage AA model; i.e., it was an application of Observation 23. We used $CA(10_{0.5}\alpha) = 0$ (by Eq. 2) and the trivial $CA(0) = 0$ to obtain $CA(R_a{:}(10_{0.5}\alpha), B_a{:}0) = (R_a{:}0, B_a{:}0)$. The violations of classical ambiguity models that we found there (formalized later) thus are violations of these models both in terms of the rAA model and in terms of the two-stage AA model.

We recommend using only the rAA model for *descriptive* purposes, while relating it to the extensive *theoretical* literature on two-stage AA models through Observation 23. Thus, whereas two-stage acts may appear as ancillary tools in proofs of mathematical theorems on V in Eq. 19, subjects are never exposed to such complex stimuli. Our recommendation explains, in formal terms, how the full literature on two-stage AA models can be used for descriptive purposes to study the function V in Eq. 19 without paying the descriptive price of the ancillary assumptions 1 (EU for risk) and 2 (backward induction) that relate to the part "$EU\circ$" in Eq. 19). We have shown that the lower right part in Figure 2 is redundant. A pragmatic objection can be raised against the rAA model. The mixture operation of outcomes is less easy to implement than in the original AA model. Now a mixture is not done by just multiplying probabilities, but it requires observing an indifference. But such observations are easy to obtain, as our experiment demonstrated. They concern stimuli that are considerably easier to understand for subjects than two-stage acts.

The objection just raised can be rephrased in a methodological way. The mixture operation in the rAA model is a derived concept. The purpose of behavioral foundations is to give conditions directly in terms of the empirical primitive, being the preference relation. Derived concepts can be used only if the resulting preference conditions can easily be (re)formulated directly in terms of primitives. Our ancillary CE mixture operation uses objective probabilities that are directly available and then certainty equivalents (CAs) that are easy to observe. The mixture operation can be implemented empirically, as demonstrated in §3. Facilitating experimental testing of an AA model was the primary motivation for us to introduce rAA models.

F Web Appendix

Web Appendix: see
http:\people.few.eur.nl/wakker/pdf/schmeidler1989.pt.webappendix.pdf

ACKNOWLEDGMENT. Horst Zank made useful comments.

References

Abdellaoui, Mohammed, Han Bleichrodt, & Corina Paraschiv (2007) "Loss Aversion under Prospect Theory: A Parameter-Free Measurement," *Management Science* 53, 1659–1674.

Abdellaoui, Mohammed, Peter Klibanoff, & Laetitia Placido (2014) "Experiments on Compound Risk in Relation to Simple Risk and to Ambiguity," *Management Science*, forthcoming.

Abdellaoui, Mohammed, Frank Vossmann, & Martin Weber (2005) "Choice-Based Elicitation and Decomposition of Decision Weights for Gains and Losses under Uncertainty," *Management Science* 51, 1384–1399.

Anscombe, Francis J. & Robert J. Aumann (1963) "A Definition of Subjective Probability," *Annals of Mathematical Statistics* 34, 199–205.

Baillon, Aurélien & Han Bleichrodt (2015) "Testing Ambiguity Models through the Measurement of Probabilities for Gains and Losses," *American Economic Journal: Microeconomics* 7, 77–100.

Baillon, Aurélien, Bram Driesen, & Peter P. Wakker (2012) "Relative Concave Utility for Risk and Ambiguity, " *Games and Economic Behavior* 75, 481–489.

Barkan, Rachel & Jeromy R. Busemeyer (1999) "Changing Plans: Dynamic Inconsistency and the Effect of Experience on the Reference Point," *Psychonomic Bulletin and Review* 6, 547–554.

Baucells, Manel & Antonio Villasis (2010) "Stability of Risk Preferences and the Reflection Effect of Prospect Theory," *Theory and Decision* 63, 193–211.

Baucells, Manel, Martin Weber, & Frank Welfens (2011) "Reference-Point Formation and Updating," *Management Science* 57, 506–519.

Borgonovo, Emanuele & Massimo Marinacci (2015) "Decision Analysis under Ambiguity," *European Journal of Operational Research* 244, 823–836.

Bruhin, Adrian, Helga Fehr-Duda, & Thomas Epper (2010) "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica* 78, 1375–1412.

