A Bayesian Approach to Uncertainty Aversion[1]

Yoram Halevy[2]
University of British Columbia
Department of Economics
1873 East Mall
Vancouver, BC CANADA V6T 1Z1
yhalevy@interchange.ubc.ca

Vincent Feltkamp[3]
Statistics Netherlands
Department of Statistical Methods
P.O. Box 4481 NL-6401 CZ Heerlen
The Netherlands
Vincent.Feltkamp@bigfoot.com

First Version: May 1998
Current Version: January 2000

[1]We would like to thank Robert Aumann, Lawrence Blume, David Easley, Jürgen Eichenberger, Dean Foster, Drew Fudenberg, Itzhak Gilboa, Kenneth Hendricks, George Mailath, Mark Machina, Steven Matthews, Nicola Persico, Andrew Postlewaite, Ariel Rubinstein, Dan Silverman, David Schmeidler, Hans Schneeweiss, Uzi Segal, Becky Stein, Guofu Tan, and Peter Wakker as well as seminar participants at the University of Pennsylvania, Cornell, Duke, University of British Columbia, Hebrew University, Tel Aviv University, Haifa University, Ben Gurion University and the Workshop on "New Themes in Decision Theory Under Uncertainty" held in Paris on June 1999 for helpful remarks.

[2]Yoram Halevy would like to thank the Department of Economics at the University of Pennsylvania for its hospitality while working on this paper.

[3]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 argue that *Uncertainty Aversion* may be viewed as a case of *"Rule Rationality"*. This paradigm claims that people's decision making has evolved to simple rules that perform well in most *regular* environments. Such an environment consists of replicas of some basic *singular* circumstance. When the rule is applied to a *singular* environment, the behavior may seem paradoxical. We claim that the regular environment in which decisions under uncertainty take place, is described by one decision that spans multiple ambiguous risks, which are positively correlated. We show that when a *risk averse* individual has a *Bayesian prior* and uses a *rule,* which is optimal for the regular ambiguous environment, to evaluate a singular vague circumstance - his behavior will exhibit *uncertainty aversion.* Thus, the behavior predicted by Ellsberg may be explained within the Bayesian expected utility paradigm.

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

1 Introduction

Daniel Ellsberg's (1963, [4]) experiments demonstrate that for many individuals *risk* (known probabilities) and *uncertainty* (or *ambiguity* - unknown probabilities) are two different notions. The economic importance of Ellsberg's examples is far beyond the distinction between risk and uncertainty, which was first suggested by Knight [10] and Keynes [9] in 1921. It is a direct criticism of Savage's [13] 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 [11] that generalizes the idea of deriving subjective probability from preferences. This assumption is critical in Economics, where the usage of subjective probability is pervasive. In many cases, not only the results depend on the existence of subjective probability, but without it defining the relevant problem would become much more difficult (if not impossible).

We argue that Ellsberg's ambiguity aversion may be viewed as a case of *"Rule Rationality"* (Aumann [2]). This paradigm claims that people's decision making has evolved to simple *rules* that perform well in most *regular* (common) environments. Such an environment consists of replicas of some basic *singular* circumstance. The rule has been determined in an evolutionary or learning process. These processes reward a behavior that utilizes a rule which works well in most environments, i.e. it is optimal for a regular environment. When applying the decision rule to a singular environment[1], the behavior may seem to be hard to rationalize. The environments in which people make decisions under uncertainty are frequently composed from bundled risks that are positively correlated. Examples of such cases are a purchase of a car, a house, and even marriage. In the case of a car, the state of each component is uncertain, and given its state there is a risk the component will malfunction before a certain mileage. The states of different components are positively correlated (e.g., may depend on previous owners). The decision is whether to buy the "car" (including all its ingredients) or not. Similar logic applies to the purchase of a house. The happiness derived from a marriage

