CARESS Working Paper 99-03

A Bayesian Approach to Uncertainty Aversion[1]

Yoram Halevy
University of Pennsylvania
Department of Economics
3718 Locust Walk
Philadelphia PA 19104-6297
halevy@econ.sas.upenn.edu

Vincent Feltkamp[2]
Statistics Netherlands
Department of Statistical Methods
P.O. Box 4481 NL-6401 CZ Heerlen
The Netherlands
Vincent.Feltkamp@bigfoot.com

April 23, 1999

[1]We would like to thank Dean Foster, George Mailath, Nicola Persico, Dan Silverman, Becky Stein and Peter Wakker for helpful remarks.

[2]This paper was written during Vincent Feltkamp s stay at Departmento de Fundamento del Análisis and 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 relate this paradox to other commonly observed anomalies, such as a rejection of the backward induction prediction in the one-shot Ultimatum Game. We argue that the pattern common to these observations is that the behavior is governed by *rational rules* . These rules have evolved and are optimal within the repeated and concurrent environments that people usually encounter. When an individual relies on these rules to analyze one-shot or single circumstances, paradoxes emerge. We show that when a *risk averse* individual has a *Bayesian prior* and uses a *rule* which is optimal for simultaneous and positively correlated ambiguous risks to evaluate a single vague circumstance, his behavior will exhibit *uncertainty aversion.* Thus, the behavior predicted by Ellsberg may be explained within the Bayesian expected utility paradigm.

JEL classi cation: D81
Keywords: Ellsberg paradox, rule rationality, ambiguity aversion, risk aversion, subjective probability.

1 Introduction

In real economic situations, almost all uncertainty is subjective. There is no objective mechanism with which economic decision makers evaluate the future return on a stock, tomorrow s interest rate or the weather in Brazil (which will determine future coffee prices). Therefore, if our goal in economic modelling is to describe and predict within these environments, we have to analyze accurately (or at least understand) how decision makers behave when confronted with subjective uncertainty.

Since the works of Frank Knight [11] and John Maynard Keynes [10] in 1921 there has been a tension between and within disciplines working to quantify subjective uncertainty (i.e., unknown or non-existent objective probability). On one side are researchers seeking a *normative* description of individual belief, i.e., a behavioral rule consistent with a reasonable set of assumptions. The most prominent representative of this group is Leonard J. Savage. In 1954 Savage [14] suggested a set of axioms which, if an individual s preferences over acts abide by them, would result in a representation of her preferences which separated preferences over consequences from the evaluation of events, and for which the evaluation of events would be consistent with an additive probability measure (probabilistic sophistication). Thus, the individual behaves as if he faced a known objective probability, and maximized his expected utility. Mark Machina & David Schmeidler [12], developed a set of axioms which did not force an expected utility structure by relaxing Savage s Sure Thing Principle, while maintaining the probabilistic sophistication part of Savage s result.

On the other side there are researchers who aspire to *describe* individual behavior in the face of uncertainty. Although Frank Knight [11] was the rst to suggest that *risk* (known probabilities) and *uncertainty* (unknown probabilities) are two different notions, he did not make clear the differences between the two in terms of individual behavior. Keynes [10] suggested that ambiguity in probability may be the result of missing information, relative to some conceivable amount of information. Savage, too, was aware that such vagueness of probability may play a role in describing one s preferences (chapter 4 in [14]), but relied on de Finetti [4] to claim that this is the reason why a normative theory of subjective probability is needed. In 1961, Daniel Ellsberg [5] suggested two hypothetical experiments which made clear that, for many individuals, risk and uncertainty are two different concepts. Thus, any attempt to describe their preferences by a (subjective) probability measure is doomed to fail. Although Ellsberg described only a thought experiment (with a very specific sample), his results were later confirmed in many experimental studies (see Camerer & Weber [3] for a comprehensive survey).

Ellsberg offered two alternative representations of the observed preferences. In Ellsberg's first explanation the decision maker behaves *pessimistically*, choosing an action to maximize his utility if the worst scenario occurs (i.e. the "maximin" over a set of priors). A set of axioms which supports such behavior was suggested by Gilboa & Schmeidler [8]. Schmeidler [15] and Gilboa [7] derived the Choquet expected utility representation, which is - if the capacity is convex - a special case of the maximin. Schmeidler [15] defined formally the concept of *Uncertainty Aversion* (to be discussed in section 3), that generalizes the behavior Ellsberg predicted. Intuitively, an individual exhibits uncertainty aversion if he prefers the objective mixing of objective lotteries over a subjective mixing. The second formulation suggested by Ellsberg is a weighted average of the maximin rule and expected probability (according to some prior). This results in behavior which is *conservative*[1] with respect to ambiguity, but not as extreme as the maximin rule. A different explanation was suggested by Uzi Segal [16]. He derived the Ellsberg behavior by considering an "Anticipated Utility" model (in which the utility from an objective lottery is not linear in probabilities) and by relaxing the "reduction of compound lotteries" assumption.