Carthy, Trevor, Susan Chilton, Judith Covey, Lorraine Hopkins, Michael Jones-Lee, Graham Loomes, Nick Pidgeon, & Anne Spencer (1999) "On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 2—The CV/SG "Chained" Approach," *Journal of Risk and Uncertainty* 17, 187–213.

Cerreia-Vioglio, Simone, Paolo Ghirardato, Fabio Maccheroni, Massimo Marinacci & Marciano Siniscalchi (2011a) "Rational Preferences under Ambiguity," *Economic Theory* 48, 341–375.

Cerreia-Vioglio, Simone, Fabio Maccheroni, Massimo Marinacci, & Luigi Montrucchio (2011b) "Uncertainty Averse Preferences," *Journal of Economic Theory* 146, 1275–1330.

Chakravarty, Sujoy & Jaideep Roy (2009) "Recursive Expected Utility and the Separation of Attitudes towards Risk and Ambiguity: An Experimental Study," *Theory and Decision* 66, 199–228.

Chambers, Robert G., Simon Grant, Ben Polak, & John Quiggin (2014) "A Two-Parameter Model of Dispersion Aversion," *Journal of Economic Theory* 150, 611–641.

Chateauneuf, Alain (1991) "On the Use of Capacities in Modeling Uncertainty Aversion and Risk Aversion," *Journal of Mathematical Economics* 20, 343–369.

Chateauneuf, Alain, Jürgen Eichberger, & Simon Grant (2007) "Choice under Uncertainty with the Best and Worst in Mind: NEO-Additive Capacities," *Journal of Economic Theory* 137, 538–567.

Chateauneuf, Alain & José Heleno Faro (2009) "Ambiguity through Confidence Functions," *Journal of Mathematical Economics* 45, 535–558.

Chew, Soo Hong, King King Li, Robin Chark, & Songfa Zhong (2008) "Source Preference and Ambiguity Aversion: Models and Evidence from Behavioral and Neuroimaging Experiments." *In* Daniel Houser & Kevin McGabe (eds.) *Neuroeconomics. Advances in Health Economics and Health Services Research* 20, 179–201, JAI Press, Bingley, UK.

Choi, Syngjoo, Raymond Fishman, Douglas Gale, & Shachar Kariv (2007) "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review* 97, 1921–1938.

Cubitt, P. Robin, Chris Starmer, & Robert Sugden (1998) "Dynamic Choice and the Common Ratio Effect: An Experimental Investigation," *Economic Journal* 108, 1362–1380.

de Finetti, Bruno (1937) "La Prévision: Ses Lois Logiques, ses Sources Subjectives," *Annales de l'Institut Henri Poincaré* 7, 1–68. Translated into English by Henry E. Kyburg, "Foresight: Its Logical Laws, its Subjective Sources." *In* Henry E. Kyburg & Howard E. Smokler (1964, eds) *Studies in Subjective Probability*, Wiley, New York, 53–118; 2nd edition 1980, Krieger, New York.

de Lara Resende, José Guilherme & George Wu (2010) "Competence Effects for Choices Involving Gains and Losses," *Journal of Risk and Uncertainty* 40, 109–132.

Dimmock, Stephen G., Roy Kouwenberg, Olivia S. Mitchell, & Kim Peijnenburg (2015) "Estimating Ambiguity Preferences and Perceptions in Multiple Prior Models: Evidence from the Field," *Journal of Risk and Uncertainty*, forthcoming.

Dobbs, Ian M. (1991) "A Bayesian Approach to Decision-Making under Ambiguity," *Economica* 58, 417–440.

Dominiak, Adam, Peter Duersch, & Jean-Philippe Lefort (2012) "A Dynamic Ellsberg Urn Experiment," *Games and Economic Behavior* 75, 625–638.

Dominiak, Adam & Jean-Philippe Lefort (2011) "Unambiguous Events and Dynamic Choquet Preferences," *Economic Theory* 46, 401–425.

Dominiak, Adam & Wendelin Schnedler (2011) "Attitudes toward Uncertainty and Randomization: An Experimental Study," *Economic Theory* 48, 289–312.