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

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. We argue that the *regular environment* in which decisions under uncertainty take place, is described by one decision that spans multiple risks, which are positively correlated. In the car example, the buyer can not buy the transmission of one car, the engine of a second car and the body of a third car. She has to decide whether to buy a specific car, including all its components. The individual's heuristic decision making under uncertainty has adapted to this "bundling" of risks by developing a "rule" that performs well in the regular environment. The outcome is that people's heuristic decision making is consistent with the regular environment where positively correlated risks are bundled in one decision. When confronted with a situation in which she has to make a decision under uncertainty, the decision maker applies that rule. Implicitly, she assigns some small probability that the environment she is confronted with is regular. We prove that if the regular environment consists of replicas of an Ellsberg type decision that are bundled together, the optimal decision *rule* for a risk averse individual, who has some Bayesian prior belief over states of the world, will exhibit uncertainty (ambiguity) aversion. If the decision maker does not know with certainty the structure of the environment, any small probability of multi-dimensional environment will lead to a decision that looks to an outside observer as "irrational". Hence, an arbitrarily small perturbation of the alleged Ellsberg paradox may be fully rationalized within a Bayesian framework. 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*[2] with respect to ambiguity, and we can bound the premium the individual 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 bundled risks. The paper concludes with

[2]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 with ambiguous consequences. In the case of risk and uncertainty, he will prefer known risky strategies over uncertain (unknown probabilities) ones. Pessimism represents the extreme form conservatism.

a discussion of the results, a comparison to the current literature on ambiguity aversion, a conjecture concerning the relation between uncertainty aversion and other behavioral anomalies, an alternative interpretation of the analytical results and suggestions for further research.

2 A Bayesian Resolution of Ellsberg's Paradoxes

This section demonstrates how the concept of "rule rationality" could be applied to the famous Ellsberg paradoxes, which motivates most of the literature on ambiguity aversion. Note that we use some simplifying assumptions that are not necessary (the more general case is analyzed in Section 3).

2.1 Ellsberg's "Two Urn" Paradox

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 or black drawn from the first urn to a bet on red or black drawn from the second urn, but she will be indifferent between betting on red or black in each urn separately (the formal definition of uncertainty aversion 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 [11] terminology, Alice is not "probabilistically sophisticated". As suggested by Ellsberg [4], and axiomatized by Gilboa & Schmeidler [7], 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 first 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[3], led Ellsberg (in his original paper [4]) to look for a less *conservative* model. As will be clear from the following, our explanation has this feature.

Alice has learned from experience (maybe not consciously) that most circumstances are not isolated (singular), but frequently similar risks are bundled. Hence, the *regular* environment in which she evaluates uncertain prospects consist of bundled risks. When asked which bet she prefers, she applies the *rule* that has evolved in this regular-bundled environment. Our goal here is to characterize the regular environment and find the optimal rule for it. The Ellsberg experiment described above constitute the singular environment in the rule rationality paradigm. For simplicity of the initial exposition we assume the regular environment consist of two Ellsberg singular experiments, which are perfectly correlated. There are two type I urns (risky), and two type II urns (ambiguous). Alice's choice set consists of betting on one color from the (two) risky urns, or on one color from the (two) uncertain urns. 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[4] apply the statistical principle of *insufficient reason*[5]. Therefore, she has a prior belief over the number of red balls contained in them, which assigns a probability of $\frac{1}{101}$ to

[3]The Maximin decision rule for the Ellsberg experiment could be adjusted to conform better with observed behavior, by changing the support of belief. However, it remains true that this decision rule, as a normative result, ignores the distribution of belief over the support. Hence pessimism is this context means: ignoring all the support except its lower bound.

[4]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.

[5]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 [13] Chapter 4 section 5 for a discussion of the principle in relation to subjective probability.

every frequency between 0 and 100 (thus p, the proportion of red balls in the ambiguous urns, 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 \$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 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:

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

$IR_{(2)}$ and $IB_{(2)}$ *second order stochastically dominate* $IIR_{(2)}$ and $IIB_{(2)}$ (i.e. the latter two are mean preserving spreads of the former)[6]. 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 the objective urns to the ambiguous ones, 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); while if she is risk neutral, she will be indifferent between the four bets.