We offer the following observation: Uncertainty aversion is not the only example of behavior where individuals' decisions do not conform to the normative predictions we (i.e. modelers) construct. Two prominent examples are the one-shot "Prisoners' Dilemma" and the "Ultimatum Game". In the first example, almost all normative notions of equilibrium (except when agents have unobserved 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

[1]Conservatism is used here and elsewhere in the paper in the ordinary sense of the word: the individual will favor traditional or current strategies over innovations whose consequences are ambiguous. In the case of risk and uncertainty, he will prefer known risky strategies over uncertain (unknown probabilities) ones.

(and 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 nd it hard to evaluate, singular situations.

The following justi cation is suggested for this bounded rational behavior: during their lives individuals frequently encounter repetitive or simultaneous circumstances, i.e. only rarely does a situation turn out to be singular (one-shot and single). Peoples way of thinking has adapted to this pattern. Furthermore, people develop rules of thumb to cope in those circumstances. These rules are determined by an evolutionary process or a learning process. These processes reward a behavior that utilizes a rule which works well in most circumstances. This type of rationalization was referred to by Aumann [2] as *Rule Rationality* (vs. Act Rationality). The outcome is that people s heuristic decision making is consistent with repeated or simultaneous situations. When the individual confronts a singular circumstance, on a subconscious level she perceives a positive probability that it will be repeated or that other simultaneous circumstances are positively correlated with it, and behaves according to the rule which is optimal for this case. To an outside observer, this behavior seems irrational . However, if the situation were to be repeated or positively correlated with concurrent situations that were to occur, this behavior would seem rationalized. Hoffman, MacCabe and Smith [9] 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 a reputation. But since the rule has been unconsciously chosen, it will not be consciously abandoned[2].

In this work, we show that this line of reasoning applies to uncertainty aversion as well. The environments in which people make decisions under uncertainty are frequently composed from multiple risks that are positively correlated. Examples of such cases are a purchase of a car or of a health insurance, and even marriage. In the case of a car, the state of each system is uncertain, and given its state there is risk the system will malfunction after certain mileage. The states of different systems are positively correlated

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

(e.g., may depend on previous owners). The decision is whether to buy the "car" (including all its ingredients) or not. In the health insurance example, the individual can not insure different risks separately, but has to choose a "package" which covers multiple risks that are positively correlated. The happiness derived from a marriage is composed of many (risky) dimensions that are positively correlated. The individual takes a decision while having some belief over the extent of these risks. The individual's heuristic decision making ("rule") under uncertainty has adapted to this "bundling" of risks, and when confronted with a single uncertain situation, employs the rule of thumb which has evolved. In this work we prove that if a risk averse individual has some Bayesian prior belief over states of the world, but instead of analyzing the single Ellsberg experiment derives an optimal decision *rule* for multiple positively correlated experiments, then she will exhibit uncertainty aversion. In this case, uncertainty aversion reduces to risk aversion, and justifies the usual response that a lottery where probabilities are unknown is "riskier" than a lottery with known probabilities. The explanation is *conservative*, and we can bound the premium the subject is willing to pay in order to discard uncertainty in favor of risk.

In the following we present the Ellsberg paradoxes, and our resolution of them. Next, we generalize the example and establish formally the relation between behavioral rules and uncertainty aversion, viz., almost every uncertainty averse behavior may be rationalized as a Bayesian optimal rule in an environment consisting of simultaneous risks. The paper ends with a discussion and suggestions for further research.

2 A Bayesian Resolution of The Ellsberg Paradoxes

Consider Ellsberg's first paradox: 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 - Alice - 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). If Alice exhibits *uncertainty aversion* she will prefer a bet on red (black) drawn from the first urn to a bet on red (black) drawn from the second urn, but she will be indifferent between betting on red or black from the rst (second) urn (the formal de nition is deferred until section 3). This pattern of behavior not only violates Savage s Sure Thing Principle (P2), but there does not exist any subjective probability (i.e., frequency of reds or blacks in the second urn) which supports these preferences. In the Machina-Schmeidler [12] terminology, Alice is not probabilistically sophisticated . As suggested by Ellsberg [5], and axiomatized by Gilboa & Schmeidler [8], this behavior can be supported by a *pessimistic* evaluation (i.e. maximin): Alice has a set of priors and for each bet she calculates her expected utility according to the worst prior belief supported in this set. In this example, if p is the proportion of red balls in the second urn, then $p \in [0, 1]$. Therefore, Alice s maximin expected utility from betting on red (black) from the second urn is zero. According to this pessimistic explanation, Alice would prefer to bet on red (black) from the rst urn, even if she knew that there is (are) one (99) red ball(s) in it, rather than bet on red (black) from the second urn. The unsatisfying predictions of this extreme pessimism, led Ellsberg [5] to suggest a more *conservative* view, and in this respect (and only in this), do we follow him.

