A Bayesian Approach to Uncertainty Aversion[1]

Yoram Halevy
University of British Columbia [2]

Vincent Feltkamp
Maastricht School of Management[3]

June 8, 2004

[1]We are indebted to the Editor, James Dow, and two anonymous referees, who contributed important insights and detailed remarks. We appreciate the helpful comments of Robert Aumann, Lawrence Blume, Stephen Coate, David Easley, Jürgen Eichenberger, Dean Foster, Drew Fudenberg, Itzhak Gilboa, Gordon Hazen, Kenneth Hendricks, Peter Klibanoff, George Mailath, Mark Machina, Steven Matthews, Stephen Morris, Nicola Persico, Andrew Postlewaite, Rafi Rob, Ariel Rubinstein, David Schmeidler, Hans Schneeweiss, Uzi Segal, Dan Silverman, Becky Stein, Guofu Tan, Peter Wakker and John Weymark. We also thank seminar participants at University of Pennsylvania, Cornell, Duke, University of British Columbia, Hebrew University, Tel Aviv University, Haifa University, Ben Gurion University, Workshop on "New Themes in Decision Theory Under Uncertainty" held in Paris on June 1999, Université de Montréal, The Canadian Economic Theory Conference held at UBC on June 2000, UC-Berkeley and Stanford University.

[2]Department of Economics, #997-1873 East Mall, Vancouver BC V6T 1Z1 Canada.yhalevy@interchange.ubc.ca, http://www.econ.ubc.ca/halevy.
Yoram Halevy thanks the Department of Economics at the University of Pennsylvania for its hospitality while working on this paper.

[3]PO box 1203, NL-6201 BE Maastricht, The Netherlands. Feltkamp@msm.nl
This paper was written during Vincent Feltkamp's stay at Departmento de Fundamentos del Análisis Económico, Alicante University, Spain. Funding from the TMR project ERB FMRX CT96 0055 is gratefully acknowledged.

Abstract

The Ellsberg paradox demonstrates that people's belief over uncertain events might not be representable by subjective probability. We show that if a risk averse decision maker, who has a well defined Bayesian prior, perceives an Ellsberg type decision problem as possibly composed of a bundle of several positively correlated problems - she will be uncertainty averse. We generalize this argument and derive sufficient conditions for uncertainty aversion.

JEL classification: D81

Keywords: Ellsberg paradox, rule rationality, ambiguity aversion, risk aversion, subjective probability, reduction of compound lotteries.

1 Introduction

Daniel Ellsberg's (1961, [8]) experiments demonstrate that for many individuals *risk* (known probabilities) and *uncertainty* (or *ambiguity* - unknown probabilities) are two different notions. Ellsberg's examples are direct criticism of Savage's [33] normative conception that uncertainty may be treated similarly to risk, when subjective probability, which is derived from preferences, replaces the objective probability in the von Neumann-Morgenstern theory of expected utility. In fact, the Ellsberg paradox is inconsistent with Mark Machina and David Schmeidler's "probabilistically sophisticated" preferences [24] that generalize the idea of deriving subjective probability from preferences. The existence of subjective probability is critical in Economics, where its usage is pervasive. In many cases, not only do the results depend on the existence of subjective probability, but without it defining the relevant problem would become much more difficult (if not impossible).

Consider Ellsberg's "Two Urns" problem: there are two urns, each containing 100 balls, which can be either red or black. It is known that the first urn holds 50 red and 50 black balls. The number of red (black) balls in the second urn is unknown. Two balls are drawn at random, one from each urn. The subject is asked to bet on the color of one of the balls. A correct bet wins her $100, an incorrect guess loses nothing (and pays nothing). The modal response exhibits *uncertainty (ambiguity) aversion:* the decision maker prefers a bet on red or black drawn from the first urn to a bet on red or black drawn from the second urn, but she is indifferent between betting on red or black in each urn separately.

In this paper we consider a perturbation of the original experiment suggested by Ellsberg, in which more than a single ball (a bundle) may be drawn from each urn. We prove that in this *regular* environment, a risk averse decision maker, who holds a Bayesian prior over possible states of the world, and has to choose on which urn to bet, will be *uncertainty averse*. Furthermore, if the decision maker does not know with certainty the structure of the environment (that is, if a single ball or a bundle will be drawn from each urn), any small probability of a *regular* environment will lead to a decision that exhibits uncertainty aversion. The explanation bounds the premium the individual is willing to pay in order to discard uncertainty in favor of risk.

To relate our perturbed environment to the actual paradox, we use the framework of *Rule Rationality,* which was suggested, among others, by Ronald Heiner [21] and Robert Aumann [2]. This paradigm claims that people's decision making has evolved to simple *rules* that perform well in most *regular* (common) environments. Heiner [21] argues that rules arise because an agent has limited cognitive abilities to identify the most preferred alternative in every environment. Hence, she faces endogenous uncertainty in choosing the optimal alternative and, under some conditions, is better off restricting her flexibility to simple alternatives that function relatively well in most environments. Although Heiner was motivated by Axelrod's findings in the repeated Prisoners' Dilemma, his claims are much more general. It should be emphasized, however, that although Heiner presents "rule rationality" as a case of "bounded rationality," this interpretation is not required for the current paper. We only show that the rule of being uncertainty averse is rational in the bundled (regular) environment, and do not derive an uncertainty averse rule as a constrained rational choice. The application of the rule to the standard Ellsberg paradox may be a result of bounded rationality (as Heiner argues) or just irrational (due to inertia or error). Another prominent advocate of "rule rationality" is Aumann [2], who restricts attention to repeated interactions and contrasts strategies in repeated games with strategies in the one shot game (what he calls *Act Rationality*). Motivated by empirical studies of the Ultimatum Game (Güth et al [18]; Binmore et al [4]), Aumann argues that the rule of rejecting low offers has been determined in an evolutionary process. This process rewards a behavior that utilizes a rule which works well in most environments, i.e. it is optimal for a regular (in Aumann's terminology - repeated) environment. When applying the decision rule to a singular (in Aumann's terminology - one shot) environment[1], the behavior may be hard to rationalize.

The *regular* environment considered in this paper consists of a bundle of several positively correlated risks. We argue that environments in which people make decisions under uncertainty are frequently *regular*. An example of a decision in such an environment is the purchase of a car. Suppose the decision maker cares about the payoff distribution of the repair cost during the first year after purchasing a car. These costs are the *sum* of repair costs of the different components of the car. The repair cost of each component is risky, but the risk that every component will malfunction during the first year, depends on the state of the car (which depends, for example, on previous

[1]Either because the individual applies a decision rule which is already "hard wired" into her decision making for similar (regular) environments, or she does not understand the singularity of the basic environment.

owners). The better the state of the car, the lower the probability that each component will need repair. Hence, the repair costs of different components are positively correlated. The decision is whether to buy the car (including all its components) and to face the uncertain aggregate repair cost, or not. Our metaphor for a risky environment is an environment in which the agent knows the state of the car, and faces the randomness implied by mechanics. In an uncertain (ambiguous) environment, the agent does not know for sure the state of the car. She may have a prior belief over the state of the car, but we show it does not collapse to the risky environment since one decision (to buy the car) spans multiple risks that are correlated through the state of the world (car). The above argument could be easily adapted to many other decision problems, such as purchasing a house, getting married, choosing a new working place and becoming a member in club.

In the following section, we present our resolution to Ellsberg's "two urns" paradox. Next, we generalize the example and establish formally the relation between behavioral rules and uncertainty aversion, viz., we derive conditions under which uncertainty aversion may be rationalized as a Bayesian rule in an environment consisting of bundled risks. The paper concludes with a discussion of the results, a comparison to the current literature on uncertainty aversion and bounded rationality, and a conjecture concerning the relation between uncertainty aversion and other behavioral anomalies.

2 A Bayesian Resolution of Ellsberg's Paradox

This section demonstrates how the concept of "rule rationality" could be applied to the famous Ellsberg's paradox, which motivates a substantial part of the literature on uncertainty aversion. We use the "Two Urns" example, which was presented in the Introduction. The "Single Urn" (with three colors) example [8] could be treated similarly. Note that we use some simplifying assumptions that are not necessary (the more general case is analyzed in Section 3).