The above explanation is *conservative.* In the case of two draws and a uniform prior, but without dependence on her risk aversion, Alice will prefer

[6]For formal definitions of first and second order stochastic dominance see [12] and Appendix A.

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

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

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 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 from 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]

---

[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: $\Gamma(k)=(k-1)!$

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

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 $IB_{(r)}$.

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 immediate 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 don't know (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, ambiguity aversion is fully rational. In other words: the slightest perturbation of the singular environment gives rise to the rationalizable behavior.

2.2 Ellsberg's "One Urn" Paradox

As will be clear from the generalization in section 3, Ellsberg's [4] second paradox (the "one urn" example) could be rationalized similarly. For the sake of completeness we cover this example too. 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's rationality abides by "rule", when his regular environment bundles risks. In this environment, uncertainty averse behavior would be the result of Bayesian prior. Assume (again, for simplicity only) that the regular environment consists of two bundled bets. Bob's payoff if he bets on red balls from two urns is: $R_{(2)} = \left(\$0, \frac{4}{9}; \$100, \frac{4}{9}; \$200, \frac{1}{9}\right)$. The probability distribution of a bet on black is: $B_{(2)}(p) = \left(\$0, (1-p)^2; \$100, 2p(1-p); \$200, p^2\right)$ where p is the relative frequency of black balls in the urns. Assume Bob's prior belief over p 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: $B_{(2)} = \left(\$0, \frac{13}{27}; \$100, \frac{10}{27}; \$200, \frac{4}{27}\right)$. 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 The General Framework

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 [14] and Gilboa & Schmeidler [7]) may be explained if agents apply a rule, which is optimal for a regular environment, to a singular environment. In the following we formulate the general framework, while relating each element directly to Ellsberg's "two urns" example.

3.1 Uncertainty Aversion

Let $\mathcal{X}$ be a set of monetary *outcomes.* In Ellsberg's setting $\mathcal{X} = \{\$0, \$100\}$. $\mathcal{R}$ is the set of finitely supported *(roulette) lotteries* over $\mathcal{X}$, i.e. ρ in $\mathcal{R}$ defines an objective mechanism of mixing among the elements of $\mathcal{X}$. In Ellsberg's setting $\mathcal{R} = \{(\$100, p; \$0, 1-p) : 0 \leq p \leq 1\}$ i.e., all lotteries between \$0 and

\$100. Assume a preference ordering over $\mathcal{R}$ that 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 $\sum_{x\in\mathcal{X}} \rho_1(x)u(x) > \sum_{x\in\mathcal{X}} \rho_2(x)u(x)$. In Ellsberg's example, if u is monotonically increasing then the decision maker simply prefers a bet with higher probability of winning \$100. Let S be a finite[8] (non-empty) set of *states of the world.* In Ellsberg's example states of the world are denoted by 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. In Ellsberg's two urns example the act IIR defines for every state s the objective lottery:

$$IIR(s) = \left(\$100, \frac{s}{100}; \$0, 1 - \frac{s}{100}\right) \tag{7}$$

The act IIB defines for every state s the lottery:

$$IIB(s) = \left(\$100, 1 - \frac{s}{100}; \$0, \frac{s}{100}\right) \tag{8}$$

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 $(f(s), \alpha; g(s), 1-\alpha)$. 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 knows to calculate probabilities[9], it is easy to verify that this compound lottery is equal to betting on IR (winning \$100 with probability of 50%). 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 (as Alice is indifferent between IIR and IIB), then she is indifferent between the two and the lottery $(f, \alpha; g, 1-\alpha)$. Alice's preferences in the Ellsberg's example violate this assumption since IR is preferred to IIR. The independence over $\mathcal{R}^*$, plus an

[8]The assumption of finiteness is not necessary, but makes the economic interpretation of the results clearer.