Alice has learned from experience that most circumstances are not isolated, but frequently risks are positively correlated. When asked which bet she prefers, she analyzes which bet would be the optimal, if similar and positively correlated draw were to be performed, and her payoff was the average of the draws. For simplicity of the initial exposition we assume two experiments and perfect correlation. The distribution of the average monetary prize if Alice bets on red (or black) from the urns with a known probability of $\frac{1}{2}$ (urn I) is:

$$IR_{(2)} = IB_{(2)} = \begin{cases} \$0 & 1/4 \\ \$50 & 1/2 \\ \$100 & 1/4 \end{cases} \tag{1}$$

When considering the ambiguous urns, Alice applies the statistical principle of *insufficient reason*[3]. Therefore, she has a prior belief about the number of red balls contained in them, which assigns a probability of $\frac{1}{101}$ to every

[3] 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 [14] Chapter 4 section 5 for a discussion of the principle in relation to subjective probability.

frequency between 0 and 100 (thus p is between 0 and 1). The assumption of perfect correlation is that the two urns have the same color composition (this is an exchangeability condition). Conditional on p, the probability that two red balls would be drawn from the ambiguous urns (i.e. winning \$100 on average 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. an average prize of \$50 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. Thus the probability of winning on average \$100 and \$0 is:

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

Thus, the expected (according to the symmetric prior) distribution of the average monetary payoff from betting on the ambiguous urns is:

$$IIR_{(2)} = IIB_{(2)} = \begin{cases} \$0 & 1/3 \\ \$50 & 1/3 \\ \$100 & 1/3 \end{cases} \tag{3}$$

$IR_{(2)}$ and $IB_{(2)}$ *second order stochastically dominate* $IIR_{(2)}$ and $IIB_{(2)}$ (i.e. the latter two are mean preserving spreads of the former)[4]. If Alice is averse to mean preserving spreads, she will prefer to bet on the rst 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 the objective urns to the ambiguous ones, and will exhibit uncertainty 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); while if she is risk neutral, she will be indifferent between the four bets.

The above explanation is of course *conservative.* In the case of two draws, and without dependence on her risk aversion, Alice will prefer to bet on the

[4]For formal de nitions of rst and second order stochastic dominance see [13] and Appendix A.

ambiguous urns, rather than bet on red from urns that contain anything less than 43 red balls. The distribution of a bet on red from an urn which contains only 42 red balls is:

$$IR\left(p=\frac{42}{100}\right)=(\$0,0.3364;\$50,0.4872;\$100,0.1764)$$

Hence, a bet on the uncertain urns would rst order stochastically dominate* a bet on red from these risky urns. Thus the *uncertainty premium* (in terms of probabilities) is bounded from above by 8%. In monetary terms, this upper bound is equivalent to $8:

$$E\left(IB_{(2)}\left(p=\frac{1}{2}\right)\right)-E\left(IB_{(2)}\left(p=\frac{42}{100}\right)\right)=\$50-\$42=\$8.$$

The only assumption relied upon in this explanation is monotonicity of the preference relation with respect to second order stochastic dominance. Therefore, this explanation is consistent with any theory of choice under risk which exhibits aversion to mean preserving spreads, including expected utility with diminishing marginal utility of wealth, Yaari s dual theory [19] (which separates risk aversion and utility of wealth) when the transformation function is convex, as well as many others.

The only relation between the two 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 rst and second ambiguous urns, respectively. It is immediate to verify that if $Corr\left(p_1,p_2\right)>0$ then $E\left(p_1p_2\right)=E\left(\left(1-p_1\right)\left(1-p_2\right)\right)>\frac{1}{4}$, 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 need not assign probability one to the simultaneous experiment in order to prefer a bet from the rst urn. She might have learned from her experience that some risks are simultaneous, but some are isolated. Even if the probability of a correlated risk is very small, she would prefer a bet on the risky urns. This is a consequence of a Sure Thing Principle argument: if there is only a single risk, she is indifferent between betting on urn I or urn II, and in the case of correlation, she strictly prefers the former. Hence the conclusion that she prefers a bet on urn I, even when she faces the slightest possibility of simultaneity.