The decision maker - Alice - has learned from experience (though maybe not consciously) that some circumstances are not isolated (singular), but that frequently similar risks are bundled. The *regular* environment in which she evaluates uncertain prospects consists of bundled risks. When asked which bet she prefers, she applies the *rule* that has evolved in this regular-bundled environment. Our goal is to characterize the regular environment and analyze the preferences the decision maker has in this environment. The original Ellsberg experiment constitutes the singular environment in this paradigm. For simplicity of the initial exposition, we assume the regular environment consists of two Ellsberg singular experiments, which are perfectly correlated. There are two type I urns (risky), and two type II urns (ambiguous). By perfect correlation, it is meant here that the two urns have the same color composition. Alice's choice set consists of betting on one color from the (two) risky urns, or on one color from the (two) uncertain urns[2]. Alice's payoff is the sum of her payoff in each draw.

The distribution of the monetary prize if Alice bets on red (or black) from the urns with a known probability of $\frac{1}{2}$ (urns of type I) is:

$$IR_{(2)} = IB_{(2)} = \begin{cases} \$0 & 1/4 \\ \$100 & 1/2 \\ \$200 & 1/4 \end{cases} \tag{1}$$

When considering the ambiguous urns, Alice might[3] apply the statistical principle of *insufficient reason*[4]. Therefore, she has a prior belief over the number of red balls contained in them, which assigns a probability of $\frac{1}{101}$ to every frequency between 0 and 100 (thus p, the proportion of red balls in the ambiguous urns, is between 0 and 1). Conditional on p, the probability that two red balls would be drawn from the ambiguous urns (i.e. winning \$200 if betting on red) is p^2, the probability of two black balls (i.e. winning \$0 if betting on red) is $(1-p)^2$, and the probability of one red ball and one black ball (i.e. a total prize of \$100 if betting on red) is $2p(1-p)$. According to the Bayesian paradigm, Alice should average these values over the different p in the support of her prior belief. Hence the probability of winning \$200

[2] Alternatively, two balls will be drawn (with replacement) from each urn.

[3] None of the results depend on this assumption. As will be clear from section 3, all that is required is that Alice will be indifferent between betting on red or black from the type II urns. This is guaranteed by any symmetric prior.

[4] The principle of insufficient reason states that if one does not have a reason to suspect that one state is more likely than the other, then by symmetry the states are equally likely, and equal probabilities should be assigned to them. The reader is referred to Savage [33] Chapter 4 section 5 for a discussion of the principle in relation to subjective probability.

and \$0 is:

$$\sum_{i=0}^{100} \frac{1}{101} \left(\frac{i}{100}\right)^2 = \sum_{i=0}^{100} \frac{1}{101} \left(1 - \frac{i}{100}\right)^2 \cong \int_0^1 p^2 dp = \frac{1}{3} \tag{2}$$

Therefore, the expected (according to the uniform prior) distribution of the monetary payoff from betting on the ambiguous urns is:

$$IIR_{(2)} = IIB_{(2)} = \begin{cases} \$0 & 1/3 \\ \$100 & 1/3 \\ \$200 & 1/3 \end{cases} \tag{3}$$

It follows that $IR_{(2)}$ and $IB_{(2)}$ *second order stochastically dominate* $IIR_{(2)}$ and $IIB_{(2)}$ (i.e. the latter two are mean preserving spreads of the former)[5]. If Alice is averse to mean preserving spreads, she will prefer to bet on the risky urns. Furthermore, if her preferences are represented by an expected utility functional (with respect to an additive probability measure), then aversion to mean preserving spreads is a consequence of risk aversion. Therefore, if Alice is *risk averse* she will prefer a bet on the objective urns to a bet on the ambiguous urns, and will exhibit uncertainty (ambiguity) aversion, as observed in the Ellsberg experiment. If she is a risk lover, she will prefer the latter to the former, and exhibit uncertainty love (also predicted behavior by Ellsberg); whereas if she is risk neutral, she will be indifferent between the four bets.

In the case of two draws and a uniform prior, but without dependence on her risk aversion, Alice will prefer to bet on the ambiguous urns, rather than bet on red from type I urns that contain anything less than 43 red balls. The distribution of a bet on red from the type I urns that contain only 42 red balls is:

$$IR_{(2)}\left(p = \frac{42}{100}\right) = (\$0, 0.3364; \$100, 0.4872; \$200, 0.1764) \tag{4}$$

Hence, a bet on the uncertain urns would *first order stochastically dominate* a bet on red from these risky urns. Thus the *uncertainty premium* (in terms of

[5] For formal definitions of first and second order stochastic dominance see [29] and Appendix A.

probabilities) is bounded from above by 8%. In monetary terms, this upper bound is equivalent to \$16[6]:

$$E\left(IB_{(2)}\left(p=\frac{1}{2}\right)\right)-E\left(IB_{(2)}\left(p=\frac{42}{100}\right)\right)=\$100-\$84=\$16 \qquad (5)$$

The only assumption relied upon in this argument is monotonicity of the preference relation with respect to first and second order stochastic dominance. Therefore, this explanation is consistent with *any* theory of choice under risk that exhibits aversion to mean preserving spreads, including expected utility with diminishing marginal utility of wealth, as well as most non-expected utility theories of choice under risk.

The logic developed above extends to regular environments composed of any number of bundled risks. Assume Alice compares the distribution of betting on r concurrent IR (IB) to r concurrent IIR (IIB) as in the Ellsberg experiment. The money gained is distributed $100X$ where X has a binomial distribution with parameters $(0.5, r)$ and (p, r), respectively. If p, the proportion of red balls in the ambiguous urns, is distributed uniformly on $[0, 1]$, then for every $0 \leq k \leq r$: [7]

$$\begin{aligned}\Pr\{X=k\} &= \binom{r}{k}\frac{1}{101}\sum_{s=0}^{100}\left(\frac{s}{100}\right)^{k}\left(1-\frac{s}{100}\right)^{r-k} \cong \qquad (6)\\ &\cong \binom{r}{k}\int_0^1 p^k(1-p)^{r-k}\,dp=\binom{r}{k}Beta\left(k+1,r-k+1\right)=\\ &=\frac{r!}{k!(r-k)!}\frac{k!(r-k)!}{(r+1)!}=\frac{1}{r+1}\end{aligned}$$

That is, the expected distribution of $IIR_{(r)}$ and $IIB_{(r)}$ is uniform, and is second order stochastically dominated by the binomial $IR_{(r)}$ and $IB_{(r)}$.

[6] These bounds depend on the uniform prior assumption. Assuming only symmetry of the prior, the lower bound on the number of red balls in the type I urn would be 29.

[7] The *Beta Integral* is defined by:
$Beta\,(m+1, n+1) = \int_0^1 p^m(1-p)^n dp = \frac{\Gamma(m+1)\Gamma(n+1)}{\Gamma(m+n+2)}$
Where $\Gamma(\alpha) = \int_0^\infty p^{\alpha-1}e^{-p}dp$ for $\alpha > 0$, and it is a well known result that when k is a natural number: $\Gamma(k) = (k-1)!$

The only relation between the two ambiguous risks needed to justify uncertainty aversion is a positive correlation. Let p_1 and p_2 be the relative frequencies of red balls in the first and second ambiguous urns, respectively. It is simple to verify that if $Corr\left(p_1, p_2\right) > 0$ then $E\left(p_1 p_2\right) = E\left(\left(1-p_1\right)\left(1-p_2\right)\right) > \frac{1}{4}$, and therefore a bet on the ambiguous urns is a mean preserving spread of a bet on the risky (known probabilities of 0.5) urns.

