Anticipated Regret as an Explanation of Uncertainty Aversion*

Daniel Krähmer† and Rebecca Stone‡

September 21, 2010

Abstract

The paper provides a psychological explanation of uncertainty aversion based on the fear of regret. We capture an agent's regret using a reference-dependent utility function in which the agent's utility depends on the performance of his chosen alternative relative to the performance of the alternative that would have been best ex post. An uncertain alternative is represented as a compound lottery. The basic idea is that selecting a compound lottery reveals information, which alters the ex post assessment of what the best choice would have been, inducing regret. We provide sufficient conditions under which regret implies uncertainty aversion in the sense of quasi-concave preferences over compound lotteries.

Keywords: Regret, Uncertainty Aversion, Reference Dependence, Counterfactual Reasoning, Information Aversion, Hindsight Bias

JEL Classification: C72, D11, D81, D83

*Aknowledgements: We would like to thank Florian Englmaier, Paul Heidhues, Rudolf Kerschbamer, Frank Riedel, and Karl Schlag for helpful suggestions and discussions. We are also greatful to seminar audiences in Bonn and Innsbruck for useful comments.

†Universität Bonn, Department of Economics and Hausdorff-Center for Mathematics, Adenauer Allee 24-42, 53113 Bonn, Germany, ++49-(0)228-736147, kraehmer@hcm.uni-bonn.de. Financial support through research grant SFB/TR 15 is gratefully acknowledged.

‡New York University School of Law, New York, NY 10012, rs2647@nyu.edu.

1 Introduction

Uncertainty aversion is a puzzling phenomenon. Why are people especially averse to uncertainty that they can't quantify? Why don't they simply assign a prior probability distribution over the possible probabilities, perhaps in accordance with the principle of insufficient reason, and thereby assimilate the uncertainty to risk? In this paper, we offer an explanation of uncertainty aversion grounded in the familiar psychological phenomenon of regret.

Regret is the negative emotion that arises from an agent's perception that he should have chosen otherwise. Uncertainty aversion refers to an agent's distaste for making choices under conditions of uncertainty in which some relevant probabilities are unknown ex ante as opposed to conditions of risk in which all relevant probabilities are known in advance. Crucial for our argument is the observation that the resolution of an uncertain lottery reveals information about the unknown probabilities of the events associated with that lottery. Thus, when an agent chooses such a lottery its resolution may alter his ex post assessment of what he optimally should have done at the moment of choice, possibly leading him to regret his past choice. By contrast, the resolution of a risky prospect reveals no such information and so leaves the agent's original assessment of the wisdom of his past choices in tact.

To illustrate, consider the two-urn example presented by Ellsberg (1961). One of the urns, the "risky urn", contains 50 black and 50 white balls. The other, the "uncertain urn", contains 100 balls that are either black or white in unknown proportions. The agent has to select the urn from which a ball is to be drawn, knowing that he will receive \$100 if a black ball is drawn from it and \$0 if a white ball is drawn. Suppose the agent represents the uncertain urn as a two-stage compound lottery in which the composition of the urn—the state—is first selected, and then a ball is drawn, conditional on the state. Suppose further that the agent adheres to the principle of insufficient reason, assigning equal probability to each of the 101 possible configurations of the uncertain urn. Thus, he perceives that the two urns yield the same probability distribution over monetary outcomes. The agent is uncertainty averse if he strictly prefers the risky urn.[1]

We model regret by assuming that the agent evaluates his past decisions in light of his ex post information. More precisely, after the ball has been drawn, the agent compares his actual payoff with the posterior expected payoff that would have been generated by the choice that is best, *conditional* on his ex post knowledge. This posterior expected payoff is the agent's reference point. We assume that in addition to his material payoff, the agent obtains an emotional payoff, which is given by the difference between his actual payoff and this reference point.

[1] In this case, due to symmetry, he also strictly prefers the known urn if he is offered a bet that wins \$100 if a white ball is drawn and \$0 otherwise, which is inconsistent with expected utility maximisation.

When the agent chooses the uncertain urn, the outcome reveals information about its true composition. Accordingly, he revises his assessment of the urn's expected payoff. If he wins the bet, he infers that *his actual choice was the best choice* since his posterior assigns greater probability to states in which there are more black balls than white balls. Thus, in this case, his reference point is the uncertain urn, and he rejoices as he compares his payoff of \$100 to the posterior mean of the uncertain urn, which, given his posterior, is strictly greater than \$50 and less than \$100. If instead he loses the bet, his posterior puts greater probability weight on states in which there are more white balls than black balls. Accordingly, he infers that *the risky urn would have been the best choice.* Thus, in this case, his reference point is the risky urn, and he experiences regret as he compares his actual payoff of \$0 to the expected payoff of the risky urn of \$50. When the agent chooses the risky urn instead, no outcome leads the agent to revise his assessment of either urn's expected payoff. Thus, whether he wins or loses, *his choice is optimal from an ex post perspective*: in either event, his reference point is the risky urn, and he compares his actual payoff to the posterior mean of the risky urn, which is \$50.

Notice that when the agent chooses the risky urn, his possible regrets equal his possible rejoicing—they exactly offset each other. By contrast, when he chooses the uncertain urn, his regret in the event of a bad outcome *more* than offsets his rejoicing in the event of a good outcome because in the latter event *he changes his assessment of what the optimal choice is* in the light of his ex post information, pushing up his reference point and diminishing his rejoicing. Accordingly, the choice of the uncertain urn leads to higher expected regret.

In this paper, we generalise the example outlined above. In the spirit of Segal (1987), we model uncertainty by compound lotteries. A compound lottery can be thought of as an urn of unknown composition. We refer to the composition of an urn as its first order distribution. Playing out a compound lottery means that in the first stage, the true first order distribution of the urn is determined, and in the second stage, an outcome is drawn according to this first order distribution. A mixture of two compound lotteries corresponds to the compound lottery where in the second stage the outcome is drawn according to the mixture of the first order distributions of the two urns. Schmeidler (1989) defines uncertainty aversion as a preference for mixtures of compound lotteries over at least one of the compound lotteries that make up the mixture.

Our two main propositions provide sufficient conditions under which regret leads to uncertainty aversion. Intuitively, regret and uncertainty aversion are linked because observing an outcome of a mixture of compound lottery provides less information than observing an outcome of one of the lotteries that make up the mixture. The reason is that the possible first order distributions of a mixture are less dispersed and therefore more similar to one another. This suggests that observing an outcome of the mixture will be less conclusive about its true first order distribution.

We identify two classes of compound lotteries for which this intuition holds. The lotteries in the first class have first order distributions that satisfy certain monotonicity conditions. These conditions hold automatically if there are only two outcomes such as "winning" or "losing" the bet as is the case in typical Ellsberg experiments. Thus, our notion of regret generates uncertainty aversion in all environments with two outcomes. The lotteries in the second class satisfy certain symmetry conditions. These conditions can capture a situation in which the agent is completely ignorant about the composition of an urn so that the principle of insufficient reason requires the agent to regard all possible urn compositions as equally likely.

Seeking psychological foundations of uncertainty aversion is worthwhile insofar as it generates new testable predictions. Our theory has implications that do not follow from other theories of uncertainty aversion and so point to ways in which our model could be empirically distinguished from its rivals. Uniquely among existing theories, our agent's behavior depends on the nature of the feedback that he expects to receive on his actual as well as his forgone choices. In particular, our model predicts that his uncertainty aversion will be mitigated when he is exposed to the same feedback about the options in his choice set regardless of the option that he chooses. Our formal analysis focuses on the case in which the agent only learns the outcome of the urn he has chosen. But our model also predicts that if the agent observes draws from both urns irrespective of his choice, he will no longer display uncertainty aversion. This is because under these conditions the effect on his hindsight knowledge is the same whether he chooses the risky or uncertain urn. Thus, our model predicts that the tendency of experimental subjects to select the unambiguous urn will be mitigated when they are told in advance that they will receive feedback on the composition of both urns regardless of the choice that they make.[2]

Insofar as we seek a psychological explanation of uncertainty aversion, our theory resembles those of Halevy and Feltkamp (2005) and Morris (1997). Halevy and Feltkamp argue that uncertainty aversion can be explained by risk aversion if agents employ a rule of thumb according to which any given gamble is evaluated as if it were bundled with another identical gamble. Halevy and Feltkamp's work is of particular relevance as they obtain uncertainty aversion under a condition similar to our monotonicity condition. We discuss the formal relation between their and our approach and how the two theories can be experimentally tested against each other in more detail in section 6 after we have presented our results. Morris explains uncertainty

[2] Curley et al. (1986) conduct an experiment in which they indeed manipulate subjects' expected feedback by revealing the contents of the uncertain urn ex post. The data reject their hypothesis that this feedback manipulation *increases* uncertainty aversion. This is actually consistent with our model, which predicts that such forced feedback will *reduce* uncertainty avoiding behaviour.

aversion as arising from the application of a different rule of thumb that instructs agents to choose as if they were making a bet against a better informed experimenter. Avoidance of uncertain urns arises from adverse selection in the style of no-trade results.

Our model of the agent's preferences is motivated by two psychological observations. The first is the observation that a person's evaluation of an outcome is often determined by its comparison with salient *reference points* as well as by its intrinsic characteristics.[3] Specifically, people appear to care directly about the comparison of outcomes of their choices to pertinent counterfactuals such as what could, should, or might have happened.[4] Second, psychological evidence indicates that reference points are often constructed not only ex ante but also *ex post*, once the outcomes of the events to which they pertain have been realised. According to Kahneman and Miller (1986), "events in the stream of experience ... are interpreted in a rich context of remembered and constructed representations of what it could have been, might have been or should have been" (p.136).