2.1 Generalization to any number of concurrent ambiguous risks

The logic developed above applies to any number of simultaneous risks. Assume Alice compares the distribution of betting on r simultaneous IR (IB) to r simultaneous IIR (IIB) as in the Ellsberg experiment. The average money gained is distributed $\frac{100X}{r}$ where X has a binomial distribution with parameters $(0.5, r)$ and (p, r), respectively. p - The proportion of red balls in the second urn, is distributed uniformly on $[0, 1]$. Therefore for $0 \leq k \leq r$: [5]

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

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 $IB_{(r)}$.

2.2 The Second Ellsberg Paradox

Ellsberg s [5] second paradox (the one urn example) is the following: An urn contains 90 balls: 30 red and 60 black or yellow (with unknown proportions). A ball is drawn at random and Bob is asked to bet on the color of the ball. A correct guess wins \$100, an incorrect guess wins \$0. Bob prefers a bet that the ball is red over a bet that the ball is black, and prefers a bet that the ball is either black or yellow over a bet that the ball is either red or yellow. Bob s preferences seem to be inconsistent with any frequency of black (yellow) balls. We claim, however, that Bob is Bayesian and assigns a uniform probability to the frequency of black balls. He is bounded rational in the sense that he is using a rule which is optimal for concurrent experiments. The average payoff to Bob if he bets on red balls from

[5] The *Beta Integral* is de ned 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: $\Gamma(k) = (k-1)!$

two urns is: $R_{(2)} = \left(\$0, \frac{4}{9}; \$50, \frac{4}{9}; \$100, \frac{1}{9}\right)$. The probability distribution of a bet on black is: $B_{(2)}(p) = \left(\$0, (1-p)^2; \$50, 2p(1-p); \$100, p^2\right)$ where p is the relative frequency of black balls in the urns. Bob's prior belief over p is derived from symmetry and is (approximately) uniform (neglecting the finite support), i.e.: $p \sim U\left[0, \frac{2}{3}\right]$. Averaging the distribution of $B_{(2)}$ over p results in: $B_{(2)} = \left(\$0, \frac{13}{27}; \$50, \frac{10}{27}; \$100, \frac{4}{27}\right)$. It is easily verified that $E\left(R_{(2)}\right) = E\left(B_{(2)}\right) = 33\frac{1}{3}$, and that $B_{(2)}$ is a mean preserving spread of $R_{(2)}$. A symmetric analysis applies to the second pattern of preferences Bob exhibits.

3 Rules and Uncertainty Aversion

The natural framework to generalize Ellsberg's examples is the Anscombe-Aumann [1] horse bets over roulette lotteries, in which objective and subjective probabilities coexist. In this section we show that almost all cases of observed uncertainty aversion (Schmeidler [15] and Gilboa & Schmeidler [8]) may be explained if the agents' preferences are defined over rules. The Anscombe-Aumann framework has some complications due to the two-stage setup. However, as long as we remain within the expected utility framework, we bypass those difficulties. A one-stage axiomatization of expected utility was suggested by Rakesh Sarin and Peter Wakker [18], but the definition of uncertainty aversion in their framework is not transparent and will be different [17] of Schmeidler's. A definition of uncertainty aversion within a Savage domain of acts was suggested recently by Larry Epstein [6]. Formulation of the results presented in the previous section within this framework, may shed light on the degree of generalization included in Epstein's definition, and remains for future work.

Let $\mathcal{X}$ be a set of *outcomes.* $\mathcal{R}$ is the set of finitely supported *(roulette) lotteries* over $\mathcal{X}$, i.e. r in $\mathcal{R}$ defines an objective mechanism of mixing among the elements of $\mathcal{X}$. Assume a preference ordering over $\mathcal{R}$ which satisfies the usual expected utility assumptions. Therefore, there exists a Bernoulli utility function $u(\cdot)$, such that lottery r_1 is preferred to lottery r_2 if and only if $\sum_{x\in\mathcal{X}} r_1(x)u(x) > \sum_{x\in\mathcal{X}} r_2(x)u(x)$. Let S be the (non-empty) set of *states of nature.* 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. Consider the set of roulette lotteries over $\mathcal{H}$, denoted by $\mathcal{R}^*$. Note that every act is a degenerate element of $\mathcal{R}^*$. An example of an element of $\mathcal{R}^*$ is the lottery: $(f, \alpha; g, 1-\alpha)$ for f and g in $\mathcal{H}$ and $0 \leq \alpha \leq 1$. The holder of this lottery will receive in every state $s \in S$ the compound lottery $(\alpha f(s) + (1-\alpha) g(s))$. Anscombe and Aumann assumed that preferences over $\mathcal{R}^*$ satisfy the independence axiom. As a result, if f and g are two acts between which the individual is indifferent, then he is indifferent between the two and the lottery $(f, \alpha; g, 1-\alpha)$. This assumption, plus an assumption on the reversal of order in compound lotteries, yield a representation of preferences over acts as an expected utility with respect to a derived subjective probability.