Note that Alice does not need to assign probability one to the regular (bundled) experiment in order to prefer a bet on the risky urns. In most cases we do not know (or do not understand) with certainty the environment in which we have to make decisions. Alice might have learned from her experience that some risks are bundled, but some are isolated. Even if the probability of a correlated risk is very small, she would prefer a bet on the risky (type I) urns. This is a consequence of a "Sure Thing Principle" argument: if there is only a singular risk, she is indifferent between betting on urn I or urn II, and in the case of bundling, she strictly prefers the former. Hence the conclusion that she prefers risk over ambiguity, even when she faces the slightest possibility of a regular environment. Thus, in the case of environmental uncertainties, the paradoxical Ellsberg choices may be fully rationalized.[8]

3 The General Framework

The natural framework to generalize Ellsberg's examples is Anscombe-Aumann's [1] horse bets over roulette lotteries, in which objective and subjective probabilities coexist. In this section we define the regular environment which consists of bundled acts. We prove that if a decision maker is risk averse, her preferences among bundled acts would exhibit "uncertainty aversion" (Schmeidler [34]).

3.1 Uncertainty Aversion

Let $\mathcal{X}$ be a finite set of monetary *outcomes,* $\mathcal{R}$ the set of finitely supported *(roulette) lotteries* over $\mathcal{X}$, and assume a preference ordering over $\mathcal{R}$ that

[8]Note, however, that as the probability of regular environment decreases, the uncertainty premium will decrease as well.

satisfies the usual expected utility assumptions. Therefore, there exists a von Neumann-Morgenstern utility function $u(\cdot)$, such that lottery ρ_1 is preferred to lottery ρ_2 if and only if $\sum_{x\in\mathcal{X}} \rho_1(x)u(x) > \sum_{x\in\mathcal{X}} \rho_2(x)u(x)$. Let S be a finite (non-empty) set of *states of the world.* In Ellsberg's "Two Urns" example, states of the world represent the number of red balls in the second urn: $S = \{0, ..., 100\}$. An *act (horse lottery)* is a function from S to $\mathcal{R}$. That is, it is a compound lottery, in which the prizes are roulette lotteries. Let $\mathcal{H}$ denote the set of acts. Define a convex combination over elements of $\mathcal{H}$ as a pointwise mixture. That is, for every $f, g \in \mathcal{H}$ and $0 \leq \alpha \leq 1$, the holder of $(f, \alpha; g, 1-\alpha)$ will receive in every state $s \in S$ the compound lottery $(f(s), \alpha; g(s), 1-\alpha)$. Assume preferences over $\mathcal{H}$ satisfy independence (Schmeidler [34]). As a result, if the decision maker is indifferent between f and g, then she is indifferent between the two and the lottery $(f, \alpha; g, 1-\alpha)$. An example of such statewise mixture in the "two urns" example is the compound lottery $\left(IIR, \frac{1}{2}; IIB, \frac{1}{2}\right)$. Assuming the decision maker abides by the Reduction of Compound Lotteries Axiom (Segal [35], [36]), it is easy to verify that this compound lottery is equal to betting on IR. Since Alice is indifferent between IIR and IIB, but prefers IR to either, her preferences in Ellsberg's example violate at least one of the assumptions: Reduction of Compound Lotteries (Segal [35]) or Independence over $\mathcal{H}$ (Schmeidler [34]).

Schmeidler [34] was the first to define uncertainty aversion, using the Anscombe-Aumann framework. Formally:

Definition 1 (Schmeidler [34]) *A decision maker is* Uncertainty Averse *if, for each pair of acts f and g, f indifferent to g implies that every convex combination of f and g is preferred to f (and to g).*

In Schmeidler's [34] model of Choquet expected utility, this can be strict only for acts that are non-comonotonic, as defined below:

Definition 2 (Schmeidler [34]) *Two acts f and g are* comonotonic *if for no $s, s' \in S$: $f(s) \succ f(s')$ and $g(s') \succ g(s)$.*

In the context of Choquet expected utility it would be reasonable to define a decision maker to be strictly uncertainty averse if she prefers any convex combination of every two non-comonotonic acts f and g, between which she is indifferent, to f and g. In Ellsberg's two urns example, IIR and IIB are not comonotonic since the higher the number of red balls in the second urn, IIR becomes more favorable and IIB becomes less favorable. Hence, strict preference of IR ($= \left(IIR, \frac{1}{2}; IIB, \frac{1}{2}\right)$) to IIR is an evidence of strict uncertainty aversion.

It should be noted that the same definition of uncertainty aversion is employed by Gilboa and Schmeidler [15] as one of their axioms in deriving the Maximin Expected Utility representation. However, in the MEU representation uncertainty aversion may be strict even for some comonotonic acts (for a characterization of the set see Ghirardato, Klibanoff and Marinacci [12]).

3.2 The *Regular* Environment

Uncertainty averse behavior is explained intuitively as the agent "hedging" between two acts. However, in the Ellsberg examples, there are opportunities for "hedging" that are in some sense stronger than those entailed by non-comonotonicity alone. In these experiments, the lotteries assigned by IIR and IIB are ranked according to First Order Stochastic Dominance criterion in every state in which they differ. That is, *every* agent with monotone preferences would prefer $IIR(s)$ to $IIB(s)$ if $51 \leq s \leq 100$ and $IIB(s)$ to $IIR(s)$ if $0 \leq s \leq 49$. Hence, we can compare the agent's utility from different acts at a specific state. Therefore, the hedging behavior could be interpreted as more fundamental, and independent of the agent's utility function. This distinction is critical in the framework of "bundled acts."

Let $\mathcal{X}$, $\mathcal{R}$, S and $\mathcal{H}$ be defined as above.

Definition 3 *Acts f and g in $\mathcal{H}$ are* Statewise Ranked by First Order Stochastic Dominance *if $f \neq g$ and at every state s in which they differ $f(s)$ First Order Stochastically Dominates (FOSD) $g(s)$ or vice versa.*

We prove that if preferences are defined over bundled acts in the *regular* environment (with more than a single lottery at every state), a seemingly uncertainty averse behavior emerges, when the original acts are Statewise Ranked by FOSD.

Definition 4 *A* Bundled Act $f_{(r)}$ *is a function from S to the sum (convolution) of $r > 1$ independent and identical lotteries over outcomes. The set of all bundled acts is the* Regular Environment *and is denoted by $\mathcal{H}_{(r)}$.*

Note, that according to Definition 4, the set of acts, $\mathcal{H}$, constitutes the *Singular Environment* in this setting. In the regular environment, every state, s, is assigned a "bundle" of lotteries. In the formal definition, we assume that conditional on the state, lotteries are independent and identically distributed. That is, the bundle consists of r independent draws from one lottery (denoted by $f(s)$). To relate Definition 4 to our resolution of the Ellsberg experiment presented above, note that a bundled act (in the regular environment) bundles a bet on all the type II or type I urns. The condition that the lotteries are conditionally (on the state) independent and identically distributed is a generalization of the "same color composition" in the type II urns. For example, the bundled act $IIR_{(2)}$ assigns to every state (frequency of red balls in the type II urns) the sum of two independent draws from the ambiguous urns. Relating to the car example presented in the introduction, the regular environment captures the idea that for a given car condition (state) the risk associated with the state of the transmission is independent of the risk associated with the state of the engine. That is, the correlation is generated by the state of the car. The dimensionality of the regular environment is indexed by r. Consider the agent's preferences over the regular environment. She is indifferent between the bundled acts $f_{(r)}$ and $g_{(r)}$ if:

$$U\left(f_{(r)}\right) = U\left(g_{(r)}\right) \tag{7}$$

Denote by $q(s)$ the subjective probability of state s. Then (7) can be written explicitly as:

$$\sum_{s \in S} q(s) E\left[u\left(f_{(r)}(s)\right)\right] = \sum_{s \in S} q(s) E\left[u\left(g_{(r)}(s)\right)\right] \tag{8}$$

where $E\left[u\left(f_{(r)}(s)\right)\right]$ is the agent's expected utility from the sum of r (objective) lotteries that f assigns to state s. In what follows we take $r = 2$ (it will be sufficient to produce uncertainty averse behavior). Then:

$$E\left[u\left(f_{(2)}(s)\right)\right] = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{X}} f(s)(x) f(s)(y) u(x+y) \tag{9}$$

where $f(s)(x)$ and $f(s)(y)$ are the probabilities of outcomes x and y respectively, according to the objective lottery $f(s)$.

The following Theorem gives a generalization of our main result. If the acts satisfy Definition 3, as the Ellsberg examples do, and preferences are defined over the regular environment (i.e. bundled acts), "uncertainty aversion" is a consequence of a Bayesian prior and risk aversion.