Our modeling of regret using a choice-dependent reference point is similar to the modeling of disappointment aversion in Koszegi and Rabin (2006). But whereas Koszegi and Rabin posit a reference point depending on an agent's ex ante expectations about what he will do, our agent's reference point depends on the agent's ex post beliefs about what he should have done ex ante if blessed with the wisdom of hindsight. In addition, Koszegi and Rabin assume loss-aversion, whereas loss-aversion in our model arises endogenously: Since the reference point depends on the agent's ex post assessment of what he should have done, the agent's expected regret looms larger than his expected rejoicing.

The assumption that the agent's evaluation is informed by knowledge gained in hindsight, is consistent with psychological evidence that indicates that when judging past decisions in hindsight, people consistently exaggerate what could have been anticipated in foresight, a tendency known as the hindsight bias (Fischhoff 1975). The classic regret theory that was developed by Loomes and Sugden (1982) and Bell (1982, 1983) also shares this assumption to some extent. In regret theory, an agent's utility depends on the ex post comparison of the outcome of his chosen option with those of unchosen alternatives. By contrast, in our model, the identity of the option whose payoff will serve as the reference point is itself determined in the light of the agent's ex post information and, thus, may actually be the agent's chosen option.

A large number of papers provide axiomatic foundations for uncertainty aversion. Several of

[3] Reference dependence is a feature of Kahneman and Tversky's (1979) Prospect Theory. For a review of some of the evidence on reference dependence see, for example, Rabin (1998).

[4] For experimental evidence that indicates that counterfactual emotions such as disappointment and regret influence choice behaviour see, for example, Mellers et al. (1999), Mellers (2000) and Mellers and McGraw (2001).

these represent uncertainty aversion as the product of vague priors.[5] For example, in an extension of Savage's (1954) minimax approach, Hayashi's (2008) considers an agent who entertains a set of subjective priors and chooses the option which minimises over this set of priors the highest possible expected "regret" defined as the difference between his actual outcome and the best possible outcome in the state (assuming the state is revealed ex post).[6,7]. In contrast to Hayashi, our model is driven by information aversion rather than vague priors. Accordingly, our approach can be experimentally discriminated from Hayashi's by manipulating the amount of feedback and/or the precision of priors that are given to subjects. However, if the prior is unique in Hayashi's setup and the state is perfectly revealed ex post in ours, Hayashi's representation and our utility function are identical. (In fact both boil down to expected utility.)

Recursive expected utility formulations of uncertainty aversion relax the reduction of compound lotteries axiom and derive preference representations whereby agents act as if they evaluate uncertainty in terms of second-order probabilities. For example, Klibanoff et al.'s (2005) agent first computes (ordinary) expected utility with respect to each possible first-order distribution and then takes the (second-order) expected utility of these first-order exptected utilities with respect to his second-order beliefs.[8] Our work complements this literature by identifying a psychological reason (regret) why an agent may directly care about the second-order risk he faces and, in turn, display uncertainty aversion. While a standard interpretation of these recursive utility models assumes that second-order risk is always subjective, our agent also fails to reduce compound lotteries even if the second-order probabilities are objective. This is in line with the experimental results of Halevy (2007), which demonstrate that there is a negative correlation between the tendency to reduce *objective* compound lotteries and uncertainty aversion: subjects who reduce such lotteries tend to be uncertainty neutral, while subjects who fail to reduce such lotteries tend to be uncertainty averse. This suggests that uncertainty aversion is associated with a more general tendency to fail to reduce compound lotteries rather than the more specific tendency to fail to perform such a reduction only when one set of probabilities is subjective.[9]

[5] This group includes the Choquet expected utility approach by Schmeidler (1989) and the maximin expected utility approach by Gilboa and Schmeidler (1988). For a generalisation of the Choquet and maximin approach see Ghirardato et al. (2004).

[6] See Milnor (1954) for an axiomatic characterisation of the minimax approach.

[7] Interestingly, Savage (1954, p. 163) himself referred to *loss* and was reluctant to use the term regret: "... some have proposed to call loss "regret," but that term seems to me charged with emotion and liable to lead to such misinterpretation as that the loss necessarily becomes known to the person."

[8] Ergin and Gul (2008), Ahn (2008) and Nau (2006) provide alternative axiomatic foundations of similar recursive expected utility formulations.

[9] But see Halevy (2007) who argues that recursive expected utility models can be viewed as consistent with his results.

This paper is organized as follows. Sections 2 presents the model. Section 3 adapts the notion of uncertainty to our context. In section 4, we give an example. Section 5 contains the formal analysis. Section 6 discusses testable implications and presents empirical support for our model from the psychological literature. Section 7 concludes. All proofs are in the appendix.

2 The model

2.1 Independent urns

We consider an agent who faces a choice problem under uncertainty. We model the choice problem as one of selecting an urn of unknown composition. Urns are represented by compound lotteries in which the composition of the urn is realized at the first stage and the payoff is determined by a draw from this urn at the second. That is, urns are represented by horse bets over roulette lotteries (Anscombe and Aumann, 1963).

There is a finite number of urns (or actions) $d \in D$. A draw form urn d delivers one of N possible payoffs $x_1 < ... < x_N$. The probabilities of these payoffs depend on the composition of urn d which is described by its state $\omega^d \in \Omega = \{1, ..., \overline{\omega}\}$.[10] Thus, there are $\overline{\omega}$ possible states per urn. Let $h^d_{n\omega^d}$ be the probability with which urn d yields payoff x_n in state ω^d. Let

$$h^d_{\omega^d} = \begin{pmatrix} h^d_{1\omega^d} \\ \vdots \\ h^d_{N\omega^d} \end{pmatrix}$$

represent the (first order) payoff distribution of urn d when the urn's true state is ω^d. The following $N \times \overline{\omega}$ Markov matrix describes the set of possible payoff distributions, one for each state, of urn d:

$$h^d = \left(h^d_1, ..., h^d_{\overline{\omega}}\right).$$

We assume that the agent does not know an urn's state but holds a prior belief over the possible states. Let $\pi^d\left(\omega^d\right)$ be the marginal probability with which the agent believes that the state of urn d is ω^d.

We will consider an agent, who as a result of his regret cares about the information that his choice reveals about the composition of an urn. In general, this includes information that a draw from one urn might reveal about a different urn. In Ellsberg-type problems, however, urns are typically independent. Therefore, we abstract from cross informational effects and

[10] The assumption that each urn has the same number of possible states is to simplify notation and is not substantial.

model urns as independent. This means that the joint probability that the agent assigns to the profile of urn states $\left(\omega^d\right)_{d\in D}$ is the product of the prior marginal probabilities: $\Pi_{d\in D}\pi^d\left(\omega^d\right)$.

We denote by X^d the random variable with values in $\{x_1, ..., x_N\}$ that describes the payoff from a draw from urn d. Let p^d be the (ex ante) distribution of X^d:

$$p^d(x_n) = \sum_{\omega\in\Omega} h^d_{n\omega}\pi^d(\omega).$$

Finally, we remark that it is without loss of generality to assume that for all $d, d' \in D$: $\pi^d = \pi^{d'}$. This is because states can always be appropriately "split".[11] Thus, we drop action indices from π.

2.2 Preferences

The agent cares about both his material payoff and his ex post evaluation of the performance of his action relative to the performance of a *reference action*. In the spirit of norm theory (Kahnemann and Miller, 1986), we suppose that the reference action is constructed *ex post*, that is, after the agent has observed the outcome of his choice. Specifically, we assume that the reference action is the *ideal* action that the agent believes that he should have chosen had he known with foresight what he knows with hindsight. The ex post comparison between what the agent actually got and what he perceives he should have gotten induces an emotional reaction: if the actual payoff falls short of his reference action's payoff, he experiences regret and his utility falls; otherwise, he rejoices and his utility rises.

Formally, the agent perceives that in period 0 Nature determines the profile of urn states $\left(\omega^d\right)_{d\in D}$ *once and for all* according to the prior probabilities in a move that the agent cannot observe. In period 1, the agent selects an urn d, and the agent's *material* instantaneous utility x_n is realized with probability $p^d(x_n)$.[12]

We define the reference action as the action that would maximise the agent's material payoff in a hypothetical (counterfactual) period 2 choice problem, in which the agent could select among the *same* urns in the light of his ex post knowledge. To fix terminology, we refer to the hypothetical agent who makes this period 2 choice as the agent's *hindsight self.* For any $d \in D$, let

$$Y^d$$

[11]To illustrate the argument, suppose D contains two actions a, b with $\pi^a \neq \pi^b$. Then define the new state space for urn d as $\widehat{\Omega} = \Omega \times \Omega$ and for $\widehat{\omega}^d = \left(\omega^a, \omega^b\right) \in \widehat{\Omega}$, let $\widehat{\pi}\left(\widehat{\omega}^d\right) = \pi^a\left(\omega^a\right)\pi^b\left(\omega^b\right)$.