Du, Ning & David Budescu (2005) "The Effects of Imprecise Probabilities and Outcomes in Evaluating Investment Options," *Management Science* 51, 1791–1803.

Eichberger, Jürgen & David Kelsey (1996) "Uncertainty Aversion and Dynamic Consistency," *International Economic Review* 37, 625–640.

Eichberger, Jürgen, Simon Grant, David Kelsey, & Gleb A. Koshevoy (2011) "The Alpha-Meu Model: A Comment," *Journal of Economic Theory* 48, 1684–1698.

Ellsberg, Daniel (1961) "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics* 75, 643–669.

Gajdos, Thibault, Takashi Hayashi, Jean-Marc Tallon, & Jean-Christophe Vergnaud (2008) "Attitude towards Imprecise Information," *Journal of Economic Theory* 140, 27–65.

Ghirardato, Paolo, Fabio Maccheroni, & Massimo Marinacci (2004) "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory* 118, 133–173.

Ghirardato, Paolo & Massimo Marinacci (2001) "Risk, Ambiguity, and the Separation of Utility and Beliefs," *Mathematics of Operations Research* 26, 864–890.

Ghirardato, Paolo, Fabio Maccheroni, Massimo Marinacci, & Marciano Siniscalchi (2003) "A Subjective Spin on Roulette Wheels," *Econometrica* 71, 1897–1908.

Gilboa, Itzhak & David Schmeidler (1989) "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics* 18, 141–153.

Grant, Simon & Ben Polak (2013) "Mean-Dispersion Preferences and Constant Absolute Uncertainty Aversion," *Journal of Economic Theory* 148, 1361–1398.

Halevy, Yoram (2007) "Ellsberg Revisited: An Experimental Study," *Econometrica* 75, 503–536.

Hayashi, Takashi & Jianjun Miao (2011) "Intertemporal Substitution and Recursive Smooth Ambiguity Preferences," *Theoretical Economics* 6, 423–475.

Hill, Brian (2013) "Confidence and Decision," *Games and Economic Behavior* 82, 675–692.

Ho, Joanna L.Y., L. Robin Keller, & Pamela Keltyka (2001) "Managers' Variance Investigation Decisions: An Experimental Examination of Probabilistic and Outcome Ambiguity," *Journal of Behavioral Decision Making* 14, 257–278.

Hogarth, Robin M. & Hillel J. Einhorn (1990) "Venture Theory: A Model of Decision Weights," *Management Science* 36, 780–803.

Jaffray, Jean-Yves (1994) "Dynamic Decision Making with Belief Functions." *In* Ronald R. Yager, Mario Fedrizzi, & Janus Kacprzyk (eds.) *Advances in the Dempster-Shafer Theory of Evidence,* 331–352, Wiley, New York.

Ju, Nengjiu & Jianjun Miao (2012) "Ambiguity, Learning, and Asset Returns," *Econometrica* 80, 559–591.

Karni, Edi & Zvi Safra (1990) "Behaviorally Consistent Optimal Stopping Rules," *Journal of Economic Theory* 51, 391–402.

Karni, Edi & David Schmeidler (1991) "Atemporal Dynamic Consistency and Expected Utility Theory," *Journal of Economic Theory* 54, 401–408.

Keeney, Ralph L. & Howard Raiffa (1976) "*Decisions with Multiple Objectives*." Wiley, New York. (Second edition 1993, Cambridge University Press, Cambridge).

Keynes, John Maynard (1921) "*A Treatise on Probability.*" McMillan, London. Second edition 1948.

Klibanoff, Peter, Massimo Marinacci, & Sujoy Mukerji (2005) "A Smooth Model of Decision Making under Ambiguity," *Econometrica* 73, 1849–1892.

Knight, Frank H. (1921) "*Risk, Uncertainty, and Profit.*" Houghton Mifflin, New York.

Kopylov, Igor (2009) "Choice Deferral and Ambiguity Aversion," *Theoretical Economics* 4, 199–225.

Köszegi, Botond & Matthew Rabin (2006) "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics* 121, 1133–1165.

Kreps, David M. (1988) "*Notes on the Theory of Choice.*" Westview Press, Boulder Colorada.

Luce, R. Duncan (2000) "*Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches.*" Lawrence Erlbaum Publishers, London.