Theorem 1 *If f and g are Statewise Ranked by FOSD and the agent is indifferent between the bundled act $f_{(2)}$ and the bundled act $g_{(2)}$, then if she is averse to mean preserving spreads and her preferences are representable by an expected utility functional, she will prefer the bundled act of $(f, \alpha; g, 1-\alpha)_{(2)}$ over the bundled act $f_{(2)}$ for every $0 < \alpha < 1$.*

Proof. See Appendix. ■

To gain intuition that motivates the Theorem, let $h_{(2)}$ be $(f, \alpha; g, 1-\alpha)_{(2)}$. That is, the bundled act where in state s the decision maker receives two independent draws from the lottery $(f(s), \alpha; g(s), 1-\alpha)$. The two draws from the lottery h will both come from f with probability α^2 and both from g with probability $(1-\alpha)^2$. Since $f_{(2)} \sim g_{(2)}$, the agent's expected utility from $h_{(2)}$ conditional on either event is equal to her conditional expected utility from $f_{(2)}$. Hence the comparison between $h_{(2)}$ and $f_{(2)}$ hinges entirely on whether h is better or worse conditional on the event that one draw comes from f and one from g. Since f and g are Statewise Ranked by FOSD, one draw from each distribution is less risky (on average) than two draws from one, so every risk averse agent will prefer h.

The implication of Theorem 1 is that if the perception of a risk averse agent is that a decision will span multiple ambiguous risks, and the acts satisfy the condition of Statewise Ranking by FOSD, then her observed behavior would exhibit uncertainty aversion.

Uncertainty averse behavior may be fully rationalized if the individual assigns a small probability that the environment she is facing is regular. The source of this belief is the agent's experience that some environments are regular and some are singular. Confronted with a new situation, if the individual's heuristic belief assigns some (possibly small) probability to the possibility she faces a regular environment, then her optimal behavior would exhibit uncertainty aversion.

Corollary 1 *Assume f and g as in Theorem 1, and suppose the individual is indifferent between the acts f and g too. Then, for every $\beta > 0$ probability of a regular environment, she will prefer a lottery between the two acts (or bundled acts - with probability β) over each act (or bundled act - with probability β).*

Proof. Since $(f, \alpha; g, 1-\alpha)_{(2)} \succ f_{(2)}$ and $(f, \alpha; g, 1-\alpha) \sim f$, it follows from the independence axiom that:

$$\left[(f,\alpha;g,1-\alpha)_{(2)},\beta;(f,\alpha;g,1-\alpha),1-\beta\right] \succ \left[f_{(2)},\beta;f,1-\beta\right] \ \blacksquare$$

The Corollary may be interpreted as a learning argument in the development of a rule. Since the agent is indifferent between the two singular acts f and $(f,\alpha;g,1-\alpha)$, the bundled acts $f_{(2)}$ and $(f,\alpha;g,1-\alpha)_{(2)}$ serve as "tie-breaking". Hence if the agent develops one rule to decide in similar environments (where the regular and the singular environments are considered subjectively similar), this rule will choose $(f,\alpha;g,1-\alpha)$.

3.3 Are the conditions necessary?

Theorem 1 shows that when f and g are statewise ranked by FOSD then preferences over bundled acts will exhibit uncertainty aversion. The following example shows that when this condition is not satisfied, uncertainty aversion or uncertainty loving among bundled acts may result (depending on the specific utility function). Hence, this condition alone is not necessary for uncertainty aversion among bundled acts. It is left for future research to fully characterize preferences on this domain.
Let the utility function be:

$$u(x)=\begin{cases} x & x\leq\gamma \\ \gamma & x>\gamma \end{cases} \tag{10}$$

for some $\gamma>0$. Assume two states of the world s,t with equal subjective probability. The two acts f, g are:

$$f(s)=g(t)=\begin{cases} 3 & 0.5 \\ 2 & 0.5 \end{cases} \qquad f(t)=g(s)=\begin{cases} 4 & 0.5 \\ 1 & 0.5 \end{cases} \tag{11}$$

The two acts are non-comonotonic (the state-lotteries are ranked by second order stochastic dominance) for $1<\gamma<4$ and the individual is indifferent between them. Therefore, uncertainty aversion would claim she prefers the mixture of the two over each act separately. However, a short calculation shows that our explanation of preference over bundled acts may or may not support uncertainty aversion in this case, depending on the parameter γ. If $1<\gamma\leqslant 2$ the individual is indifferent between $f_{(2)}$ and $(f,0.5;g,0.5)_{(2)}$, while if $2<\gamma<4$ the individual prefers the latter bundled act to the former (that is, exhibits uncertainty aversion).
A utility function that exhibits strict uncertainty loving for the above acts is

given by:

$$u(x) = \begin{cases} x & x \leqslant 3 \\ \frac{x+3}{2} & 3 < x < 5 \\ 4 & x \geqslant 5 \end{cases} \tag{12}$$

Here, $U\left(f_{(2)}\right) > U\left((f, 0.5; g, 0.5)_{(2)}\right)$.

The intuition that motivates the above examples is that in the absence of statewise ranking by FOSD, diminishing marginal utility of wealth does not impose enough restrictions on the preference over bundled acts to imply ambiguity aversion.

4 Discussion and Conclusion

This work shows that a perturbation of the Ellsberg paradox's environment leads to uncertainty averse behavior which is consistent with expected utility theory and Bayesian rationality. If one uses "rule rationality," then human behavior may exhibit insensitivity to the details of the environment, and uncertainty aversion becomes a very plausible prediction even in the standard environment.

4.1 Comparison with the Literature

The Ellsberg paradox motivated an extensive literature that tried to explain this predicted behavior. In this section we shall discuss only few alternative resolutions.

The Maximin Expected Utility (MEU) model, which was axiomatized by Gilboa and Schmeidler [15] and Casadesus-Masanell, Klibanoff and Ozdenoren [6], derives from individual's preferences a convex set of priors. The decision maker chooses the act that maximizes her expected utility if the worst prior, included in the set of priors, occurs (Maximin over a convex set of priors). Note that MEU does not imply extreme pessimism, since the set of priors itself is endogenously derived from preferences. Hence conservatism in the Maximin framework is measured by the size of this set. For example, the set $[0, 1]$ corresponds to extreme pessimism, while smaller sets correspond to more moderate conservatism. Schmeidler [34] and Gilboa [14] derived the Choquet expected utility representation, which is a special case of the Maximin if the capacity is convex. Uncertainty Aversion was first defined in this

context. We point out that the preferences over bundled acts suggested in this paper, and the MEU are two distinct representations, and are not equivalent. The following thought experiment may sharpen the difference (beyond the example in the previous section):
Suppose a third urn containing 100 balls (red or black) is added to the original two urns in the Ellsberg example. The composition of this urn is determined by lottery that assigns probabilities $0 \leqslant \pi_j \leqslant 1$ that the number of red balls is $j = 0, \ldots, 100$ and $\sum_{j=0}^{100} \pi_j = 1$. Furthermore, assume π is symmetric, that is: $\pi_{100-j} = \pi_j$. The subject is asked to bet on the color of a ball drawn, before she knows the result of the lottery π. Note that urn III is completely objective and is composed of two-stage lotteries. According to MEU, the decision maker should be indifferent between betting on the first urn (known 50-50 composition) and the third, and as long as the set of priors is symmetric and non-singleton, a bet on either should be preferred over a bet on the ambiguous (second) urn. This is a result of the "reduction of compound lotteries" assumption, included in the expected utility treatment of objective uncertainty (risk) within MEU. According to the theory of preferences over bundled acts proposed here, the subject will rank the first urn highest (as long as the subjective and the objective priors are not a point mass on a composition of 50-50), and then rank urns II and III according to the dispersions of $\mathbf{q}$ (the subjective prior on urn II) and π (the objective prior on urn III). For example, if $\pi_0 = \pi_{100} = 0.5$, all risk averse individuals will weakly prefer the ambiguous urn over the third urn. The above predictions may enable us to compare empirically between the theories.