[9]This is the Reduction of Compound Lotteries assumption. It is a necessary part of expected utility theory. However, in theories of non-expected utility this assumption may be relaxed (see Segal [15]).

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 [1].

The literature on uncertainty aversion in the Anscombe-Aumann framework follows Schmeidler [14], in focusing on how individual's preferences among lotteries, which are assigned by an act to different states, change. Formally:

Definition 1 (Schmeidler) *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 $g(s) \succsim g(s')$.*

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. Following this logic, it seems natural to generalize and define uncertainty aversion in terms of comonotonic acts:

Definition 2 (Schmeidler [14], Gilboa-Schmeidler [7]) *A decision maker is 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.*

Previous generalizations of expected utility built on these definitions. In his axiomatization of expected utility with respect to a non-additive probability, Schmeidler [14] constrained independence to hold only for comonotonic acts. Gilboa and Schmeidler [7] assumed weak uncertainty aversion as one of their axioms in deriving the Maximin representation.

3.2 The *Regular* Environment

Uncertainty averse behavior is explained intuitively as the agent "hedging" between the two acts. However, this generalization *ignores* the unique symmetry in the Ellsberg examples. 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 $IIR(s)$ over $IIB(s)$ if $51 \leq s \leq 100$ and $IIB(s)$ over $IIR(s)$ 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 "rule rationality", in which we adopt all Ancombe-Aumann assumptions.

Let $\mathcal{X}$ be a finite set of monetary outcomes. $\mathcal{R}$ is the set of finitely supported *(roulette) lotteries* over $\mathcal{X}$. Let S be a finite (non-empty) set of *states of the world*. For every state $s \in S$ let $q(s)$ be the subjective probability of state s. An *Act* is a function from states to lotteries over outcomes.

Definition 3 *Acts f and g 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 rules in the *regular* environment, with more than a single lottery at every state, a seemingly uncertainty averse behavior emerges.

Definition 4 *A* Rule *$f_{(r)}$ is a function from S to the sum (convolution) of $r > 1$ exchangeable*[10] *lotteries over outcomes. The set of all rules is the* Regular Environment *and is denoted by $\mathcal{H}_{(r)}$.*

Note, that according to Definition 4, the set of acts constitute the *Singular Environment* in this setting. In the regular environment, every state (s) is assigned a "bundle" of lotteries, which are positively correlated. In the formal definition we assume exchangeability, i.e. 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 rule (in the regular environment) bundles a bet on all the type II or type I urns. We assume here that the lotteries at every state are exchangeable, which is a generalization of the "same color composition" in the type II urns above. For example, the rule $IIR_{(2)}$ 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 positively correlated with the risk associated with the state of the engine. The dimensionality of the regular environment is indexed by r. Consider the agent's preferences over the regular environment. She is indifferent between the rules $f_{(r)}$ and $g_{(r)}$ if:

[10] Exchangeable means here that *conditional* on the state the lotteries over outcomes are independent and identical.

$$U\left(f_{(r)}\right)=U\left(g_{(r)}\right) \tag{9}$$

Or, explicitly:

$$\sum_{s\in S} q(s)E\left[u\left(f_{(r)}\left(s\right)\right)\right]=\sum_{s\in S} q(s)E\left[u\left(g_{(r)}(s)\right)\right] \tag{10}$$

where $E\left[u\left(f_{(r)}\left(s\right)\right)\right]$ 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:

$$E\left[u\left(f_{(2)}\left(s\right)\right)\right]=\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{X}} f\left(s\right)\left(x\right)f\left(s\right)\left(y\right)u(x+y) \tag{11}$$

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

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. rules), "uncertainty aversion" is a consequence of a Bayesian prior and risk aversion.