[12]Identifying the agent's material utility with payoff is without loss of generality. If the agent has "material risk preferences" $v(x_n)$, then we can replace x_n by $x'_n = v(x_n)$.

be the random variable that describes the hypothetical payoff from selecting urn d in period 2—the payoff of the hindsight self. We assume that the agent perceives that in his hindsight self's choice problem the composition (the state) of the urn is the same as in his actual choice problem, but that a new, conditionally independent outcome is drawn from the urn. If the agent chose urn d and observed $X^d = x_n$ in period 1, then his hindsight self's expected material payoff from choice $e \in D$ is given by the posterior mean[13]

$$E\left[Y^e | X^d = x_n\right] = \sum_{\omega\in\Omega}\sum_{m=1}^{N} x_m h^d_{m\omega}\frac{h^d_{n\omega}\pi(\omega)}{p^d(x_n)}.$$

The agent's *reference action* is the hindsight self's best period 2 action, and the payoff from this action is the agent's *reference point* $r_d(x_n)$:

$$r_d(x_n) = \max_{e\in D} E\left[Y^e | X^d = x_n\right].$$

The emotional component of the agent's instantaneous utility—his regret—is given by the difference between his actual material payoff and his reference point. His *overall* instantaneous utility from action d and outcome x_n, $u^d(x_n)$, is a linear combination of his material utility and his regret:

$$u^d(x_n) = x_n + \rho(x_n - r_d(x_n)),$$

where $\rho \geq 0$ measures the agent's regret concerns.[14]

2.3 Behaviour

We assume that the agent anticipates his emotional response and maximises his expected utility taking his regret concerns into account. Thus, the agent chooses d to maximise

$$\begin{aligned} U^d &= E\left[u^d\left(X^d\right)\right] \\ &= E\left[X^d\right] + \rho E\left[X^d - \max_{e\in D} E\left[Y^e|X^d\right]\right]. \end{aligned} \tag{1}$$

Our independence assumptions imply that when the agent chooses d, he revises his beliefs about the payoff distribution of urn d while maintaining his ex ante beliefs about the payoff distribution of urn e, for all other $e \neq d$. Thus,

$$E\left[Y^e|X^d\right] = \begin{cases} E\left[Y^d|X^d\right] & \text{if} \quad e = d \\ E\left[X^e\right] & \text{if} \quad e \neq d \end{cases}. \tag{2}$$

[13] If $p^d(x_n) = 0$, we set $E\left[Y^e|X^d = x_n\right] = E\left[Y^e\right]$.

[14] Strictly speaking, reference point and utility depend all on the choice set D. Since we do not consider variations across choice sets, we suppress this dependency.

Our objective in this paper is to isolate the effects that are driven by the agent's regret concerns. Therefore, we focus on the case in which a standard expected utility maximiser ($\rho = 0$) is indifferent between all actions:

$$E\left[X^d\right] = E\left[X^{d'}\right] \qquad \forall\ d, d' \in D.$$

Equation (2) then implies that $\max_{e\in D} E\left[Y^e|X^d\right] = \max\left\{E\left[X^d\right], E\left[Y^d|X^d\right]\right\}$, and thus, (1) simplifies to

$$U^d = E\left[X^d\right] + \rho E\left[X^d - \max\left\{E\left[X^d\right], E\left[Y^d|X^d\right]\right\}\right]. \tag{3}$$

Remarks Since the agent's utility depends directly on his beliefs about the actions of his hindsight self, who moves in period 2 and seeks to maximise the reference payoff $E\left[Y^e|X^d\right]$, the agent's behaviour can be formally viewed as the equilibrium outcome of an intra-personal psychological game in the sense of Geanakoplos et al. (1989) where, in equilibrium, the agent correctly predicts his hindsight self's action. When the agent has chosen d, then the hindsight self's optimal strategy is to stick to d for all x_n with $E\left[Y^d|X^d = x_n\right] \geq E\left[X^d\right]$ and to switch actions otherwise.

As in Halevy and Feldkamp (2005), we are agnostic about how the agent's prior comes about. In the classic Ellsberg problem, the "horse"-component of the lottery reflects the agent's subjective uncertainty about the composition of the uncertain urn. However, our theory is also applicable if horse-bets represent objective risk as would be the case if, for example, the experimenter told the agent the prior probabilities of the possible urn compositions. For our purposes, it is important only that the agent distinguishes between his uncertainty about the composition of an urn and the uncertainty about the ball that will be drawn from an urn of a given composition. Crucially, the agent's regret concerns imply that he does not reduce the compound lottery to a simple probability distribution over final outcomes (even if the horse bets represent objective risk).

3 Independent mixtures and uncertainty aversion

In the spirit of Schmeidler (1989), we define uncertainty aversion to be a general preference for mixtures of urns over the worst of these urns. Schmeidler refers to a mixture as a *statewise* mixture, that is, urn c is a mixture of urns a and b if, conditional on the state, the payoff of urn c is determined by drawing from urn a with probability λ and from urn b with probability $1-\lambda$ for some $\lambda \in [0, 1]$. When c is a mixture of a and b, then it is less uncertain than urn a and urn b. Intuitively, this is so since, conditional on the state, the composition of urn c is an average of the compositions of urn a and urn b. Therefore, the compositions of urn c across all possible states are more similar to one another than the compositions of urn a or urn b, and thus there is less uncertainty about the true composition of urn c than about the true composition of urn a or urn b.

If c is a statewise mixture of a and b, then the composition of urn c is correlated with those of urn a and urn b. Thus, a draw from c is informative about the composition of a and b, and the random variable X^c is not independent of X^a and X^b. Since we want to abstract from such cross informational effects, we introduce the concept of an *independent mixture.*

Definition 1 *Let $\lambda \in [0,1]$. Action c is an independent λ-mixture of actions a and b if:*
(i) $h^c_{n\omega} = \lambda h^a_{n\omega} + (1-\lambda) h^b_{n\omega}$ for all $n \in \{1, ..., N\}, \omega \in \Omega$.
(ii) The family $\left(X^a, X^b, X^c\right)$ is stochastically independent.

An independent mixture preserves the averaging feature of a statewise mixture, but avoids the implication that urns are correlated. We define uncertainty aversion with respect to independent mixtures.

Definition 2 *Let c be an independent λ-mixture of actions a and b. An agent is uncertainty averse with respect to actions $a, b,$ and c if for all $\lambda \in [0,1]$ it holds that*

$$U^c \geq \min\left\{U^a, U^b\right\},$$

and if the inequality is strict for at least one $\lambda \in (0,1)$.

4 Example

We illustrate our model by a simple two-urn, two-state, two-outcome example. There are two urns, a and c, which contain white and black balls. Drawing a white ball yields $x_1 = 0$, and drawing a black ball yields $x_2 = 100$. There are two possible states of each urn, the first corresponding to a "good" composition and the second to a "bad" composition. In the "good" state $\omega^d = 1$, the fraction of black balls is $\frac{1}{2} + \eta^d$, and in the "bad" state $\omega^d = 2$, the fraction of black balls is $\frac{1}{2} - \eta^d$ for $\eta^d \in [0, 1/2]$. Thus,

$$h^d = \begin{pmatrix} \frac{1}{2} - \eta^d & \frac{1}{2} + \eta^d \\ \frac{1}{2} + \eta^d & \frac{1}{2} - \eta^d \end{pmatrix}.$$

Figure 1 illustrates the compound lottery that describes urn d.

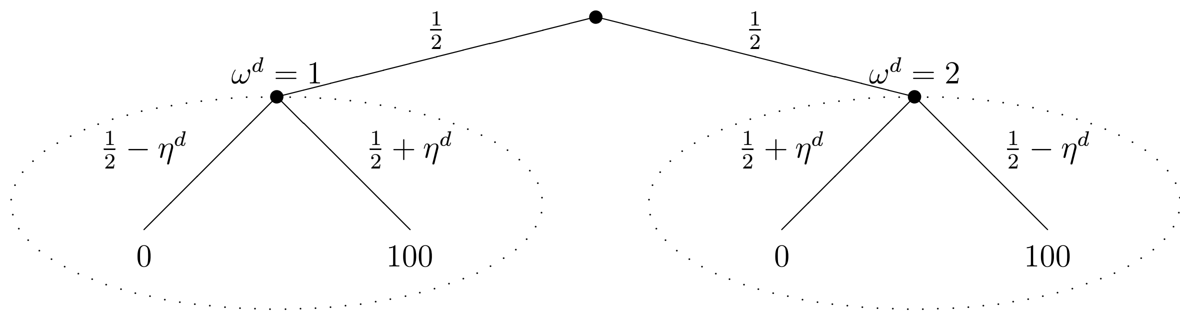

Figure 1: urn d

We assume that the states of each urn are equally likely: $\pi^d(1) = \pi^d(2) = 1/2$. Thus, by independence, the joint probability of any profile $(\omega^a, \omega^c) \in \{1,2\}^2$ is always equal to $\frac{1}{4}$.

We now show that if $\eta^a > \eta^c$, then the agent strictly prefers urn c to a and that this preference corresponds to uncertainty aversion. Note that $\eta^a > \eta^c$ intuitively means that urn a is the more uncertain urn since it implies that the possible compositions of a are more dissimilar than those of urn c, which means that the outcome of urn a is less predictable ex ante. In the extreme case in which $\eta^c = 1/2$, there is no uncertainty about the composition of urn c.

Suppose that the agent has chosen urn a. He uses the information contained in the outcome to update his belief about the state of the urns and then constructs the hypothetical payoff from the optimal period 2 choice. Since the outcome of urn a contains no information about the state of urn c, his expected payoff from switching to c in period 2 is just the ex ante expectation of c, irrespective of the outcome $X^a = x_n$:

$$E\left[Y^c | X^a = x_n\right] = 1/2 \cdot 100 + 1/2 \cdot 0 = 50.$$

In contrast, his expected material payoff from sticking with a does depend on the outcome. If the outcome is good ($X^a = 100$), his posterior assigns probability $1/2 + \eta^a$ to urn a being in the "good" state and accordingly the expected material payoff from choosing urn a again is

$$E\left[Y^a | X^a = 100\right] = (1/2 + \eta^a) \cdot 100 + (1/2 - \eta^a) \cdot 0 > 50.$$

It follows that the hindsight self would stick with urn a and so the agent's reference point is given by $r_a(100) = E\left[Y^a | X^a = 100\right]$. Thus, his instantaneous utility is

$$u^a(100) = 100 + \rho\left(100 - E\left[Y^a | X^a = 100\right]\right).$$

Conversely, if the outcome is bad ($X^a = 0$), his posterior assigns probability $1/2 - \eta^a$ to urn a being in the "good" state, and accordingly the expected material payoff from choosing urn a again is

$$E[Y^a | X^a = 0] = (1/2 - \eta^a) \cdot 100 + (1/2 + \eta^a) \cdot 0 < 50.$$

It follows that the hindsight self would switch to urn c, and so the agent's reference point is given by $r_c(0) = E[Y^c | X^a = 0] = 50$. Thus, his instantaneous utility is

$$u^a(0) = 0 + \rho(0 - E[Y^c | X^a = 0]) = -50\rho.$$

Therefore, the agent's expected utility from choosing urn a is given by

$$\begin{aligned} U^a &= 1/2 \cdot u^a(100) + 1/2 \cdot u^a(0) \\ &= 50 + 1/2 \cdot \rho(100 - E[Y^a | X^a = 100]) + 1/2 \cdot (-50\rho) \\ &= 50 - \rho(1/2 + \eta^a) 50. \end{aligned}$$

Analogously,

$$U^c = 50 - \rho(1/2 + \eta^c) 50.$$

Therefore, if $\eta^c < \eta^a$, then $U^c > U^a$, which means that the agent strictly prefers urn c to urn a.