As shown in the above thought experiment, the theory presented in this paper allows for an aversion to known second order probabilities, through the bundling effect. This main feature of our theory is present in Uzi Segal's [35] work as well. He analyzes ambiguous prospects as two-stage lotteries (similar to the framework here): first a probability is chosen according to some prior belief distribution, and then a second lottery is performed. Segal relaxes the Reduction of Compound Lotteries Axiom, and replaces the one stage Mixture Independence with Compound Independence (Segal [36]). This allows him to consider utility functions that are more general than expected utility. An ambiguous lottery is evaluated by replacing each second-stage lottery with its certainty equivalent. Segal shows that for Anticipated Utility[9], risk aversion and reasonable restrictions on the transformation of probabilities

[9]Similar analysis could be done for other non-expected utility functions.

function, may rationalize the Ellsberg paradox. Hence, both theories share the causation between risk aversion and uncertainty aversion. However, these are different explanations: Segal's theory relies on non-expected utility. Both the second and first stage lotteries are evaluated according to a non-expected utility model. If the utility function would be linear in probabilities, the ambiguous lottery would give the same payoff as its expected risky counterpart. Under the theory presented here, the lottery at each state is replaced by the sum (convolution) of two conditionally independent lotteries, which we call regular act. Even if the convolution is evaluated using expected utility function (as in Section 3), the convolution operator itself makes the evaluation of the second stage non-linear in probabilities, and leads to the violation of the reduction axiom. More specifically, the bundling effect causes the payoff at the second stage to be a quadratic function of the probabilities:

$$U\left(IIR_{(2)}\middle|\, p\right) = p^2\left[u\left(200\right) - 2u\left(100\right) + u\left(0\right)\right] + \\ + 2p\left[u\left(100\right) - u\left(0\right)\right] + u\left(0\right) \tag{13}$$

Monotonicity and risk aversion imply that this is a concave function of p. Since the first stage is evaluated using expected utility (which is linear in p), the concavity in p implies that the decision maker will be averse to mean preserving spreads in p. If one calculates an expression similar to (13) for an arbitrary prize x, then $\partial U\left(IIR_{(2)}\middle|\, p\right)/\partial x = 2p^2 u'\left(2x\right) + 2p\left(1-p\right)u'\left(x\right)$. That is, the weight attached to the marginal utility at $2x$ is higher than the weight the consequence $2x$ receives in calculating (13). Furthermore, the second derivative of (13) assigns even a higher weight to $u''(2x)$.

To compare these findings to Segal, note that the expression corresponding to (13) in Segal is:

$$V\left(IIR|\, p\right) = v\left(0\right) + \left[v\left(100\right) - v\left(0\right)\right] f\left(p\right) \tag{14}$$

where $v\left(\cdot\right)$ is the decision maker's cardinal utility index, and $f\left(\cdot\right)$ is her decision weights function satisfying $f\left(0\right) = 0$ and $f\left(1\right) = 1$. It is easy to see that the curvature of (14) as a function of the prize is determined by the curvature of v at the prize. Unlike the bundling model, the first stage evaluation in Segal's model is not linear in probabilities. Segal [35] derived sufficient conditions on $f\left(\cdot\right)$ that will generate Ellsberg type behavior. Those conditions are slightly stronger than risk aversion in the theory of anticipated utility, and are related to conditions that can generate Allais type behavior using this functional form.

The two theories (Segal [35] and the bundling theory presented here) have different predictions for the thought experiment presented above (using a third urn). A decision maker who follows Segal's theory would be indifferent between the risky (first) urn and a third urn with extreme dispersion of $\pi_0 = \pi_{100} = 0.5$. Another decision maker who has preferences over bundled acts as this paper suggests, and is risk averse, will strictly prefer the first urn to the third extreme urn, and will weakly prefer the ambiguous urn to the extreme urn. In an experiment conducted by Halevy [19] he found strong experimental support for both pattern of choices in the population. Yates and Zukowski [41] considered a similar third urn with a uniform π. They found that, on average, subjects valued urn I more than urn III, and urn III more than urn II[10].

Recently, Ergin and Gul [11], Klibanoff et al [23] and Nau [27] developed Segal's [35] approach, focusing on the violation of the reduction axiom in explaining ambiguity aversion. Although the goals of the three papers are different they all assume that the decision maker has a prior belief over the possible realization of the probability distribution, but since she does not reduce probabilities between this stage and the second stage in which objective lotteries are performed, it allows to capture her ambiguity attitudes. Nau [27] emphasizes that the utility could be state dependent, Ergin and Gul [11] allow for probabilistically sophisticated preferences and Klibanoff et al [23] concentrate on expected utility valuation without imposing reduction.

It is of interest to study the degree of risk aversion implied by the proposed theory of bundling. For simplicity, consider an arbitrary lottery (with no ambiguity) p. Let $c(p)$ be the certainty equivalent of p defined by the implicit relation: $p \sim \delta_{c(p)}$ were δ_x is the degenerate lottery that pays x with certainty. Now consider bundling two such lotteries, and compare the certainty equivalent of the bundled lottery to the certainty equivalent of the original lottery [11]. If $u(\cdot)$ is exponential - that is, exhibits constant absolute

[10] Yates and Zukowski [41] averaged minimum selling price of a chosen lottery, for different individuals. Hence, their results involve interpersonal comparisons, and should be treated carefully.

[11] This problem is closely related to Samuelson's "Fallacy of Large Numbers" [32], which shows that if an expected utility maximizer rejects a certain bet at all wealth levels, she should reject any sum of independent repetitions of this bet. However, as pointed out by many authors (e.g. Ross [30]), allowing for a wealth effect may lead to the acceptance of the sum of bets. If, on the other hand, expected utility is relaxed, Chew and Epstein [7] have shown that acceptance of the sum of bets may be consistent with Weighted Utility and Anticipated Utility.

risk aversion, it can be shown that the certainty equivalent of the bundled (compound) lottery, $p_{(2)}$ is exactly twice the certainty equivalent of p. That is, bundling does not change the risk attitudes of the decision maker. However, if one allows for wealth effects, this will not be true anymore. That is, the risk attitudes of the agent will be a function of how "big" the bundle she evaluates is. This is an important line of research, which we plan to pursue in the future, since it relates ambiguity aversion (the current paper) to risk attitudes over a sequence of lotteries (Samuelson [32]).

Stephen Morris [25] takes a strategic approach, and argues that the unattractiveness of the ambiguous urn is a result of the asymmetry in information between the experimenter and the subject. This approach may rationalize a lower willingness to bet on one color from the ambiguous urn, but when bets on both colors are offered (as in the original Ellsberg example) the individual should behave non-strategically. Morris argues, along the lines of "rule rationality," that individuals utilize their experience from situations of asymmetric information in responding to Ellsberg's paradox.

Note that the formal model presented in this paper is silent as to whether the risks are bundled or repeated. In the latter case, the results could be interpreted as a "policy" of preferring a sequence of risky bets to a sequence of ambiguous bets. Hazen [20] shows (similarly to the example in Section 2) that a risk averse policy maker who is an expected utility maximizer and faces more than a single repetition of the Ellsberg problem will exhibit uncertainty aversion. Schneeweiss [39] analyzes the Ellsberg paradox assuming the number of repetitions approaches infinity and the utility function is quadratic[12]. Both works are limited to the Ellsberg example, and do not explore how general the result is[13]. Furthermore, the "policy" interpretation of the results is vulnerable to considerable limitations on the rules considered. For example, the decision maker cannot learn from one repetition to the next (for the optimal strategy in this case see Müller and Scarsini [26]), and can not alternate (hedge) between different ambiguous risks[14]. Hence, the decision maker is

[12]We thank Gordon Hazen and Hans Schneeweiss for bringing their works to our attention.

[13]Both works assume expected utility (and even more restrictive functional forms), while as shown in Section 2, the only requirment is aversion to mean preserving spreads.

[14]It is easy to show that if the decision maker faces repeated draws (with replacement) from the two Ellsberg urns, then even if she cannot learn (that is, has to have a "policy"), mixing between bets on the uncertain (II) urn - that is, choosing a policy of $(IIR, IIB, IIR, IIB, ...)$ - second order stochastically dominates bets on the risky (I) urn.

not rational even in the repeated environment.