The independence assumption over elements of $\mathcal{R}^*$ is challenged by the Ellsberg paradox. In this setting the set of outcomes is $\mathcal{X} = \{\$0, \$100\}$. States of the world are denoted by the number of red balls in the second urn: $S = \{0, ..., 100\}$. The act IIR de nes for every state the objective lottery:

$$IIR(s) = \left(\$100, \frac{s}{100}; \$0, 1 - \frac{s}{100}\right)$$

The act IIB de nes for every state the lottery:

$$IIB(s) = \left(\$100, 1 - \frac{s}{100}; \$0, \frac{s}{100}\right)$$

The typical individual is indifferent between IIR and IIB, but prefers the lotteries IR and IB to either of the former. A simple calculation reveals that $IR = \left(IIR, \frac{1}{2}; IIB, \frac{1}{2}\right)$, and thus the independence axiom over $\mathcal{R}^*$ is violated. This observation led Schmeidler [15] to relax independence in favor of *comonotonic independence*. Two acts f and g are comonotonic if for every two states $s, s' \in S$: $f(s) \succsim f(s')$ if and only if $g(s) \succsim g(s')$. Schmeidler [15] constrained independence to hold true only for comonotonic acts.

Following David Schmeidler [15], it seems natural to generalize and de- ne a decision maker to be *strictly uncertainty averse* if for *every* two non-comonotonic acts f and g, between which she is indifferent, the decision maker prefers any convex combination of them to each one separately. Indeed, Gilboa and Schmeidler [8] assumed weak uncertainty aversion as one of their axioms in deriving the maximin representation. This behavior is explained intuitively as the agent hedging between the two acts. Note that this is a generalization of Ellsberg s observation; in his two urns example the *utility* of an act at state s is proportional to the probability of receiving \$100 in this state, according to that act (normalizing $u(0)$ to 0). It is easy to verify that our explanation of the Ellsberg example is consistent with this de nition of strict uncertainty aversion: the bet on red or black from two ambiguous urns is a mean preserving spread of the compound bet on $(IIR, \alpha; IIB, 1-\alpha)_{(2)}$ from these urns, which for every $0 < \alpha < 1$ has the distribution:

$$(IIR, \alpha; IIB, 1-\alpha)_{(2)} = \begin{cases} \$0 & \left(\alpha^2 - \alpha + 1\right)/3 \\ \$50 & \left(-2\alpha^2 + 2\alpha + 1\right)/3 \\ \$100 & \left(\alpha^2 - \alpha + 1\right)/3 \end{cases}$$

However, this generalization *ignores* the unique symmetry in the Ellsberg example because, in Ellsberg, the acts are non-comonotonic in *every* two states. Furthermore, in the Ellsberg example, the lotteries assigned by IIR and IIB are ordered according to First Order Stochasic Dominance criterion in every state in which they differ. That is, *every* agent with monotone preferences would prefer $IIR(s)$ over $IIB(s)$ if $51 \leq s \leq 100$ and $IIB(s)$ over $IIR(s)$ if $0 \leq s \leq 49$. Therefore, the hedging behavior could be interpreted as more fundamental, and independent of the agent s utility function. This observation is generalized and proved in our framework of rules : we adopt all Anscombe-Aumann assumptions, but by considering multiple simultaneous lotteries, a seemingly uncertainty averse behavior emerges.

Formally, let $\mathcal{X}$ be a nite set of monetary outcomes. For every $s \in S$ let $q(s)$ be the subjective probability of state s. A *rule* is de ned as an act with r concurrent lotteries at every state, $r > 1$. We assume the lotteries at every state are exchangeable. This is a generalization of the same color composition in the Ellsberg example. Thus, the environment in which the rule is evaluated is represented by r. An agent s preferences are de ned over rules (although an outside observer might view only acts). She is indifferent between the rules f and g if:

$$U\left(f_{(r)}\right) = U\left(g_{(r)}\right) \tag{4}$$

Or, explicitly:

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

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

$$E\left[u\left(f_{(2)}\left(s\right)\right)\right] = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{X}} f\left(s\right)\left(x\right) f\left(s\right)\left(y\right) u(x+y) \tag{5}$$

Let p_1 and p_2 be probability measures over $\mathcal{X}$ with distribution functions F_1 and F_2, respectively. The probability measure p_1 would *First Order Stochastic Dominate (FOSD)* p_2 if $F_1(x) \leq F_2(x)$ for every $x \in \mathcal{X}$ and there exists at least one outcome in which the inequality is strict.