Theorem 5 *If f and g are Statewise Ranked by FOSD and the agent is indifferent between the rule $f_{(2)}$ and the rule $g_{(2)}$, then if she is averse to mean preserving spreads and her preferences are representable by an expected utility functional, she will prefer the rule of $(f,\alpha;g,1-\alpha)_{(2)}$ over the rule $f_{(2)}$ for every $0<\alpha<1$.*

Proof. See Appendix. ∎

The implication of Theorem 5 is that almost all seemingly uncertainty averse behavior may be rationalized if the agent's perception is that a decision will span multiple ambiguous risks. Confronted with this environment, if she is risk averse her observed behavior would exhibit uncertainty aversion.

Uncertainty averse behavior may be fully rationalized if the individual assigns a small probability that the environment he 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. In this case uncertainty averse behavior is fully rationalized, within the expected utility framework.

Corollary 6 *Assume f and g as in Theorem 5, 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 rules - with probability β) over each act (or rule - 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:

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

■

4 Discussion and Conclusion

Increasing the dimensionality of the environments a decision maker considers, places the Ellsberg "paradox" as most singular. Arbitrary small perturbation of the environment leads to uncertainty averse behavior which is fully consistent with expected utility theory and Bayesian rationality. We argue that if one uses "rule rationality" then human behavior will exhibits "inertia" (cannot adjust to this singularity in the environment), and ambiguity aversion becomes a very plausible prediction.

4.1 Comparison with the literature

The Ellsberg paradox motivated a big literature that tried to explain this predicted behavior. In this short section we shall discuss only two alternative resolutions. The Maximin model, which was suggested (and rejected) by Ellsberg, and later axiomatized by Itzhak Gilboa and David Schmeidler [7], assumes the decision maker behaves pessimistically, choosing the act that maximizes her expected utility if the worst scenario occurs (maximin over a convex set of priors). Schmeidler [14] and Gilboa [6] derived the Choquet expected utility representation, which is - if the capacity is convex - a special case of the maximin. Uncertainty Aversion was first defined in this context. Although in this paper we show that almost every empirically observed uncertainty averse behavior may be rationalized as rule rationality, we point out that these are two distinct representations, and are not equivalent. We shall prove it by a short example. Let the utility function be:

$$u(x) = \begin{cases} x+5 & x \le -5 \\ \frac{x+5}{2} & -5 < x \le -3 \\ 1 & -3 < x \end{cases} \tag{12}$$

and assume two states of the world s, t with equal subjective probability. The two acts f, g are:

$$f(s) = g(t) = \begin{cases} -3 & 0.5 \\ -2 & 0.5 \end{cases} \qquad f(t) = g(s) = \begin{cases} -4 & 0.5 \\ -1 & 0.5 \end{cases} \tag{13}$$

The two acts are non-comonotonic and the individual is indifferent between them, therefore uncertainty aversion would claim he prefers the mixture of the two over each act separately. However, a short calculation shows that our explanation of preference over rules does not support uncertainty aversion in this case. Formally, these acts are not statewise ranked by FOSD.

In a second attempt, Uzi Segal [15] showed that if expected utility together with the axiom of reduction of compound lotteries are relaxed, the Ellsberg type behavior may be rationalized. The solution he obtains is conservative and ties uncertainty aversion to risk aversion - two features that are common with our work. The difference is that his explanation depends crucially on the non-expected utility structure of preferences.

4.2 Uncertainty Aversion

The Anscombe-Aumann framework has some complications due to its 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 [17], but the definition of uncertainty aversion in their framework is not transparent and will be different [16] of Schmeidler's. A definition of uncertainty aversion within a Savage domain of acts was suggested recently by Larry Epstein

[5]. His definition, defines uncertainty aversion relative to probabilistically sophisticated preferences. Generalization of the results presented in the previous section within Epstein's definition, may shed light on the degree of singularity of Ellsberg experiments and our generalization, and remains for future work.