Note that our notion of bounded rationality included in the "Rule Rationality" description is distinct from the Case Based Decision Theory (CBDT) studied by Gilboa and Schmeidler [16]. Their theory is aimed at describing situations where the state space is unknown to the decision maker. A case in their theory is described by the triplet: (problem, act, outcome). The decision maker evaluates each act by her average payoff when this act was taken at "similar" problems she can recall, weighted by how similar the problems are. Clearly, the theory presented here is not formally related to CBDT. Here, the decision maker knows the possible states of the world, and uses this information extensively. However, the notion of "similarity, " which is used in Rubinstein [31] as well, may be incorporated into the current model. The decision maker views the singular and the regular environments as similar, and it leads her to prefer risky acts over ambiguous acts in the singular environment as well.

4.2 Uncertainty Aversion

Schmeidler's [34] definition of uncertainty aversion (Definition 1), which was used by Gilboa and Schmeidler [15] as well, is nested within the Anscombe-Aumann [1] framework. This framework has some drawbacks due to its compound lottery structure. Casadesus-Masanell, Klibanoff and Ozdenoren [6], who present an axiomatization of Maximin Expected Utility in a completely subjective world without lotteries, provide an analogue to Schmeidler's definition of uncertainty aversion - without objective lotteries.

Other definitions of uncertainty aversion, which differ from Schmeidler's, have appeared in the literature. Epstein [9] defines uncertainty aversion relative to probabilistically sophisticated preferences, while Ghirardato and Marinacci [13] define it relative to subjective expected utility. A one-stage axiomatization of expected utility, that allows for objective lotteries, was suggested by Sarin and Wakker [38], but the definition of uncertainty aversion in their original framework is not transparent and will be different [37] from Schmeidler's. Klibanoff, Marinacci and Mukerji [23] employ this framework while relaxing the Reduction of Compound Lotteries axiom, to define *smooth ambiguity aversion* as an aversion to mean preserving spreads in the ex-ante

Hence, risk averse decision maker who faces a sequence of draws will prefer to bet using the uncertain urn to bets using the risky urn.

evaluation of an act (similar to risk aversion in objective expected utility). The exact way in which the behavior described in this paper relates to those alternative definitions, remains for future work.

4.3 "Rule Rationality" and Other Experimental Anomalies

An underlying feature of the explanation presented here is that individuals treat a single draw from an Ellsberg urn (act) the same way they treat multiple draws from the urn (bundled act). Many studies in psychology have focused on how individuals update their belief. Although there is no updating of belief per se in this paper, we believe that there is a close connections between the two phenomena[15]. Tversky and Kahneman [40] first noted that people consistently overestimate the distributional similarity between a small sample and the population, and named this behavior: "the law of small numbers." This goes both ways - from the sample to the population, and vice versa. Bar Hillel and Wagenaar [3] concentrate on "local representativeness" bias, when people expect even a short sequence of signals to have the same proportions of signals as a much longer (or infinite) sequence. Grether [17] and Camerer [5] test if individuals and markets are Bayesian and find support for bias in a direction of "exact representativeness," in which agents tend to believe that the (unknown) population's distribution is similar to the small sample's distribution. These observations were modeled and applied by Rabin [28] to a variety of economic scenarios. These behavioral regularities may illuminate our current study and give our explanation an alternative motivation: if we reinterpret the prior as representing a "population," then the risky population (generated by urn I) second order stochastically dominates the uncertain population (generated by urn II). If the decision maker exhibits "local representativeness" she expects these relative properties of the distributions to be maintained even for a small sample (in our case of size one), and hence will be uncertainty averse. We believe that the relation between the two phenomena requires further experimental study.

As discussed in the Introduction, two other prominent experimental anomalies, that initially seem unrelated to uncertainty aversion, are the one-shot "Prisoners' Dilemma" and the "Ultimatum Game". In the first example, almost all normative notions of equilibrium (except when agents have unob-

[15] We thank the Editor James Dow for pointing out this relation.

served utility from cooperation) predict that individuals will not cooperate. Yet, in practice, many subjects do indeed cooperate. In the Ultimatum Game, the normative backward induction argument predicts that the individual who makes the offer will leave a minimal share to his opponent, and the latter will accept any positive offer. In practice, most offers are "fair," and most respondents reject "unfair" (albeit positive) splits. Explanations for these phenomena vary, but the one explanation we find most compelling (and which may be viewed as a strategic basis for other explanations), claims that people do not "understand" that these are one-shot games. Individuals play a strategy which is perfectly reasonable (according to some equilibrium notion) for a repeated game. Thus, people are, in some sense, not "programmed" for, and therefore find it hard to evaluate, singular situations. Aumann [2] contrasted this *"Rule Rationality"* with "Act Rationality". Hoffman, MacCabe and Smith [22] have suggested that in the Ultimatum Game, the rule to "reject anything less than thirty percent" may be rationalized as building up a reputation in an environment where the interaction is repeated. This rule does not apply to the one-shot Ultimatum Game because in that situation the player does not build up reputation. But since the rule has been unconsciously chosen, it will not be consciously abandoned[16].

The (speculative) relation between the decision theoretic problem studied in this paper, and other anomalies in game theory, leads us to hypothesize that rule rationality is a form of limited rationality that should be studied carefully. Specifically, experiments could determine whether certain individuals rely more than others on behavioral rules. If rule rationality is found to be common, it may call for reconsidering the structure of experiments in economics and psychology. Currently, most of the experimental literature identifies a singular environment as a good experimental design, since it enables concentration on a specific issue. However, if individuals use in this environment their experience from more "regular" environments, the designer should consider whether the behavior in the experiment is robust to small perturbations of the environment. An evolutionary model in which "rule rationality" emerges may illuminate the set of procedures for which this notion of limited rationality is viable.

[16] It may be argued that "manners" have evolved in a similar way, and explain the Proposer behavior in Dictator Games as a result of expected "reciprocity".

References

[1] ANSCOMBE, F. J. and AUMANN, R. J. (1963), "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199-205.

[2] AUMANN, R. J. (1997), "Rationality and Bounded Rationality," *Games and Economic Behavior*, 21, 2-14.

[3] BAR-HILLEL, M. and WAGENAAR, W. A. (1991), "Perception of Randomness," *Advances of Applied Mathematics*, 12, 428-454.

[4] BINMORE, K., SHAKED, A. and SUTTON, J. (1985), "Testing Noncooperative Bargaining Theory: A Preliminary Study," *American Economic Review*, 75, 1178-1180.

[5] CAMERER, C. F. (1987), "Do Biases in Probability Judgements Matter in Markets? Experimental Evidence," *American Economic Review*, 77 (5), 981-997.

[6] CASADESUS-MASANELL, R., KLIBANOFF, P. and OZDENOREN, E. (2000), "Maxmin Expected Utility over Savage Acts with a Set of Priors," *Journal of Economic Theory*, 92, 35-65.

[7] CHEW, S. H. and EPSTEIN, L. R. (1988), "The Law of Large Numbers and the Attractiveness of Compound Gambles," *Journal of Risk and Uncertainty*, 1, 125-132.

[8] ELLSBERG, D. (1961), "Risk, Ambiguity and the Savage Axioms," *The Quarterly Journal of Economics,* 75 (4), 643–669.

[9] EPSTEIN, L. G. (1999), "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66, 579-608.

[10] EPSTEIN, L. G. and ZHANG, J. (2001), "Subjective Probabilities on Subjectively Unambiguous Events," *Econometrica*, 69 (2), 265-306.

[11] ERGIN, H. and GUL, F. (2004), "A Subjective Theory of Compound Lotteries," mimeo.

[12] GHIRARDATO, P., KLIBANOFF P. and MARINACCI, M. (1998), "Additivity with Multiple Priors," *Journal of Mathematical Economics*, 30, 405-420.

[13] GHIRARDATO, P. and MARINACCI M. (2002), "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102 (2), 251-289.

[14] GILBOA, I. (1987), "Expected Utility with Purely Subjective Non-Additive Probabilities," *Journal of Mathematical Economics,* 16, 65–88.

[15] GILBOA, I. and SCHMEIDLER, D. (1989), "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics,* 18, 141–153.