De nition 1 *Acts f and g are* Non-Comonotonic in the Strong Sense *if in every state s in which they differ $f(s)$ FOSD $g(s)$ or vice versa, and there exist at least two states in which the order of FOSD is reversed.*

Note that this de nition is stronger than Schmeidler s non-comonotonicity since f and g will be non-comonotonic for every agent with a strictly monotone utility function.

Following is a generalization of our main result which derives uncertainty averse behavior as rule rationality :

Theorem 1 *If f and g are Non-Comonotonic in the Strong Sense and the agent is indifferent between the rule f and the rule g, then if she is averse to mean preserving spreads and her preferences are representable by an expected utility functional, she will prefer the rule of $(f, \alpha; g, 1-\alpha)$ over the rule f for every $0 < \alpha < 1$.*

P roof. See Appendix. ■

The implication of Theorem 1 is that almost all seemingly uncertainty averse behavior may be rationalized if the agent s perception is that she is facing concurrent correlated risks. Confronted with this environment, if she is risk averse her observed behavior would exhibit uncertainty aversion.

Uncertainty averse behavior may be rationalized even if the individual thinks there is only a small probability of simultaneity. The source of this belief is the agent s experience that some circumstances consist of multiple risks and some from a singular risk. Confronted with a new situation, if the individual s heuristic belief assigns some (possibly small) probability to the possibility he faces a complex circumstance, then her optimal behavior would exhibit uncertainty aversion.

Corollary 2 *Assume f and g as in Theorem 1, and suppose the individual is indifferent between the act f and the act g. Then, for every $\beta > 0$ probability of concurrent risks, she will prefer $(f, \alpha; g, 1-\alpha)$ over f, for every $0 < \alpha < 1$.*

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

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

■

4 Discussion and Conclusion

A pattern of behavior which exhibits uncertainty aversion is related to other puzzling cases of *irrationality* , such as evidence from experiments with the one-shot Prisoners Dilemma or the Ultimatum Game. In each case the explanation for the observed anomaly in a single experiment is that people (as opposed to abstract decision makers) nd it hard to analyze circumstances that are not recurrent or concurrent. The cause for this behavior may be found in individual experience which teaches that most circumstances are not singular. Human decision making have adapted to this reality by adopting rule rationality procedures. The result is that when confronted with a new circumstance the individual typically assigns a positive probability it will be repeated or that his decision may affect other simultaneous risks.

Our assumption is that, in this environment, every time the agent encounters a new circumstance he behaves according to the *rule of thumb* that has evolved in environments which consist of simultaneous risks. This results in a *conservative* behavioral rule, as opposed to the current pessimistic explanation of the maximin. Our explanation of uncertainty aversion builds directly on people being *risk averse*, and analyzing the *concurrent* ambiguous risks with a Bayesian *prior*. Therefore, this explanation does not depend on the Expected Utility structure, and is applicable to other theories of choice under risk in which agents preferences exhibit aversion to mean preserving spreads. This work sheds new light on the structure of the Ellsberg paradoxes , that triggered the axiomatic literature of the maximin and nonadditive expected utility. In particular, Schmeidler s de nition of uncertainty aversion, seem to be stronger than the observed behavior. This work calls for new experimental work which will bridge the gap or differentiate between the Ellsberg experiment and the formal de nition of uncertainty aversion. Future work could analyze similarly the generalization embodied in Epstein s de nition.

This work opens new research opportunities in the area of uncertainty aversion, by using standard tools, as expectation with respect to an additive probability. Formal de nition and study of the uncertainty premium will be useful in information economics, nance, and other economic elds. One interesting experiment would be to compare preferences between an ambiguous Ellsberg bet and the reduced risky bet on r simultaneous urns (as in (3)), for uncertainty averse individuals. Indifference between the two would support the hypothesis presented in this work. Finally, a formal evolutionary model in which rule rationality , as described above, emerges may illuminate the set of procedures for which this notion of bounded rationality is viable.

References

[1] Anscombe, F. J. and R. J. Aumann (1963): A De nition of Subjective Probability , *Annals of Mathematical Statistics*, 34, 199-205.

[2] Aumann, Robert J. (1997): Rationality and Bounded Rationality , *Games and Economic Behavior*, 21, 2-14.