Another extension of our work is to study more carefully the "uncertainty premium" presented in Section 2. We believe this concept has important applications in information economics, finance, and other economic fields, which could not have been analyzed until now due to the pessimistic characterization of the maximin.

4.3 "Rule Rationality" and other experimental anomalies

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 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 (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 find it hard to evaluate singular situations. Aumann [2] contrasted this *"Rule Rationality"* with "Act Rationality". Hoffman, MacCabe and Smith [8] 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[11].

The (speculative) relation between the decision theoretic problem studied in this paper, and other anomalies in interactive game theory, leads us to hypothesize that rule rationality is a form of bounded rationality that should be studied carefully. Specifically, experiments could determine whether certain individuals rely (more than others) on this sort of rationality. If this sort of "bounded rationality" is 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 to concentrate 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 perturbation of the environment.

Note that the repetitive structure of the other "game theoretic" cases could be imported into our framework, but at a cost. The formal model is silent whether the risks are bundled or repetitive. Then, the results could be interpreted as a "policy" for a sequence of ambiguous risks[12]. However, this interpretation is vulnerable to considerable limitations on the rules considered: the decision maker can not learn from one risk to the next, and can not alternate (hedge) between different ambiguous risks. Hence, the decision maker is not rational even in the repeated environment.

An 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, an evolutionary model in which "rule rationality" 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 Definition of Subjective Probability", *Annals of Mathematical Statistics*, 34, 199-205.

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

[12] A previous version of this work emphasized this aspect. Schneeweiss [18] analyzes the Ellsberg paradox assuming the number of repetitions approaches infinity and the utility function is quadratic.

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

[3] 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 Probability*, New York: John Wiley and Sons.

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

[5] Epstein, Larry G. (1999): "A Definition of Uncertainty Aversion", *Review of Economic Studies*, 66, 579-608.

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

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

[8] 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.

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

[10] Knight, Frank H. (1921): *Risk, Uncertainty and Profit*, Boston: Houghton Mifflin Company.

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

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

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

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

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

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

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

[18] Schneeweiss, Hans (1999): "Resolving the Ellsberg paradox by Assuming that People Evaluate Repetitive Sampling" (memo).

A Preliminaries

Let ψ and τ be finite measures on $\mathcal{X}$. Define[13]:

$$F_\psi(x) = \sum_{t \leq x} \psi(t) \text{ and } F_\tau(x) = \sum_{t \leq x} \tau(t) \tag{14}$$

Assume ψ and τ are such that:

$$F_\psi(+\infty) = F_\tau(+\infty) \tag{15}$$

Assumption (15) would hold true if, for example, ψ and τ are probability measures (then (15) is equal to one), or when each is a difference of two probability measures (then (15) is equal to zero).

Definition 7 *Let ψ and τ be two finite measures defined over $\mathcal{X}$, and let F_ψ and F_τ be defined as in (14) and satisfy (15). 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.*

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

Definition 7 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:

$$\int_{-\infty}^{+\infty} F_{\psi}(x)dx = \int_{-\infty}^{+\infty} F_{\tau}(x)dx \tag{16}$$

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.

Definition 8 ψ Second Order Stochastically Dominates (SOSD) τ *if (16) 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 9 *If ψ SOSD τ then:*

$$U(\psi) = \int_{-\infty}^{+\infty} u(x)\psi(x)dx > \int_{-\infty}^{+\infty} u(x)\tau(x)dx = U(\tau)$$

for all strictly monotone and strictly concave u.

Proof. The proof is similar to Rothschild and Stiglitz's [12]: using (15) instead of assuming probability measures, and (16) instead of assuming equal expectations. ■

B Proof of Theorem 5

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

$$U\left(f_{(2)}\right) = U\left(g_{(2)}\right) \tag{9'}$$

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

$$h(s)\,(x) = \alpha f(s)\,(x) + (1-\alpha)\,g(s)\,(x) \tag{17}$$

Then we need to show that:

$$U(h_{(2)}) > U\left(f_{(2)}\right) \tag{18}$$

Consider the function θ defined as:

$$\theta(s)\,(x) = f\,(s)\,(x) - g\,(s)\,(x) \tag{19}$$

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.