[16] GILBOA, I. and SCHMEIDLER, D. (1995), "Case-Based Decision Theory," *The Quarterly Journal of Economics*, 110 (3), 605-639.

[17] GRETHER, D. M. (1980), "Bayes Rule as a Descriptive Model: The Representativeness Heuristic," *The Quarterly Journal of Economics*, 95 (3), 537-557.

[18] GÜTH, W., SCHMITTBERGER, R. and SCHWARZE, B. (1982), "An Experimental Study of Ultimatum Bargaining," *Journal of Economic Behavior and Organization*, 3, 367-388.

[19] HALEVY, Y. (2004), "Ambiguity Aversion and Reduction: an Experiment on Heterogeneous Population," (mimeo, University of British Columbia).

[20] HAZEN, G. B. (1992), "Decision versus Policy: An Expected Utility Resolution of the Ellsberg Paradox," in: *Decision Making Under Risk and Uncertainty: New Models and Empirical Finding*, J. Geweke (Ed.), Kluwer Academic Publishers, Dordrecht, The Netherlands.

[21] HEINER, R. A. (1983), "The Origin of Predictable Behavior," *The American Economic Review,* 73 (4), 560-595.

[22] HOFFMAN, E., MCCABE, K. and SMITH, V. L. (1996), "Social Distance and Other-Regarding Behavior in Dictator Games," *The American Economic Review*, 86 (3), 653-660.

[23] KLIBANOFF, P., MARINACCI, M. and MUKERJI, S. (2004), "A Smooth Model of Decision Making Under Ambiguity," mimeo.

[24] MACHINA, M. J. and SCHMEIDLER, D. (1992), “A More Robust Definition of Subjective Probability,” *Econometrica*, 60 (4), 745-780.

[25] MORRIS, S. (1997), “Risk, Uncertainty and Hidden Information,” *Theory and Decision*, 42, 235-269.

[26] MÜLLER, A. and SCARSINI, M. (2002), “Even Risk-Averters May Love Risk,” *Theory and Decision,* 52, 81-99.

[27] NAU, R. F. (2002), “Uncertainty Aversion with Second Order Probabilities and Utilities,” mimeo.

[28] RABIN, M. (2002), “Inference by Believers in The Law of Small Numbers,” *The Quarterly Journal of Economics*, 117 (3), 775-816.

[29] ROTHSCHILD, M. and STIGLITZ, J. E. (1970), “Increasing Risk: I. A Definition,” *Journal of Economic Theory*, Vol. 2 (3), 225-243.

[30] ROSS, S. A. (1999), “Adding Risks: Samuelson’s Fallacy of Large Numbers Revisited,” *Journal of Financial and Quantitative Analysis*, 34 (3), 323-339.

[31] RUBINSTEIN, A. (1988), “Similarity and Decision-making under Risk (Is There a Utility Theory Resolution to the Allais Paradox?),” *Journal of Economic Theory*, 46, 145-153.

[32] SAMUELSON, P. A. (1963), “Risk and Uncertainty: A Fallacy of Large Numbers,” *Scientia*, 98, 108-113.

[33] SAVAGE, L. J. (1954) *The Foundations of Statistics* (New York: John Wiley and Sons). Revised and enlarged edition (1972): (New York: Dover).

[34] SCHMEIDLER, D. (1989), “Subjective Probability and Expected Utility without Additivity,” *Econometrica,* 57 (3), 571–587.

[35] SEGAL, U. (1987), “The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach,” *International Economic Review*, 28 (1), 175-202.

[36] SEGAL, U. (1990), “Two-Stage Lotteries Without The Reduction Axiom,” *Econometrica*, 58 (2), 349-377.

[37] SARIN, R. and WAKKER, P. (1992), "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica*, 60 (6), 1255-1272.

[38] SARIN, R. and WAKKER, P. (1997), "A Single-Stage Approach to Anscombe and Aumann's Expected Utility," *Review of Economic Studies*, 64, 399-409.

[39] SCHNEEWEISS, H. (1999), "Resolving the Ellsberg paradox by Assuming that People Evaluate Repetitive Sampling," in: *Modelling and Decisions in Economics: Essays in Honor of Franz Ferschi*, Edited by: G. Feichtinger, F. Ferschi, K-P Kistner and U. Leopold-Wildburger, Physica Verlag.

[40] TVERSKY, A. and KAHNEMAN, D. (1971), "Belief in the Law of Small Numbers," *Psychological Bulletin*, 78, 105-110.

[41] YATES, F. J. and L. ZUKOWSKI, G. (1976), "Characterization of Ambiguity in Decision Making," *Behavioral Science*, 21, 19-25.

A Preliminaries

Let ψ and τ be countably additive and finite set-functions on $\mathcal{X}$. Define:

$$F_{\psi}(x) = \sum_{t \leq x} \psi(t) \text{ and } F_{\tau}(x) = \sum_{t \leq x} \tau(t) \tag{15}$$

Assume ψ and τ are such that:

$$F_{\psi}(+\infty) = F_{\tau}(+\infty) \tag{16}$$

Assumption (16) would hold true if, for example, ψ and τ are probability measures (then (16) is equal to one), or when each is a difference of two probability measures (then (16) is equal to zero).

Definition 5 *Let ψ and τ be countably additive and finite set-functions on $\mathcal{X}$, and let F_{ψ} and F_{τ} be defined as in (15) and satisfy (16). The function ψ* First Order Stochastic Dominates (FOSD) *the function τ if for every $x \in \mathcal{X}$: $F_{\psi}(x) \leq F_{\tau}(x)$ with strict inequality for at least one x.*

Definition 5 is a generalization of the standard definition of first order stochastic dominance, and it includes the probability measure as a special case. It is well known that every decision maker with monotone preferences, choosing between two distributions ordered by FOSD, will prefer the dominant one.
Assume[17]:

$$\int_{-\infty}^{+\infty} F_{\psi}(x)dx = \int_{-\infty}^{+\infty} F_{\tau}(x)dx \tag{17}$$

That is, the mean of ψ is equal to the mean of τ. For example, if ψ is the difference of two probability measures and $\tau \equiv 0$ then it implies that the two probability distributions from which ψ was derived have the same expected value.

Definition 6 ψ Second Order Stochastically Dominates (SOSD) τ *if (17) holds and:*

$$\int_{-\infty}^{x} F_{\psi}(t)dt \leq \int_{-\infty}^{x} F_{\tau}(t)dt \quad \forall\ x \in \mathcal{X}$$

with strict inequality for at least one x.

Claim 1 *If ψ SOSD τ then:*

$$U(\psi) = \sum_{x\in\mathcal{X}} u(x)\psi(x)dx > \sum_{x\in\mathcal{X}} u(x)\tau(x)dx = U(\tau)$$

for all strictly monotone and strictly concave u.

Proof. The proof is similar to Rothschild and Stiglitz's [29]: using (16) instead of assuming probability measures, and (17) instead of assuming equal expectations. ■

B Proof of Theorem 1

Let f and g be statewise ranked by FOSD, and:

$$U\left(f_{(2)}\right) = U\left(g_{(2)}\right) \tag{7'}$$

[17] Since all set-functions we shall deal with have finite variation, all the integrals converge.