[3] Camerer, Colin and Martin Weber (1992): Recent Developments in Modeling Preferences: Uncertainty and Ambiguity , *Journal of Risk and Uncertainty*, 5, 325-370.

[4] de Finetti, Bruno (1937): La prevision: ses lois logiques, ses sources subjectives , *Annales de l Institute Henri Poincaré,* Vol. 7, 1-68. Translation to English in: Kyburg, Henry E. Jr. and Howard E. Smokler, eds. (1964): *Studies in Subjective Probablility*, New York: John Wiley and Sons.

[5] Ellsberg, Daniel (1961): Risk, Ambiguity and the Savage Axioms , *The Quarterly Journal of Economics,* 75 (4), 643 669.

[6] Epstein, Larry G. (1998): A De nition of Uncertainty Aversion , WP # 453 Rochester Center for Economic Research (forthcoming in: *Review of Economic Studies*).

[7] Gilboa, Itzhak (1987): Expected Utility with Purely Subjective Non-Additive Probabilities , *Journal of Mathematical Economics,* 16, 65 88.

[8] Gilboa, Itzhak and David Schmeidler (1989): Maxmin Expected Utility with a Non-Unique Prior , *Journal of Mathematical Economics,* 18, 141 153.

[9] Hoffman, Elizabeth, Kevin McCabe, and Vernon L. Smith (1996): Social Distance and Other-Regarding Behavior in Dictator Games , *The American Economic Review*, 86 (3), 653-660.

[10] Keynes, John M. (1921): *A Treatise on Probability*, London: McMillan.

[11] Knight, Frank H. (1921): *Risk, Uncertainty and Pro t*, Boston: Houghton Mi- in Company.

[12] Machina, Mark J. and David Schmeidler (1992): A More Robust De - nition of Subjective Probability , *Econometrica*, 60 (4), 745-780.

[13] Rothschild, Michael and Joseph E. Stiglitz (1970): Increasing Risk: I A De nition , *Journal of Economic Theory*, Vol. 2 (3), 225-243.

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

[15] Schmeidler, David (1989): Subjective Probability and Expected Utility without Additivity , *Econometrica,* 57 (3), 571 587.

[16] Segal, Uzi (1987): The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach , *International Economic Review*, 28 (1), 175-202.

[17] Sarin, Rakesh and Peter Wakker (1992): A Simple Axiomatization of Nonadditive Expected Utility , *Econometrica*, 60 (6), 1255-1272.

[18] Sarin, Rakesh and Peter Wakker (1997): A Single-Stage Approach to Anscombe and Aumann s Expected Utility , *Review of Economic Studies*, 64, 399-409.

[19] Yaari, Menahem E. (1987): The Dual Theory of Choice Under Risk , *Econometrica*, 55, 95-115.

A Preliminaries

Let ψ and τ be finite measures on $\mathcal{X}$. Define[6]:

$$F_{\psi}(x) = \int_{-\infty}^{x} \psi(t)dt \text{ and } F_{\tau}(x) = \int_{-\infty}^{x} \tau(t)dt \tag{6}$$

Assume ψ and τ are such that:

$$F_{\psi}(+\infty) = F_{\tau}(+\infty) \tag{7}$$

Assumption (7) would hold true if, for example, ψ and τ are probability measures (then (7) is equal to one), or when each is a difference of two probability measures (then (7) is equal to zero).

Definition 2 *Let ψ and τ be two finite measures defined over $\mathcal{X}$, and let F_{ψ} and F_{τ} be defined as in (6) and satisfy (7). The measure ψ* First Order Stochastic Dominates (FOSD) *the measure τ if for every $x \in \mathcal{X}$: $F_{\psi}(x) \leq F_{\tau}(x)$ with strict inequality for at least one x.*

Definition 2 is a generalization of the standard definition of first order stochastic dominance, and of course 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:

$$\int_{\mathcal{X}} F_{\psi}(x)dx = \int_{\mathcal{X}} F_{\tau}(x)dx \tag{8}$$

[6] Since all the measures we shall deal with have finite variation, all the integrals converge.

That is, the mean measure of ψ is equal to the mean measure 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.

De nition 3 ψ Second Order Stocastically Dominates (SOSD) τ *if (8) 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.

If ψ SOSD τ then:

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

for all strictly monotone and strictly concave u.