Now let action b be an urn with the same Markov matrix as a except with the columns reversed. (This is, of course, exactly the same urn with a different labelling of states given our independence assumption.) Thus,

$$h^b = \begin{pmatrix} \frac{1}{2} + \eta^a & \frac{1}{2} - \eta^a \\ \frac{1}{2} - \eta^a & \frac{1}{2} + \eta^a \end{pmatrix}.$$

Notice that c is an independent mixture of a and b since for $\lambda = \frac{1}{2}(1 + \eta^c/\eta^a)$, we have $h^c = \lambda h^a + (1 - \lambda) h^b$. Therefore, since the agent prefers c to a, we have shown that the agent displays uncertainty aversion with respect to actions a, b, and c. We will generalise this example in section 5.2.

5 Analysis and results

For the rest of the paper, we consider a fixed choice set $D = \{a, b, c\}$ where c is an independent λ-mixture of a and b. We study environments in which regret leads to uncertainty aversion.[15] First, we show that the agent displays uncertainty aversion when the first-order distributions of the urns are ranked by first-order stochastic dominance and a second condition which establishes

[15] A similar approach is adopted by Halevy and Feltkamp (2005).

the uniformity of the hindsight self's response to the draw from the urn. Second, we show that when the agent's uncertainty about the states of the urns can be characterized by a Markov matrix that exhibits a certain symmetry property in a sense to be more precisely defined, then the agent will display uncertainty aversion.

5.1 Uniform switching and statewise first order stochastic dominance

In this section, we show that uncertainty aversion arises when two conditions hold. The first requires that the hindsight self chooses the same urn as his actual self whenever his observation exceeds a uniform threshold that is independent of the agent's actual choice d. The second requires that the first-order distributions of urn a and urn b are ranked in terms of first-order stochastic dominance and at least one distribution strictly first-order stochastically dominates another. The conditions are formally stated in Proposition 1.

Proposition 1 *If the following conditions jointly hold, then the agent displays uncertainty aversion:*
(i) There is a threshold $x_{n^*} \in \{x_1, ..., x_N\}$ *such that for all* $d \in D : E\left[Y^d | X^d = x_n\right] > E\left[X^d\right]$ *if and only if* $x_n > x_{n^*}$.
(ii)(a) For all $\omega \in \Omega$, *either* h^a_ω *first-order stochastically dominates* h^b_ω *or vice versa, and (b) there is an* $\omega \in \Omega$ *with* $\pi(\omega) > 0$ *such that* h^a_ω *strictly first-order stochastically dominates* h^b_ω *or vice versa.*

Notice that conditions (i) and (ii) are automatically satisfied if the number of payoff outcomes is $N = 2$. Thus, we can infere the following important corollary:

Corollary 1 *If there are only two outcomes* ($N = 2$)*, then the agent (typically*[16]*) displays uncertainty aversion.*

Accordingly, fear of regret can explain behaviour in a typical Ellsberg experiment, in which the only outcomes are winning or losing the bet.

To understand the intuition behind Proposition 2, recall that the reference point represents the expected payoff of the hindsight self who makes a hypothetical choice ex post. Upon observing outcome x_n from urn d, the hindsight self sticks to urn d if the posterior mean $E\left[Y^d | X^d = x_n\right]$ of urn d exceeds its prior mean $E\left[X^d\right]$. In this case, the agent's reference point is pushed up, exceeding his prior expectation. Otherwise, the hindsight self chooses some

[16] If $N = 2$, condition (ii) (b) is violated in degenerate cases only.

other urn, and the agent's reference point is the same as his prior expectation ($E\left[X^d\right]$). We refer to an observation which pushes up the agent's reference point as "critical". Condition (i) guarantees that all urns have the same critical observations.

Because urn c is a mixture, choosing urn c essentially amounts to choosing urn a with probability λ and urn b with probability $1-\lambda$. Therefore, if the agent selects urn c, then there is a positive probability that the outcome is drawn from urn a, it is critical, and the hindsight self selects the other urn b. Likewise, there is a positive probability that the outcome is drawn from urn b, it is critical, and the hindsight self selects the other urn a. In other words, choosing urn c amounts to mixing an observation from a with the selection of b by the hindsight self (and vice versa) with positive probability. It is these possibilities that make the utility associated with urn c different from the utilities of the "pure" urns a and b. If one of the pure urns is chosen by contrast, urn a for example, then if a critical observation arises, the hindsight self never selects urn b because the critical observation induces him to choose the same pure urn again.

This raises the question when mixing an observation from a with the selection of b by the hindsight self could give rise to more regrets than a draw from urn a that is followed by selection of urn a. This can be the case only if there are some states in which a critical observation from a is more likely than a critical observation from b and, at the same time, the payoff from b in this state is larger than the payoff from a in this state (and vice versa). But the latter is ruled out by condition (ii) which guarantees that in any state in which a is more likely to generate a critical observation than b, the payoff from a in this state is also larger than the payoff from b in this state (and vice versa).

Condition (i) in Proposition 1 depends on the unconditional distribution of payoffs. The next lemma provides a sufficient condition for (i) in terms of the primitives h^a_ω and h^b_ω. It states monotonicity conditions, which guarantee that the posterior means $E\left[Y^a \middle| X^a = x_n\right]$ and $E\left[Y^b \middle| X^b = x_n\right]$ are increasing in x_n and "cut through" $E\left[X^d\right]$ at the same threshold x_{n^*}.

Lemma 1 *The following conditions are sufficient for condition (i) in Proposition 1. For* $d = a, b$*:*
(i) There is a threshold $x_{n^*} \in \{x_1, ..., x_N\}$ *such that* $h^d_{n\omega}$ *is strictly increasing in* ω *if* $x_n > x_{n^*}$ *and decreasing in* ω *if* $x_n \leq x_{n^*}$*.*
(ii) For all $\omega > \theta$*, the difference* $h^d_{n\omega} - h^d_{n\theta}$ *is increasing in* n*.*

The following example illustrates the conditions in Lemma 1:

$$h^a = \frac{1}{10}\begin{pmatrix} 6 & 3 & 1 \\ 3 & 2 & 0 \\ 1 & 5 & 9 \end{pmatrix}, \quad h^b = \frac{1}{10}\begin{pmatrix} 5 & 4 & 2 \\ 2 & 2 & 1 \\ 3 & 4 & 7 \end{pmatrix}.$$

Condition (i) in Lemma 1 holds for $n^* = 2$. It is easy to verify (ii). Observe also that the first-order stochastic dominance ranking condition from Proposition 1 is satisfied: h_1^b dominates h_1^a, h_2^a dominates h_2^b, and h_3^a dominates h_3^b.

5.2 Strong Symmetry

In this subsection, we show that regret leads to uncertainty aversion when the urns in the agent's choice problem display symmetry properties. The conditions generalise the example discussed in section 4. From a methodological point of view, under the conditions that we identify in this subsection, the theory of Halevy and Feltkamp (2005) does not necessarily generate uncertainty aversion (as is demonstrated by an example in the appendix). Thus, our theory can in principle be tested against Halevy and Feltkamp's by exposing subjects to the choice problems discussed in this section.

We will use the basic insight that the integrand

$$X^d - \max\left\{E\left[X^d\right], z\right\} \tag{4}$$

in the objective function (3) is concave in z. Since the agent chooses d so as to maximise $E\left[X^d - \max\left\{E\left[X^d\right], E\left[Y^d|X^d\right]\right\}\right]$, and since all urns have the same mean by assumption, this implies that action c is preferred to action a if $E\left[Y^a|X^a\right]$ is a mean preserving spread of $E\left[Y^c|X^c\right]$. Intuitively, the more dispersed the distribution of the posterior means, the more highly correlated they are with the true material payoff from the urn and, thus, the better the decision the hindsight self will make. Since the agent's utility is decreasing in the hindsight self's payoff, the agent therefore dislikes a more dispersed distribution of the posterior means. We summarize this basic property in the following lemma.

Lemma 2 *The agent prefers action c over action a if $E\left[Y^a|X^a\right]$ is a mean preserving spread of $E\left[Y^c|X^c\right]$.*

We now identify conditions such that $E\left[Y^a|X^a\right]$ is a mean preserving spread of $E\left[Y^c|X^c\right]$. We use the following definition of a mean preserving spread (see Rothschild and Stiglitz, 1970).