Maafi, Hela (2011) "Preference Reversals under Ambiguity," *Management Science* 57, 2054–2066.

Maccheroni, Fabio, Massimo Marinacci, & Aldo Rustichini (2006) "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica* 74, 1447–1498.

Machina, Mark J. (1989) "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty," *Journal of Economic Literature* 27, 1622–1688.

Machina, Mark J. (2014) "Ambiguity Aversion with Three or More Outcomes," *American Economic Review* 104, 3814–3840.

McCord, Mark R. & Richard de Neufville (1986) ""Lottery Equivalents": Reduction of the Certainty Effect Problem in Utility Assessment," *Management Science* 32, 56–60.

Nau, Robert F. (2006) "Uncertainty Aversion with Second-Order Utilities and Probabilities," *Management Science* 52, 136–145.

Neilson, William S. (2010) "A Simplified Axiomatic Approach to Ambiguity Aversion," *Journal of Risk and Uncertainty* 41, 113–124.

Ozdenoren, Emre & James Peck (2008) "Ambiguity Aversion, Games against Nature, and Dynamic Consistency," *Games and Economic Behavior* 62, 106–115.

Pulford, Briony D. (2009) "Is Luck on My Side? Optimism, Pessimism, and Ambiguity Aversion," *Quarterly Journal of Experimental Psychology* 62, 1079–1087.

Raiffa, Howard (1961) "Risk, Uncertainty and the Savage Axioms: Comment," *Quarterly Journal of Economics* 75, 690–694.

Ryan, Matthew J. (2009) "Generalizations of SEU: A Geometric Tour of Some Non-Standard Models," *Oxford Economic Papers* 61, 327–354.

Savage, Leonard J. (1954) "*The Foundations of Statistics*." Wiley, New York. (Second edition 1972, Dover Publications, New York.)

Sayman, Serdar & Ayşe Öncüler (2009) "An Investigation of Time-Inconsistency," *Management Science* 55, 470–482.

Schmeidler, David (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica* 57, 571–587.

Schmidt, Ulrich (2003) "Reference-Dependence in Cumulative Prospect Theory," *Journal of Mathematical Psychology* 47, 122–131.

Schneider, Florian Schneider & Martin Schonger (2015) "Allais at the Horse Race: Testing Models of Ambiguity Aversion," mimeo.

Siniscalchi, Marciano (2009) "Vector Expected Utility and Attitudes toward Variation," *Econometrica* 77, 801–855.

Siniscalchi, Marciano (2011) "Dynamic Choice under Ambiguity," *Theoretical Economics* 6, 379–421.

Skiadas, Costis (2013) "Scale-Invariant Uncertainty-Averse Preferences and Source-Dependent Constant Relative Risk Aversion," *Theoretical Economics* 8, 59–93.

Starmer, Chris (2000) "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk," *Journal of Economic Literature* 38, 332–382.

Strzalecki, Tomasz (2011) "Axiomatic Foundations of Multiplier Preferences," *Econometrica* 79, 47–73.

Sugden, Robert (2003) "Reference-Dependent Subjective Expected Utility," *Journal of Economic Theory* 111, 172–191.

Thaler, Richard H. & Cass R. Sunstein (2008) *"Nudge: Improving Decisions About Health, Wealth, and Happiness."* Yale University Press, New Haven.

Trautmann, Stefan T. & Gijs van de Kuilen (2015) "Ambiguity Attitudes." *In* Gideon Keren & George Wu (eds) *Handbook of Judgment and Decision Making* , forthcoming, Blackwell, Oxford.

Tversky, Amos & Daniel Kahneman (1981) "The Framing of Decisions and the Psychology of Choice," *Science* 211, 453–458.

Tversky, Amos & Daniel Kahneman (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297–323.

Wakker, Peter P. (2010) "*Prospect Theory: for Risk and Ambiguity.*" Cambridge University Press, Cambridge, UK.

Yechiam, Eldad, Julie C. Stout, Jerome R. Busemeyer, Stephanie L. Rock, & Peter R. Finn (2005) "Individual Differences in the Response to Forgone Payoffs: An Examination of High Functioning Drug Abusers," *Journal of Behavioral Decision Making* 18, 97–110.