P roof. he proof is similar to Rothschild and Stiglitz s [13]: using (7) instead of assuming probability measures, and (8) instead of assuming equal expectations. ■

B Proof of Theorem 1

Let f and g be non-comonotonic in the strong sense, and:

$$U\left(f_{(2)}\right) = U\left(g_{(2)}\right) \tag{4 }$$

De ne for every $s \in S$:

$$h(s)(x) = \alpha f(s)(x) + (1-\alpha) g(s)(x) \tag{9}$$

Then we need to show that:

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

Consider the function θ de ned as:

$$\theta(s)(x) = f(s)(x) - g(s)(x) \tag{10}$$

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

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

$$= \sum_s q(s) \sum_x \sum_y \left[\begin{array}{c} \alpha^2 (f(s)(x)) (f(s)(y)) + \\ + (1-\alpha)^2 (g(s)(x))(g(s)(y)) + \\ +2\alpha(1-\alpha)(f(s)(x))(g(s)(y)) \end{array} \right] u(x+y) \tag{11}$$

Let $\theta_{(2)}$ be the convolution of θ with θ in 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):

$$\begin{aligned} U\left(\theta_{(2)}\right) &= \sum_s q(s) U\left[\theta(s) * \theta(s)\right] = &(12) \\ &= \sum_s q(s) \sum_x \sum_y \theta(s)(x)\,\theta(s)(y)\, u(x+y) = \\ &= \sum_s q(s) \sum_x \sum_y \left[f(s)(x) - g(s)(x)\right]\left[f(s)(y) - g(s)(y)\right] u(x+y) = \\ &= \sum_s q(s) \sum_x \sum_y \left[\begin{array}{c} (f(s)(x))(f(s)(y)) + \\ +(g(s)(x))(g(s)(y)) - \\ -2(f(s)(x))(g(s)(y)) \end{array} \right] u(x+y) &(13) \end{aligned}$$

By substitution of (11) and (13) and utilizing (4) it follows that:

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

Thus, the theorem will be true if and only if $U\left(\theta_{(2)}\right) < 0$.

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

P roof. ince f and g are non-comonotonic in the strong sense, then if they differ they are ordered according to FOSD. Assume $f(s)$ FOSD $g(s)$. Therefore:

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

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

Lemma 3 *Any function ξ which is the difference of two probability mass measures can be written as a nite sum of measures:*

$$\xi = \sum_{l=1}^{L} \xi_l \tag{14}$$

where:

$$\xi_l(x) = \xi_{a_l,b_l,p_l}(x) = \begin{cases} p_l & \text{if} \quad x = a_l \\ -p_l & \text{if} \quad x = b_l \\ 0 & \quad OTHERWISE \end{cases} \tag{15}$$

with $a_l < b_l$ and $|p_l| \leq 1$. p_l is negative (positive) if and only if ξ_l FOSD $\mathbf{0}$ ($\mathbf{0}$ FOSD ξ_l). Furthermore, if $\mathbf{0}$ FOSD ξ (ξ FOSD $\mathbf{0}$) then all p_l can be chosen positive (negative) in the decomposition (15).

P roof. ecall that since ξ is a difference of probability mass measures, it is a nite measure 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

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

De ne:

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

De ne $\overline{\xi}_1 = \xi - \xi_{a_1 b_1 p_1}$. 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 $F_\xi \geq \xi(a_1) \geq p_1$. Note that $\overline{\xi}_1$ is a measure with at least one less mass point than ξ.

Hence if $\overline{\xi}_1 \not\equiv 0$ then **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 nitely supported (i.e. there exist only nitely many points x such that $\xi(x) \neq 0$), the sequence is nite. 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 **0**. ■

Lemma 4 *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 (15) structure.*

P roof. he measure $\xi_l * \xi_k$ is given by:

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

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

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

Therefore:

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

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

Corollary 5 *In every state in which f and g differ, the zero function SOSD $\theta(s) * \theta(s)$.*

P roof. ince f and g are non-comonotonic in the strong sense, by Claim B the zero function FOSD $\theta(s)$ or vice versa. By Lemma 3, we can decompose every difference measure $\theta(s)$ into $L(s)$ measures with all p_l $(l = 1, \dots, L(s))$ positive (if $\mathbf{0}$ FOSD $\theta(s)$) or negative (if $\theta(s)$ FOSD $\mathbf{0}$). Therefore:

$$\theta(s) * \theta(s) = \left(\sum_{l=1}^{L(s)} \theta_l(s) \right) * \left(\sum_{k=1}^{L(s)} \theta_k(s) \right) = \sum_{l=1}^{L(s)} \sum_{k=1}^{L(s)} \theta_l(s) * \theta_k(s) \tag{16}$$

By Lemma 4 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. ■

P roof. Proof of Theorem 1] Recall from (12) that $U\left(\theta_{(2)}\right)$ is additive across states. By Corollary 5 and Claim A: $U\left[\theta(s) * \theta(s)\right] < 0$ 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 $U\left(\theta_{(2)}\right) < 0$. ■