Definition 3 *$E\left[Y^a|X^a\right]$ is a mean preserving spread (MPS) of $E\left[Y^c|X^c\right]$ if there is a $N \times N$ transition matrix $(\tau_{nm})_{n,m}$ with $\tau_{nm} \geq 0$, $\sum_n \tau_{nm} = 1$ for all n, m such that*

$$E\left[Y^c|X^c = x_m\right] = \sum_n E\left[Y^a|X^a = x_n\right]\tau_{nm} \qquad \text{for all } m \in \{1, ..., N\}, \tag{5}$$

$$p^a(x_n) = \sum_m p^c(x_m)\,\tau_{nm} \qquad \text{for all } n \in \{1, ..., N\}.\text{[17]} \tag{6}$$

We show that (5) and (6) hold if the urns a and b satisfy specific symmetry conditions. Recall that the columns of the matrix h^d represent all possible (first-order) payoff distributions of urn d. We will look at urns that are *strongly symmetric* in the following sense: For any possible state ω there is another state ω' such that the highest payoff is as likely in state ω as is the lowest payoff in state ω', the second highest payoff is as likely in state ω as is the second to lowest payoff in state ω', etc.

If the agent has a uniform prior over states, then strong symmetry captures the intuitive idea behind the principle of insufficient reason which prescribes that the agent consider *any* first order distribution of an uncertain urn as a possible state and place equal probability weight on each of them.

The urns in the example discussed in section 4 are strongly symmetric. For a further illustration, we consider a specific numerical example:

$$h^a = \begin{pmatrix} .2 & .1 & .2 & .5 \\ .3 & .7 & .7 & .3 \\ .5 & .2 & .1 & .2 \end{pmatrix}.$$

Notice that, e.g., the worst outcome x_1 has the same probability .2 in state ω_1 as the best outcome x_3 in state ω_4. Observe that a matrix with this the property remains the same if one first reverses the order of the columns and then the order of the rows. We now give the formal definition. For an integer k, let

$$\sigma^k = \begin{pmatrix} 0 & & 1 \\ & \cdots & \\ 1 & & 0 \end{pmatrix}$$

be the $k \times k$ matrix with 1's on the secondary diagonale and 0's elsewhere. When the dimension is clear from the context, we omit k. Multiplication of σ from the right reverses the order of the columns of a matrix. Multiplication of σ from the left reverses the order of the rows of a matrix. Thus, an urn d is strongly symmetric if its associated matrix h^d is preserved under these two operations.

Definition 4 *A $k \times l$ matrix h is strongly symmetric if*

$$h = \sigma^k h \sigma^l.$$

We now consider strongly symmetric matrices h^a and ask under what assumptions on h^b the agent prefers the resulting mixture over h^a, thus displaying uncertainty aversion. The assumption we impose is that h^b is the columnwise mirror image of h^a, that is, $h^b = h^a\sigma$. Of course, h^a and h^b are utility equivalent, but if we mix h^a and h^b, we obtain another strongly symmetric matrix h^c whose payoff distribution in state $\omega = 1$ is a mixture of the payoff distributions of urn a in the states $\omega = 1$ and $\omega = \overline{\omega}$, whose payoff distribution in state $\omega = 2$ is a mixture of the payoff distributions of urn a in the states $\omega = 2$ and $\omega = \overline{\omega} - 1$, etc. To see how this works, return to the above example. The matrix h^b is the columnwise mirror image of h^a, and if h^c is a, say, .5-mixture of h^a and h^b, we have:

$$h^b = h^a \sigma = \begin{pmatrix} .5 & .2 & .1 & .2 \\ .3 & .7 & .7 & .3 \\ .2 & .1 & .2 & .5 \end{pmatrix}, \qquad h^c = \begin{pmatrix} .35 & .15 & .15 & .35 \\ .3 & .7 & .7 & .3 \\ .35 & .15 & .15 & .35 \end{pmatrix}.$$

Importantly note that if there are only two states or two payoffs, a .5-mixture of h^a and h^b yields an urn c whose first order payoff distributions are the same in all states, which means that choosing c involves no second-order risk.

We now prove the general claim that if urn c can be constructed from urn a in such a manner, then the agent prefers urn c to urn a and therefore displays uncertainty aversion. To do so, we also require that the prior π be strongly symmetric. To avoid rather uninteresting case distinctions, we assume that $p^a(x_n) > 0$ for all n.[18]

Proposition 2 *Let π and h^a be strongly symmetric. Let $h^b = h^a \sigma$. Then $E[Y^a | X^a]$ is an MPS of $E[Y^c | X^c]$. Thus, the agent displays uncertainty aversion.*

Intuitively, since the first-order distributions of urn c are mixtures of the first-order distributions of urn a, observations from urn a enable the hindsight self to make sharper subsequent predictions about the outcomes of a than he could make about the outcomes of urn c based on observations from c. This implies that the posterior mean of urn a, $E[Y^a | X^a]$, is more dispersed than the posterior mean of urn c, $E[Y^c | X^c]$. Strong symmetry guarantees that the degree of dispersion in the posterior means of the two urns can, in fact, be compared in the mean preserving spread sense.

The proof of Proposition 2 actually reveals more. We can interpret a matrix h^d as a random variable whose realisations h^d_ω are (first-order) distributions over $\{x_1, ..., x_N\}$. In this sense, we can ask when h^a is an MPS of h^c. It turns out that under strong symmetry this is in fact the case.[19] This means that the first-order distributions of urn a are riskier than the first-order distributions of urn c in the mean preserving spread sense. In other words, urn a exhibits

[18] This is actually without loss of generality. As apparent from the proof of Proposition 2, strong symmetry implies that $p^a = p^b = p^c$. Thus, we can delete zero-probability outcomes from the outcome space.

[19] To establish this, we have to verify the conditions of Definition 3 which can now be written: (i) $h^c = h^a \tau$, and (ii) $\pi^a = \tau \pi^c$. (i) is implied by strong symmetry of h^a, and (ii) is implied since, by assumption, $\pi = \pi^a = \pi^c$ is strongly symmetric.

more second-order risk than urn c. Therefore, the agent's preference for urn c is, in effect, an aversion to second-order risk. In other words, under strong symmetry, uncertainty aversion and second-order risk aversion are equivalent in our setup.

In general, $E[Y^a|X^a]$ being an MPS of $E[Y^c|X^c]$ does not guarantee a *strict* preference of c over a for some λ as is required in our definition of uncertainty aversion. Note first that the mean of $E\left[Y^d|X^d\right]$ equals $E\left[X^d\right]$ for $d = a, b$ and that the function $E\left[X^d\right] - \max\left\{E\left[X^d\right], z\right\}$ in (4) has a kink in $z = E\left[X^d\right]$. Thus, the preference of c over a is strict, if (a) $E[Y^a|X^a]$ is non-degenerate and if (b) $E[Y^a|X^a]$ is not identical to $E[Y^c|X^c]$. While both conditions are generically true in the class of strongly symmetrc matrices[20], the following lemma provides a sufficient condition for (b).

Lemma 3 *Under the conditions of Proposition 2, let $E[Y^a|X^a]$ be non-degenerate. Let the number of states be larger than the number of outcomces: $\overline{\omega} \geq N$, let h^a have rank N, and let $\pi(\omega) > 0$ for all ω. Then the preference of c over a is strict for all $\lambda \in (0, 1)$.*

6 Discussion

In this section we clarify the relationship between our work and Halevy and Feltkamp (2005), henceforth HF, and point out testable implications of our model. We also provide evidence from the psychology literature that supports our model.

6.1 Testable Implications

Like us, HF assume that the agent represents an urn as a compound lottery with a given (second-order) prior belief. HF assume that the agent mistakenly perceives that he faces a series of draws with replacement from the urn instead of a single draw only. Therefore, the successive realisations of the uncertain (compounded) urn are correlated with one another due to their common dependence on the unknown state. This correlation increases the overall risk to which the decision maker is exposed in comparison to the risky urn whose successive realisations are independent.

HF obtain uncertainty aversion under the FOSD ranking condition (ii) of Proposition 1. In fact, the driving force behind our result and HF's result is similar. When HF's agent prefers urn a over urn b, then he does so because the first and the second draw from urn a are more correlated than the first and the second draw from urn b. But this implies that the first draw

[20] If (a) or (b) are violated, then a slight perturbation of the entries of h^a changes $E[Y^a|X^a]$ and turns the equalities into inequalities. Similarly, (a) and (b) are true for generic payoff vector $(x_1, ..., x_N)$.

from urn a is a better predictor for the second draw from urn a than is the first draw from urn b for the second draw from urn b. Thus, observing the first draw from urn a provides more useful information for a second (hypothetical) decision than does observing the first draw from urn b. This intuitively suggests that our information-averse agent would also prefer urn b over urn a. Our results show that this intuition is formally true if in addition to (ii), condition (i) of Proposition 1 is imposed.

Despite these apparent similarities between our and HF's approach, the implications are different in at least two respects. First, as indicated at the beginning of subsection 5.2, it is possible to construct examples that satisfy our strong symmetry conditions of Proposition 2, for which the HF framework is without bite. We present such an example in the appendix. Therefore, by exposing subjects to such choice problems the two theories can be tested against each other.

There is a second perhaps more important implication that permits our model to be experimentally distinguished from that of HF. As explained in the Introduction, our agent directly cares about the feedback he receives. The driving force of our theory is that an uncertain urn generates more information about its composition than does a risky urn, and so selecting the risky urn enables the agent to avoid feedback about his choice. Therefore, our theory, unlike HF's, predicts that experimental subjects' tendency to select the risky urn will be mitigated when they are told in advance that they will receive feedback on the composition of both urns regardless of the choice that they make. Indeed, uniquely among theories of uncertainty aversion, the behaviour of our agent depends on the nature of the feedback he expects to receive on his actual as well as his forgone choice. Thus, by manipulating subjects' anticipated feedback, we can distinguish our model from other competing theories.