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

$$= \sum_s q(s) \sum_x \sum_y \begin{bmatrix} \alpha^2\,(f(s)(x))\,(f(s)(y)) + \\ +\,(1-\alpha)^2\,(g\,(s)\,(x))(g\,(s)\,(y)) + \\ +2\alpha(1-\alpha)(f(s)(x))(g(s)(y)) \end{bmatrix} u(x+y) \tag{20}$$

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):

$$
\begin{aligned}
U\left(\theta_{(2)}\right) &= \sum_s q(s)U\left[\theta(s) * \theta(s)\right] = \qquad (21)\\
&= \sum_s q(s)\sum_x \sum_y \theta\left(s\right)\left(x\right)\theta\left(s\right)\left(y\right)u(x+y) = \\
&= \sum_s q(s)\sum_x \sum_y \left[f(s)(x) - g\left(s\right)\left(x\right)\right]\left[f(s)(y) - g\left(s\right)\left(y\right)\right]u(x+y) = \\
&= \sum_s q(s)\sum_x \sum_y \left[\begin{array}{c} (f(s)(x))(f(s)(y)) + \\ +(g(s)(x))(g(s)(y)) - \\ -2(f\left(s\right)\left(x\right))(g\left(s\right)\left(y\right)) \end{array}\right]u(x+y) \qquad (22)
\end{aligned}
$$

By substitution of (20) and (22) and utilizing (9') it follows that:

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

Thus, (18) holds if and only if $U\left(\theta_{(2)}\right) < 0$.

Claim 10 *In every state in which f and g differ: $\theta(s)$ FOSD* **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 $f\left(s\right)$ FOSD $g(s)$. Then:

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

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

Lemma 11 *Let ξ be a function, which is the difference of two probability mass measures and assume ξ and* **0** *are ranked according to first order stochastic dominance. Then ξ can be written as a finite sum of measures:*

$$\xi = \sum_{l=1}^{L} \xi_l \qquad (23)$$

where:

$$\xi_l(x) = \xi_{a_l,b_l,p_l}(x) = \begin{cases} p_l & if \quad x = a_l \\ -p_l & if \quad x = b_l \\ 0 & \quad OTHERWISE \end{cases} \tag{24}$$

with $a_l < b_l$ and $|p_l| \leq 1$. If **0** *FOSD ξ (ξ FOSD* **0***) then all p_l can be chosen positive (negative) in the decomposition (24).*

Proof. Recall that since ξ is a difference of probability mass measures, it is a finite measure with $F_\xi(+\infty) = 0$. Assume **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:

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

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 $F_\xi \geq \xi(a_1) \geq p_1$. 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 finitely supported (i.e. there exist only finitely many points x such that $\xi(x) \neq 0$), 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 **0**. ■

Lemma 12 *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 (24) structure.*

Proof. The measure $\xi_l * \xi_k$ 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 13 *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 10 the zero function FOSD $\theta(s)$ or vice versa. By Lemma 11, we can decompose every difference measure $\theta(s)$ into $L(s)$ measures 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:

$$\theta(s) * \theta(s) = \left(\sum_{l=1}^{L(s)} \theta_l(s)\right) * \left(\sum_{k=1}^{L(s)} \theta_k(s)\right) = \sum_{l=1}^{L(s)} \sum_{k=1}^{L(s)} \theta_l(s) * \theta_k(s) \qquad (25)$$

By Lemma 12 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 5. Recall from (21) that $U\left(\theta_{(2)}\right)$ is additive across states. By Corollary 13 and Claim 9: $U\left[\theta(s) * \theta(s)\right] < 0$ 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 (18) holds. ■