Therefore, there exist at least two states in which f and g differ. Define for every $s \in S$:

$$h(s)\,(x) = \alpha f(s)\,(x) + (1-\alpha)\,g(s)\,(x) \tag{18}$$

Then we need to show that:

$$U(h_{(2)}) > U\left(f_{(2)}\right) \tag{19}$$

Consider the function θ defined as:

$$\theta(s)\,(x) = f\,(s)\,(x) - g\,(s)\,(x) \tag{20}$$

for every x and s.
Let $h_{(2)}$ be the convolution (denoted by '$*$') of h with h at every state. $U\left(h_{(2)}\right)$ is the expected utility from this convolution, averaged over all states.

$$\begin{aligned} U\left(h_{(2)}\right) &= \sum_s q(s) U\left[h\,(s) * h\,(s)\right] = \\ &= \sum_s q(s) \sum_x \sum_y \begin{bmatrix} \alpha f(s)\,(x) + \\ (1-\alpha)\,g\,(s)\,(x) \end{bmatrix} \begin{bmatrix} \alpha f(s)\,(y) + \\ (1-\alpha)\,g\,(s)\,(y) \end{bmatrix} u(x+y) = \end{aligned}$$

$$= \sum_s q(s) \sum_x \sum_y \begin{bmatrix} \alpha^2\,(f(s)(x))\,(f(s)(y)) + \\ +\,(1-\alpha)^2\,(g\,(s)\,(x))(g\,(s)\,(y)) + \\ +2\alpha(1-\alpha)(f(s)(x))(g(s)(y)) \end{bmatrix} u(x+y) \tag{21}$$

Let $\theta_{(2)}$ be the convolution of θ with θ at every state. We can view $U\left(\theta_{(2)}\right)$ as the "expected utility" from this convolution (note that it is additive in the states):

$$U\left(\theta_{(2)}\right) = \sum_s q(s) U\left[\theta(s) * \theta(s)\right] = \tag{22}$$

$$\begin{aligned} &= \sum_s q(s) \sum_x \sum_y \theta\,(s)\,(x)\,\theta\,(s)\,(y)\,u(x+y) = \\ &= \sum_s q(s) \sum_x \sum_y \left[f(s)(x) - g\,(s)\,(x)\right]\left[f(s)(y) - g\,(s)\,(y)\right] u(x+y) = \end{aligned}$$

$$= \sum_s q(s) \sum_x \sum_y \begin{bmatrix} (f(s)(x))(f(s)(y)) + \\ +(g(s)(x))(g(s)(y)) - \\ -2(f\,(s)\,(x))(g\,(s)\,(y)) \end{bmatrix} u(x+y) \tag{23}$$

By substitution of (21) and (23) and utilizing (7') it follows that:

$$U(h_{(2)}) - U\left(f_{(2)}\right) = -\alpha\,(1-\alpha)\,U\left(\theta_{(2)}\right) \tag{24}$$

Thus, (19) holds if and only if $U\left(\theta_{(2)}\right) < 0$.

Claim 2 *In every state in which f and g differ: $\theta(s)$ FOSD $\mathbf{0}$ (the zero function) or vice versa.*

Proof. Since f and g are statewise ranked by FOSD, then if they differ at state s, they are ranked according to FOSD. Assume $f(s)$ FOSD $g(s)$. Then:

$$F_{\theta(s)}(x) = F_{f(s)}(x) - F_{g(s)}(x) \leq 0$$

The symmetric argument holds when $g(s)$ FOSD $f(s)$. ■

Lemma 1 *Let ξ be a function, which is the difference of two probability mass measures and assume ξ and $\mathbf{0}$ are ranked according to first order stochastic dominance. Then ξ can be written as a finite sum of functions:*

$$\xi = \sum_{l=1}^{L} \xi_l \tag{25}$$

where:

$$\xi_l(x) = \xi_{a_l,b_l,p_l}(x) = \begin{cases} p_l & \text{if} \quad x = a_l \\ -p_l & \text{if} \quad x = b_l \\ 0 & \quad\quad OTHERWISE \end{cases} \tag{26}$$

with $a_l < b_l$ and $|p_l| \leq 1$. If $\mathbf{0}$ FOSD ξ (ξ FOSD $\mathbf{0}$) then all p_l can be chosen positive (negative) in the decomposition (26).

Proof. Recall that since ξ is a difference of probability mass measures, it is a finite set function with $F_\xi(+\infty) = 0$. Assume $\mathbf{0}$ FOSD ξ, i.e.: $F_\xi(x) \geq 0$ $\forall$ $x \in \mathcal{X}$ with strict inequality for at least one x. Then:

$$a_1 \equiv \min\{x|\xi(x) > 0\}$$

exists. Since $F_\xi(x) \geq 0$, it follows that for all $x < a_1$: $F_\xi(x) = 0$. Therefore $F_\xi(a_1) = \xi(a_1)$. Similarly, there exists

$$b_1 \equiv \min\{x > a_1|\xi(x) < 0\}$$

Define:

$$p_1 \equiv \min\left\{\xi\left(a_1\right), \left|\xi\left(b_1\right)\right|\right\} > 0$$

Define $\overline{\xi}_1 = \xi - \xi_{a_1 b_1 p_1}$. It is still true that $F_{\overline{\xi}_1}(x) \geq 0$, since $F_{\overline{\xi}_1}(\cdot)$ differs from $F_\xi(\cdot)$ only in the interval $[a_1, b_1]$, and there $F_\xi \geq \xi\left(a_1\right) \geq p_1$. Note that $\overline{\xi}_1$ is a set-function with at least one less mass point than ξ.

Hence if $\overline{\xi}_1 \not\equiv 0$ then $\mathbf{0}$ FOSD $\overline{\xi}_1$ and we can repeat the process, obtaining iteratively $(\xi_2, \xi_3, \ldots, \xi_L)$. Because each $\overline{\xi}_l$ has at least one less mass point than $\overline{\xi}_{l-1}$, and ξ is finitely supported (i.e. there exist only finitely many points x such that $\xi\left(x\right) \neq 0$), the sequence is finite. The sequence has to stop, at some stage L with $\overline{\xi}_L \equiv 0$. Hence $\xi \equiv \sum_{l=1}^{L} \xi_l$, with $p_l > 0$ for all l.

A similar proof holds for the case where ξ FOSD $\mathbf{0}$. ■

Lemma 2 *If $p_l p_k > 0$ then **0** (the zero function) SOSD $\xi_l * \xi_k$ (the convolution of ξ_l and ξ_k), when ξ_l and ξ_k have the (26) structure.*

Proof. The convolution $\xi_l * \xi_k$ is given by:

$$\left(\xi_l * \xi_k\right)(x) = \begin{cases} p_l p_k & if \quad x = a_l + a_k \\ -p_l p_k & if \quad x = a_l + b_k \\ -p_l p_k & if \quad x = b_l + a_k \\ p_l p_k & if \quad x = b_l + b_k \end{cases}$$

$F_{\xi_l * \xi_k}(x) = \int_{-\infty}^{x} \left(\xi_l * \xi_k\right)(t)dt$ is equal to:

$$F_{\xi_l * \xi_k}(x) = \begin{cases} p_k p_l & if \quad x \in \left[a_l + a_k, \min\left\{a_k + b_l, b_k + a_l\right\}\right] \\ -p_k p_l & if \quad x \in \left[\max\left\{a_k + b_l, b_k + a_l\right\}, b_k + b_l\right] \\ 0 & OTHERWISE \end{cases}$$

Therefore:

$$\int_{-\infty}^{x} F_{\xi_l * \xi_k}(t)dt \geq 0$$

That is, the zero function SOSD $\xi_l * \xi_k$. ■

Corollary 2 *In every state in which f and g differ, the zero function SOSD $\theta(s) * \theta(s)$.*

Proof. Since f and g are statewise ranked by FOSD, by Claim 2 the zero function FOSD $\theta(s)$ or vice versa. By Lemma 1, we can decompose every difference of probability measures set-function $\theta(s)$ into $L(s)$ functions with all p_l $(l = 1, \ldots, L(s))$ positive (if $\mathbf{0}$ FOSD $\theta(s)$) or negative (if $\theta(s)$ FOSD $\mathbf{0}$). Therefore:

$$\theta(s) * \theta(s) = \left(\sum_{l=1}^{L(s)} \theta_l(s) \right) * \left(\sum_{k=1}^{L(s)} \theta_k(s) \right) = \sum_{l=1}^{L(s)} \sum_{k=1}^{L(s)} \theta_l(s) * \theta_k(s) \qquad (27)$$

By Lemma 2 each convolution element of the above sum is second order stochastically dominated by the zero function. Therefore, the zero function SOSD the sum of those convolutions. ■

Proof of Theorem 1. Recall from (22) that $U\left(\theta_{(2)}\right)$ is additive across states. By Corollary 2 and Claim 1: $U\left[\theta(s) * \theta(s)\right] < 0$ in every state in which f and g differ. In states in which f and g are equal, $\theta(s) \equiv 0$, and therefore: $U\left[\theta(s) * \theta(s)\right] = 0$. It follows that $U\left(\theta_{(2)}\right) < 0$ and (19) holds. ■