6.2 Empirical Evidence

Our theory requires that the agent decomposes the overall uncertainty to which he is exposed into two components: the uncertainty about the first-order distribution and the risk about the outcome, conditional on the first-order distribution. He blames himself only for his failure to know what he now knows about the first-order distribution, but not for failing to predict the actual outcome per se. Thus, he implicitly categorizes uncertainty into that which he believes he could know, and so may be blamed for not knowing, and that which he regards as inherently unknowable. The idea that people distinguish between uncertainty that arises from ignorance and the uncertainty that arises from intrinsic randomness has received backing in the psychology literature. Frisch and Baron (1988) conceive of uncertainty as precisely "the subjective experience of missing information relevant for a prediction", and, in support of this conception, Brun and Teigen (1990) find that subjects prefer guessing the outcome of an uncertain event before it has occurred to guessing it after it has occured but before they know it. Notice that while an event that has not yet occurred may be regarded as unknowable in principle, an event that has already occurred may be regarded as in principle knowable, in which case any failure to know the outcome is attributable to ignorance rather than inherent randomness. Thus, our model would predict that subjects would experience greater regret when engaging in postdiction than prediction, leading them to prefer the latter over the former. Indeed, Brun and Teigen's subjects commonly cite as a reason for their preference for prediction that wrong postdictions are much more embarassing than wrong predictions.

In addition, psychologists have conducted experiments in which subjects' behaviour is influenced by the amount of feedback they expect to obtain. Specifically, subjects become more willing to choose a riskier gamble if they will learn the outcome of the gamble regardless of the choice they make, while such feedback on the safer alternative is provided only in the event that they choose it.[21] This suggests that people make choices to minimise their exposure to information about the outcomes of unchosen alternatives, and supports a version of regret theory that was first suggested by Bell (1983) according to which people are averse to such feedback.[22] As previously discussed, our agent also displays information aversion. However, whereas Bell's agent is averse to feedback about outcomes, our agent is averse to feedback about the true probability distributions associated with the options in his choice set.

Finally, Ritov and Baron (2000) provide some evidence which points to a link between uncertainty aversion and regret. In their experiments, they find that the omission bias is intensified when subjects choose under conditions of uncertainty rather than mere risk. An omission bias is a tendency to prefer inaction to action even if the consequences of inaction are worse than the consequences of action. Psychological evidence suggests that this may be due to the tendency of commissions to elicit greater regret than omissions, at least in the short run.[23] If, as our model suggests, regrets may be intensified under conditions of ambiguity, this would help to explain why the omission bias is also amplified under these conditions.

7 Conclusion

This paper represents an attempt to gain a better understanding of the psychological forces that drive uncertainty aversion. We have proposed a model of regret that can account for

[21] For a review of this evidence, see Zeelenberg (1999).

[22] For a detailed discussion of feedback effects in a Bell-type framework, see Krähmer and Stone (2008).

[23] According to Gilovich and Medvec (1994), commissions tend to elicit greater regret in the short run while ommissions elicit greater regret in the long run.

Ellsberg type behaviour in choice under uncertainty. We have argued that uncertainty aversion arises from an agent's aversion to discovering that his choices are suboptimal from an ex post perspective.

We view our paper as forming part of a broader research agenda that seeks to find emotional underpinnings of behavioural phenomena. In general, this is a worthwhile enterprise if the resulting psychological models generate novel predictions about the phenomena they seek to explain. Our model of regret meets this test by suggesting that a kind of information aversion accounts for uncertainty aversion under specific conditions. Accordingly, it identifies situations in which agents will be more or less likely to make uncertainty avoiding choices. According to our model, agents will display greater uncertainty aversion when conditions are such that choosing the more uncertain option reveals more information about the (first order) distribution of the chosen option. In particular, as we have previously argued, uniquely among existing theories of uncertainty aversion our model predicts that an agent's uncertainty aversion will be mitigated when he is exposed to the same feedback about the (first-order) distributions of options in his choice set regardless of the option that he chooses.

Appendix

Proof of Proposition 1 Without loss of generality, suppose $U^a \geq U^b$. We have to show that $U^c \geq U^b$ for all $\lambda \in [0,1]$ with strict inequality for one $\lambda \in (0,1)$. Note first that the material term $E\left[X^d\right]$ in (3) is the same for all actions. Thus, only differences in the regret term

$$V^d = E\left[X^d - \max\left\{E\left[X^d\right], E\left[Y^d|X^d\right]\right\}\right]$$

matter. Thus, $V^a \geq V^b$, and we need to show that $V^c \geq V^b$ for all $\lambda \in [0,1]$ with strict inequality for one $\lambda \in (0,1)$. To this end, we first re-write V^d for all $d \in D$ by conditioning on observations $X^d = x_n$. Note that by (i), for all $d \in D$, we have $E\left[Y^d|X^d\right] > E\left[X^d\right]$ if and only if $X^d > x_{n^*}$. Thus,

$$V^d = \sum_{n=1}^{n^*} P\left[X^d = x_n\right]\left(E\left[X^d - E\left[X^d\right]\right]\right) + \sum_{n=n^*+1}^{N} P\left[X^d = x_n\right]\left(E\left[X^d\right] - E\left[Y^d|X^d = x_n\right]\right).$$

The first term on the right hand side disappears. The second term depends only on marginal states ω^d. Thus, since ω^d is independent of $\omega^{d'}$ for all $d' \neq d$, we can sum over the conditional expectation, conditional on $\omega^d \in \Omega^d$. Hence, we obtain (by dropping the index d from ω^d)

$$\begin{aligned} V^d &= \sum_{\omega=1}^{\overline{\omega}} \pi(\omega) \sum_{n=n^*+1}^{N} P\left[X^d = x_n | \omega\right] \left(E\left[X^d\right] - E\left[Y^d | X^d = x_n, \omega\right]\right) \\ &= \sum_{\omega=1}^{\overline{\omega}} \pi(\omega) \sum_{n=n^*+1}^{N} h^d_{n\omega} \left(E\left[X^d\right] - E\left[Y^d | \omega\right]\right), \end{aligned}$$

where, in the second line, we have used that $P\left[X^d = x_n|\omega\right] = h^d_{n\omega}$, and the fact that $E\left[Y^d|X^d = x_n, \omega\right] = E\left[Y^d|\omega\right]$. Now, define $t^d_\omega = E\left[X^d\right] - E\left[Y^d|\omega\right]$, and $s^d_\omega = \sum_{n=n^*+1}^{N} h^d_{n\omega}$. Thus,

$$V^d = \sum_{\omega=1}^{\overline{\omega}} \pi(\omega)\, s^d_\omega t^d_\omega.$$

Note that $t^d_\omega = \sum_{n=1}^{N} h^d_{n\omega} \left(E\left[X^d\right] - x_n\right)$. Therefore, since c is a λ-mixture of a and b:

$$s^c_\omega = \lambda s^a_\omega + (1-\lambda)\, s^b_\omega, \quad \text{and} \quad t^c_\omega = \lambda t^a_\omega + (1-\lambda)\, t^b_\omega.$$

Using this in V^c and multiplying out, we obtain:

$$\begin{aligned} V^c &= \sum_{\omega=1}^{\overline{\omega}} \pi(\omega) \left(\lambda^2 s^a_\omega t^a_\omega + (1-\lambda)^2 s^b_\omega t^b_\omega + \lambda(1-\lambda)\left(s^a_\omega t^b_\omega + s^b_\omega t^a_\omega\right)\right) \\ &= \lambda^2 V^a + (1-\lambda)^2 V^b + \lambda(1-\lambda) \sum_{\omega=1}^{\overline{\omega}} \pi(\omega) \left(s^a_\omega t^b_\omega + s^b_\omega t^a_\omega\right). \end{aligned} \tag{7}$$

We can now show that $V^c \geq V^b$. Indeed, since $V^a \geq V^b$ by assumption, we can estimate the first two terms in (7) against $\left(\lambda^2 + (1-\lambda)^2\right) V^b$, and it follows that $V^c \geq V^b$ if the last sum in (7) is larger than $2V^b$. In fact, we show that it is larger than $V^a + V^b \geq 2V^b$, that is, we show:

$$\sum_{\omega=1}^{\overline{\omega}} \pi(\omega) \left(s^a_\omega t^b_\omega + s^b_\omega t^a_\omega\right) \geq \sum_{\omega=1}^{\overline{\omega}} \pi(\omega) \left(s^a_\omega t^a_\omega + s^b_\omega t^b_\omega\right).$$

Indeed, by the previous expression is equivalent to

$$\sum_{\omega=1}^{\overline{\omega}} \pi(\omega) \left(s^a_\omega - s^b_\omega\right) \left(t^b_\omega - t^a_\omega\right) \geq 0. \tag{8}$$

Now suppose that h^a_ω first order stochastically dominates h^b_ω. Since x_n is increasing in n, this implies that $t^a_\omega \leq t^b_\omega$. Observe that we can write $s^d_\omega = \sum_{n=1}^{N} h^d_{n\omega} \cdot 1_{\{n>n^*\}}(n)$, and since $1_{\{n>n^*\}}(n)$ is an increasing function in n, it is also true that $s^a_\omega \geq s^b_\omega$. If h^b_ω first order stochastically dominates h^a_ω, the argument is reversed. Thus, with (ii) (a), we have for all ω:

$$s^a_\omega \geq s^b_\omega \Longleftrightarrow t^a_\omega \leq t^b_\omega.$$

But this implies (8).

To complete the proof, it remains to show that $V^c > V^b$ for some $\lambda \in (0,1)$. But this follows from (ii) (b), which ensures that at least one term under the sum in (8) is strictly positive. This establishes the claim. □

Proof of Lemma 1 We proceed in two steps. In Step 1, we show that (i) implies

$$E\left[Y^d | X^d = x_n\right] > E\left[X^d\right] \Longleftrightarrow x_n > x_{n^*}. \tag{9}$$

In Step 2, we show that if (i) and (ii) hold for $d = a, b$, then they hold for $d = c$. Thus, steps 1 and 2 imply the first condition in Proposition 1.

As for Step 1. We suppress the action index d. Let

$$\widehat{\pi}_{\omega n} = \frac{h_{n\omega}\pi(\omega)}{p(x_n)}$$

be the posterior probability of state ω conditional on $X = x_n$. For $p(x_n) = 0$, we set $\widehat{\pi}_{\omega n} = \pi_\omega$.[24] We denote the posterior distribution on Ω conditional on $X = x_n$ by the $\overline{\omega} \times 1$ vector $\widehat{\pi}_n$.

We now proceed in two steps. In step (A), we show that if $h_{n\omega}$ is strictly increasing in ω, then the posterior $\widehat{\pi}_n$ strictly first order stochastically dominates the prior π. In step (B), we show that (c) implies that the function

$$E[Y|\omega]$$

is increasing in ω. Thus, it follows by the strict dominance relation established in (A) that

$$\sum_\omega E[Y|\omega]\,\widehat{\pi}_{\omega n} > \sum_\omega E[Y|\omega]\,\pi(\omega)$$

if $h_{n\omega}$ is strictly increasing in ω. But note that by the law of iterated expectation, the left hand side is equal to $E[Y|X = x_n]$, and the right hand side is equal to $E[X]$. If $h_{n\omega}$ is decreasing in ω, an identical argument applies with all inequalities being weak and reversed. Thus, (i) implies (9).

As for (A). The argument is the same as in Milgrom (1981), proof of Proposition 1.

As for (B). Let $\omega > \theta$, we have to show that

$$E[Y|\omega] - E[Y|\theta] = \sum_{n=1}^{N} x_n (h_{n\omega} - h_{n\theta}) \geq 0.$$

To see this, observe first that since $h_{n\omega} - h_{n\theta}$ is increasing in n by assumption (ii), there is a unique $\widetilde{n}$ such that $h_{n\omega} - h_{n\theta} \leq 0$ if and only if $n \leq \widetilde{n}$. Hence, since x_n is increasing in n, it follows for all n:

$$x_n (h_{n\omega} - h_{n\theta}) \geq x_{\widetilde{n}} (h_{n\omega} - h_{n\theta}).$$

[24] This is consistent with our convention that $E\left[Y^d | X^d = x_n\right] = E\left[X^d\right]$ for $p(x_n) = 0$.

Observe second that $\sum_{n=1}^{N}(h_{n\omega}-h_{n\theta})=0$. Together, the two observations imply that

$$\sum_{n=1}^{N} x_n (h_{n\omega}-h_{n\theta}) \geq x_{\tilde{n}} \sum_{n=1}^{N} (h_{n\omega}-h_{n\theta}) = 0.$$

This completes Step 1.

As for Step 2. Note that since $h^c_{n\omega}$ is the convex combination of $h^a_{n\omega}$ and $h^b_{n\omega}$, monotonicity properties carry over to $h^c_{n\omega}$. This completes the proof. □

Proof of Proposition 2 We begin by writing the conditions (i) and (ii) Definition 3 in matrix notation. For x_n with $p^d(x_n)>0$, Bayes' rule yields that

$$\widehat{\pi}^d_{\omega n} = \frac{h^d_{n\omega}\pi(\omega)}{p^d(x_n)} \tag{10}$$

is the posterior probability that the agent assigns to state ω conditional on having observed $X^d = x_n$. Define the vectors

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}, \quad p^d = \begin{pmatrix} p^d(x_1) \\ \vdots \\ p^d(x_N) \end{pmatrix}, \quad \pi = \begin{pmatrix} \pi(1) \\ \vdots \\ \pi(\overline{\omega}) \end{pmatrix}, \quad \widehat{\pi}^d_n = \begin{pmatrix} \widehat{\pi}^d_{1n} \\ \vdots \\ \widehat{\pi}^d_{\overline{\omega}n} \end{pmatrix}.$$

We can thus write

$$p^d = h^d\pi^d, \quad E\left[X^d\right] = x^T p^d,$$

where x^T is the transposed of x. We summarise all possible posterior beliefs in the $\overline{\omega}\times\overline{x}$ Markov matrix

$$\widehat{\pi}^d = \left(\widehat{\pi}^d_{\omega n}\right)_{\omega,n}. \tag{11}$$

Thus, the (posterior) probability with which $Y^d = x_m$, conditional on $X^d = x_n$, is given by

$$\widehat{p}^d_{mn} = \sum_{\omega\in\Omega} h^d_{m\omega}\widehat{\pi}^d_{\omega n}.$$

We denote the conditional distribution of Y^d, conditional on $X^d = x_n$, by the $N\times 1$ vector $\widehat{p}^d_n$. Thus, we can write the posterior mean as

$$E\left[Y^d | X^d = x_n\right] = x^T \widehat{p}^d_n.$$

By defining the $N\times N$ matrix

$$\widehat{p}^d = \left(\widehat{p}^d_1, ..., \widehat{p}^d_N\right) = h^d\widehat{\pi}^d,$$

we can express the $N\times 1$ vector of posterior means as

$$\left(E\left[Y^d | X^d = x_n\right]\right)_n = x^T\widehat{p}^d.$$

We can now write the conditions (i) and (ii) from Definition 3 in matrix form:

$$x^T \widehat{p}^c = x^T \widehat{p}^a \tau, \quad \text{(i')}$$

$$p^a = \tau p^c. \quad \text{(ii')}$$

We now prove (i'). Evidently, (i') is implied by the condition $\widehat{p}^c = \widehat{p}^a \tau$, or equivalently,

$$h^c \widehat{\pi}^c = h^a \widehat{\pi}^a \tau. \quad \text{(i")}$$

To establish (i"), we first compute $\widehat{\pi}^d$. Since the prior is strongly symmetric, i.e. $\sigma\pi = \pi$, and since $h^b = h^a \sigma$ by assumption, we have

$$p^b = h^b \pi = h^a \sigma \pi = h^a \pi = p^a.$$

Moreover, $p^c = h^c \pi = \lambda h^a \pi + (1 - \lambda) h^b \pi = p^a$. Hence, we have shown that $p^a = p^b = p^c$. Let $p = p^a$. By (10), it holds for all $d \in D$ that $\widehat{\pi}^d_{\omega n} = \left(h^d_{xn} \pi(\omega)\right) / p(x_n)$ if $p(x_n) > 0$. Since by assumption $p(x_n) > 0$ for all n, we can write

$$\widehat{\pi}^d = \delta_\pi \left(h^d\right)^T \delta_p^{-1},$$

where δ_v is the $k \times k$ matrix with diagonale elements v_i for $i = 1, ..., k$ and off-diagonale elements 0 for an arbitrary $v \in \mathbb{R}^k$. Thus, with $h^c = \lambda h^a + (1 - \lambda) h^b$, we obtain

$$\widehat{\pi}^c = \delta_\pi \left(\lambda h^a + (1 - \lambda) h^b\right)^T \delta_p^{-1} = \lambda \widehat{\pi}^a + (1 - \lambda) \widehat{\pi}^b.$$

Because $h^b = h^a \sigma$, we have

$$\widehat{\pi}^b = \delta_\pi \left(h^b\right)^T \delta_p^{-1} = \delta_\pi \left(h^a \sigma\right)^T \delta_p^{-1} = \delta_\pi \sigma^T \left(h^a\right)^T \delta_p^{-1} = \sigma \delta_\pi \left(h^a\right)^T \delta_p^{-1} = \sigma \widehat{\pi}^a.$$

Therefore, we can write the left hand side of (ii') as

$$\begin{aligned} h^c \widehat{\pi}^c &= \left[\lambda h^a + (1 - \lambda) h^b\right] \left[\lambda \widehat{\pi}^a + (1 - \lambda) \widehat{\pi}^b\right] \\ &= \left[\lambda h^a + (1 - \lambda) h^a \sigma\right] \left[\lambda \widehat{\pi}^a + (1 - \lambda) \sigma \widehat{\pi}^a\right] \\ &= \lambda^2 h^a \widehat{\pi}^a + 2\lambda (1 - \lambda) h^a \sigma \widehat{\pi}^a + (1 - \lambda)^2 h^a \sigma \sigma \widehat{\pi}^a. \end{aligned}$$

Since $\sigma\sigma$ is the identity matrix, and since $h^a \sigma = \sigma h^a$ by strong symmetry, this implies

$$\begin{aligned} h^c \widehat{\pi}^c &= h^a \widehat{\pi}^a \left(\left[\lambda^2 + (1 - \lambda)^2\right] \delta + 2\lambda (1 - \lambda) \sigma\right) \\ &= h^a \widehat{\pi}^a \tau, \end{aligned}$$

where δ is the identity matrix, and $\tau = \left[\lambda^2 + (1 - \lambda)^2\right] \delta + 2\lambda (1 - \lambda) \sigma$ is a transition matrix satisfying $\sum_n \tau_{nm} = 1$, $\tau_{nm} \geq 0$ for all m, n, and this establishes (i").

As for (ii'). We first show that $\sigma p^a = p^a$. Indeed, since $p^a = h^a \pi^a$,

$$\sigma p^a = \sigma h^a \pi^a = h^a \sigma \pi^a = h^a \pi^a = p^a,$$

where the third and fourth equality follows since h^a and π are strongly symmetrc by assumption. Thus,

$$\tau p^a = \left(\left[\lambda^2 + (1-\lambda)^2\right]\delta + 2\lambda(1-\lambda)\sigma\right)p^a = \left(\lambda^2 + (1-\lambda)^2 + 2\lambda(1-\lambda)\right)p^a = p^a.$$

Because $p^c = p^a$, this proves (ii'). □ .

Proof of Lemma 3 We have to show that $E[Y^a|X^a = x_n] \neq E[Y^c|X^c = x_n]$ for some n. Recall that we can write $\left(E\left[Y^d|X^d = x_n\right]\right)_n = x^T\widehat{p}^d$, and that we have $x^T\widehat{p}^c = x^T\widehat{p}^a\tau$ by assumption. Thus, we have to show that $x^T\widehat{p}^a \neq x^T\widehat{p}^a\tau$.

We drop the action index. With $\tau = \left[\lambda^2 + (1-\lambda)^2\right]\delta + 2\lambda(1-\lambda)\sigma$, it follows that $x^T\widehat{p} = x^T\widehat{p}\tau$ if and only if $x^T\widehat{p} - x^T\widehat{p}\sigma = 0$ and $\lambda \in (0,1)$. Hence, we have to show that $x^T\widehat{p} - x^T\widehat{p}\sigma \neq 0$. Recall that $x^T\widehat{p} = x^T h\widehat{\pi}$. A tedious but straightforward computation yields that $\widehat{\pi}\sigma = \sigma\widehat{\pi}$. This together with $h\sigma = \sigma h$ delivers:

$$x^T\widehat{p} - x^T\widehat{p}\sigma = x^T h\widehat{\pi} - x^T\sigma h\widehat{\pi} = \left(x^T - x^T\sigma\right)h\widehat{\pi}.$$

Recall that $\widehat{\pi} = \delta_\pi h^T \delta_p^{-1}$. Since $\pi(\omega) > 0$ for all ω by assumption, δ_π has rank $\overline{\omega}$. Likewise, δ_p^{-1} has rank N. Hence, since h has rank N by assumption, it follows that $\widehat{\pi}$ has rank $\min\{N, \overline{\omega}\} = N$, and consequently, the $N \times N$ matrix $h\widehat{\pi}$ has rank N. Therefore, $\left(x^T - x^T\sigma\right)h\widehat{\pi} = 0$ if and only if $x^T - x^T\sigma = 0$. But since $x = (x_n)_n$ is increasing in n by assumption, the latter cannot be true. This establishes the claim. □

Example distinguishing our model from HF's model The following example illustrates that there are settings, in which our model generates uncertainty aversion while HF's model does not. There are two equally likely (marginal) states and five outcomes. We set

$$h^a = \frac{1}{10}\begin{pmatrix} 1.0 & 1.5 \\ 3.0 & 0.5 \\ 4.0 & 4.0 \\ 0.5 & 3.0 \\ 1.5 & 1.0 \end{pmatrix}, \quad h^b = \frac{1}{10}\begin{pmatrix} 1.5 & 1.0 \\ 0.5 & 3.0 \\ 4.0 & 4.0 \\ 3.0 & 0.5 \\ 1.0 & 1.5 \end{pmatrix}, \quad h^c = \frac{1}{10}\begin{pmatrix} 1.25 & 1.25 \\ 1.75 & 1.75 \\ 4.00 & 4.00 \\ 1.75 & 1.75 \\ 1.25 & 1.25 \end{pmatrix}.$$

Actions a and b clearly satisfy the strong symmetry conditions of Proposition 2 and c is a .5-mixture of a and b. Thus, in a pairwise choice between c and either a or b our agent will certainly prefer urn c to either a or b. However, actions a and b do not satisfy HF's requirement that corresponding states of a and b be ranked by first order stochastic dominance.

Denote by Z^d the sum of outcomes when d is chosen twice and played out conditionally independently. (In the notation of our model: $Z^d = X^d + Y^d$.) According to HF's model, in a choice between action a and action c, any agent with a concave utility function will prefer the mixture c to a if and only if Z^c second order stochastically dominates Z^a, which is equivalent to

$$\sum_{z=2}^{\widetilde{z}} P\left[Z^a \leq z\right] \leq \sum_{z=2}^{\widetilde{z}} P\left[Z^c \leq z\right] \quad \text{for all } \widetilde{z} \in \{2, ..., 10\}.$$

Now notice that $P\left[Z^a = 2\right] = (1/2) \times (1/10)^2 + (1/2) \times (15/100)^2 = 0.01625$ and $P\left[Z^c = 2\right] = (125/100)^2 = 0.015625$. Thus,

$$\sum_{z=2}^{2} P\left[Z^a \leq z\right] = 0.01625 > 0.015625 = \sum_{z=2}^{z} P\left[Y^c \leq z\right].$$

Similarly, we obtain that

$$\sum_{z=2}^{3} P\left[Z^a \leq z\right] = 0.07000 < 0.075000 = \sum_{z=2}^{3} P\left[Z^c \leq z\right].$$

Thus, Z^c and Z^a are not ranked by second order stochastic dominance, so it is not true that any risk-averse agent will necessarily prefer c to a. □

References

Ahn, D. S. (2008). "Ambiguity without state space." *Review of Economic Studies*, 75(1): 3-28..

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

Bell, D. E. (1982). "Regret in decision making under uncertainty." *Operations Research*, 30: 961-981.

Bell, D. E. (1983). "Risk premiums for decision regret." *Management Science*, 29: 1156-1166.

Brun, W. and K. H. Teigen (1990). "Prediction and postdiction preferences in guessing." *Journal of Behavioral Decision Making*, 3: 17-28.

Curley, S. P., F. Yates and R. A. Abrams (1986). "Psychological Sources of Ambiguity Avoidance." *Organizational Behavior and Human Decision Processes*, 38: 230-256.

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

Ergin, H. and F. Gul (2008). "A subjective theory of compound lotteries." mimeo.

Frisch, D., and J. Baron (1988). "Ambiguity and rationality." *Journal of Behavioral Decision Making*, 1: 149-157.

Fischhoff, B. (1975). "Hindsight ≠ foresight: The effect of outcome knowledge on judgment." *Journal of Experimental Psychology*, 1: 288-299.

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989). "Psychological games and sequential rationality." *Games and Economic Behavior*, 1: 60-79.

Ghirardato, P., Maccheroni, F., and Marincacci, M. (2004). "Differentiating ambiguity and ambiguity attitude." *Journal of Economic Theory*, 118; 133-173.

Gilboa, I. and D. Schmeidler (1988). "Maxmin expected utility with a non-unique prior." *Journal of Mathematical Economics*, 18: 141-153.

Gilovich, T. and V. H. Medvec (1994). "The temporal pattern to the experience of regret." *Journal of Personality and Social Psychology*, 67(3): 357-365.

Halevy, Y. and V. Feltkamp (2005). "A Bayesian approach to uncertainty aversion." *Review of Economic Studies*, 72: 449-466.

Halevy, Y. (2007). "Ellsberg revisited: an experimental study." *Econometrica*, 75(2): 503-536.

Harless, D. (1992). "Actions versus prospects: the effect of problem representation on regret." *American Economic Review*, 82: 634-649.

Hayashi, T. (2008). "Regret aversion and opportunity dependence." *Journal of Economic Theory*, 139(1): 242-268.

Kahneman, D. and D. T. Miller (1986). "Norm theory: Comparing reality to its alternatives." *Psychological Review*, 93: 136-153.

Kahneman, D. and A. Tversky (1979). "Prospect theory: An analysis of decision under risk." *Econometrica*, 47: 263-291.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). "A smooth model of decision making under ambiguity." *Econometrica*, 73(6): 1849-1892.

Koszegi, B. and M. Rabin (2006). "A model of reference-dependent preferences." *Quarterly Journal of Economics*, 121: 1133-1165.

Krähmer, D. and R. Stone (2008). "Regret in sequential decision problems." mimeo

Loomes, G. and R. Sugden (1982). "Regret theory: An alternative approach to rational choice under uncertainty." *The Economic Journal*, 92: 805-824.

Mellers, B. (2000). "Choice and the relative pleasure of consequences." *Psychological Bulletin*, 126: 910-924.

Mellers, B., A. Schwartz and I. Ritov (1999). "Emotion-based choice." *Journal of Experimental Psychology: General*, 128: 332-345.

Mellers, B. and A. P. McGraw (2001). "Anticipated emotions as guides to choice." *Current Directions in Psychological Science*, 10(6): 210-214.

Milgrom, P. (1981). "Good news and bad news: Representation theorems and applications." *Bell Journal of Economics*, 13: 380-391.

Milnor, J. (1954). "Games against nature." In R. R. Thrall, C. H. Coombs, and R. L. Davis (Eds), *Decision Processes* (pp. 49-60), New York: Wiley.

Morris, S. (1997). "Risk, uncertainty, and hidden information." *Theory and Decision*, 42: 235-269.

Nau, R. F. (2006). "Uncertainty aversion with second-order utilities and probabilities." *Management Science*, 52(1): 136-145.

Rabin, M. (1998). "Psychology and Economics." *Journal of Economic Literature*, 36: 11-46.

Ritov, I. and J. Baron (2000). "Reluctance to vaccinate: omission bias and ambiguity." In C. R. Sunstein (ed.), *Behavioral Law and Economics*, (pp. 168-186). New York: Cambridge University Press.

Rothschild, M. and J. E. Stiglitz (1970). "Increasing Risk." *Journal of Economic Theory*, 2: 225-243.

Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.

Schmeidler, D. (1989). "Subjective probability and expected utility without additivity." *Econometrica*, 57: 571-587.

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

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

Starmer, C., and R. Sugden (1993). "Testing for juxtaposition and event-splitting effects." *Journal or Risk and Uncertainty*, 6: 235-254.

Starmer, C., and R. Sugden (1998). "Testing alternative explanations for cyclical choice." *Economica*, 65: 347-361.

Tversky, A., and Kahneman, D. (1982). "Causal schemas in judgments under uncertainty." In D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 117-128). New York: Cambridge University Press.

Zeelenberg, M. (1999). "Anticipated regret, expected feedback, and behavioral decision making." *Journal of Behavioral Decision Making*, 12: 93-106.