*Econometrica*, Vol. 75, No. 2 (March, 2007), 503–536

# ELLSBERG REVISITED: AN EXPERIMENTAL STUDY

BY YORAM HALEVY[1]

An extension to Ellsberg's experiment demonstrates that attitudes to ambiguity and compound objective lotteries are tightly associated. The sample is decomposed into three main groups: subjective expected utility subjects, who reduce compound objective lotteries and are ambiguity neutral, and two groups that exhibit different forms of association between preferences over compound lotteries and ambiguity, corresponding to alternative theoretical models that account for ambiguity averse or seeking behavior.

KEYWORDS: Uncertainty aversion, probabilistic sophistication, reduction of compound lotteries, nonexpected utility, maxmin expected utility, anticipated utility, rank dependent utility, recursive utility, compound independence, bundling, rule rationality.

## 1. INTRODUCTION

IN 1961, ELLSBERG SUGGESTED several ingenious experiments that demonstrated that Savage's (1954) normative approach, which allows a modeler to derive subjective probabilities from the decision maker's preferences, faces severe descriptive difficulties. In particular, Ellsberg's (1961) examples showed that many decision makers have a nonneutral attitude to ambiguity: their choices reveal preferences that differentiate between *risk* (known probabilities) and *uncertainty* (unknown probabilities). The approach in which subjective belief (derived from preferences) substitutes for objective probability in the evaluation of an uncertain prospect—distilled in Savage's (1954) model of subjective expected utility and Machina and Schmeidler's (1992) work on *probabilistically sophisticated* preferences—has been challenged. The existence of subjective probability is crucial in economics, where its usage is pervasive. In many cases, not only do the results depend on the assumption that decision makers' preferences are described by the subjective expected utility model, but without the existence of subjective probability (probabilistic sophistication), defining the relevant problem would become much more difficult.

To cope with the apparent descriptive problem raised by Ellsberg, several normative and behavioral models have been suggested. Gilboa and Schmeidler (1989) formulated the influential model of maxmin expected utility (MEU), in

[1]A co-editor and three anonymous referees provided thoughtful comments and suggestions that improved this work substantially. I appreciate helpful discussions with Paul Beaudry, Gorkem Celik, Larry Epstein, Haluk Ergin, Itzhak Gilboa, David Green, Edi Karni, Dan Levin, Sujoy Mukerji, Emre Ozdenoren, Mike Peters, Zvi Safra, and Uzi Segal. Special thanks to Peter Wakker. I also thank workshop participants in Michigan, University of British Columbia (UBC), Hebrew University, Tel Aviv, Haifa, RUD 2003 in Milan, The 2004 Canadian Experimental and Behavioral Economics Workshop in Calgary, The 2005 Annual International Meeting of the ESA in Montréal, and the 2005 Econometric Society World Congress. Rodney Chan, Fernando Har, and Haifang Huang provided superb research assistance. The CASSEL staff, and especially Walter Yuan, provided outstanding support and assistance. Financial support from UBC's Hampton Fund is gratefully acknowledged.

which a decision maker has a set of prior beliefs (core of belief) and her utility of an act is the minimal expected utility in this set. Schmeidler (1989) derived the Choquet expected utility (CEU) model axiomatically. When these preferences exhibit global ambiguity aversion (convex capacity) they can be represented by the MEU model. Segal (1987) showed in an early thought-provoking paper that when ambiguity is modeled as a two-stage lottery, relaxing the axiom of reduction of compound lotteries (ROCL) and applying rank dependent utility (Quiggin (1982)) to evaluate the lotteries recursively leads (under reasonable assumptions) to ambiguity aversion (recursive nonexpected utility preferences (RNEU)). Several recent papers (Klibanoff, Marinacci, and Mukerji (2005), Ergin and Gul (2004), Nau (2006), Ahn (2003), Seo (2006)) studied the preferences of a decision maker who is expected utility on first-(subjective) and second-(objective) stage lotteries, and her attitude toward ambiguity is determined by the relative concavities of the two utility functions (the recursive expected utility model (REU)). Halevy and Feltkamp (2005) suggested that the decision maker's preferences may have been molded in an environment of bundled risks, where a behavior similar to ambiguity aversion is a consequence of aversion to mean preserving spreads. Applying the decision rule to the Ellsberg example leads to ambiguity aversion.

The goal of this work is to compare the performance of these theories in a controlled experimental environment, which is an extension of the original Ellsberg experiment. Subjects were presented with four urns that contained 10 (red or black) balls each. The first two urns were the standard Ellsberg urns: the first contained 5 red balls and 5 black balls (risky), and the second's color composition was unknown (ambiguous). The number of red balls in the third urn was uniformly distributed between 0 and 10. The fourth urn contained either 10 red or 10 black balls (with equal probability). The subjects were asked to bet on a color in each urn (before the color composition in urns 3 and 4 was known). Then, a reservation price for each urn was elicited using the Becker–DeGroot–Marschak (1964) mechanism.

While the first two urns test the known distinction between risk and uncertainty, the latter two urns (together with the first urn) are two-stage objective lotteries that have varying degrees of dispersion in the first stage: urn 1's first-stage lottery is degenerate (all the risk is resolved in the second stage), while urn 4's second-stage lotteries are degenerate (all the risk is resolved in the first stage). The experimental setup makes it possible to test the association between a nonneutral attitude to ambiguity and different violations of reduction of compound (objective) lotteries. The former states that subjective uncertainty cannot be reduced to risk or, in the language of Machina and Schmeidler (1992), the agent is not "probabilistically sophisticated" (see Epstein (1999)). Violation of ROCL implies that an agent who faces compound lotteries, does not calculate probabilities of final outcomes according to the laws of probability (e.g., multiply probabilities).

The theories considered have different predictions on the relative attractiveness of bets on the four urns. The MEU and the CEU, derived within the Anscombe and Aumann (1963) framework, assume that when faced with objective lotteries, the decision maker's choices abide by the expected utility axioms. Therefore, a decision maker described by these theories (interpreted within the preceding framework) will satisfy the ROCL axiom and will be indifferent between the three objective urns. The RNEU model assumes that the decision maker is indifferent between urns 1 and 4, and if she is ambiguity averse, then she should rank urn 3 as worse than the other two objective urns. An ambiguity averse decision maker who uses a rule that considers a bundle of lotteries (and may be consistent with an interpretation of the REU in which reduction of compound objective lotteries does not hold) will prefer urn 1 to urn 3 to urn 4.

The main premise used in analyzing the sample is that the population may be heterogeneous: different subjects have different patterns of choice that correspond to different theories that describe, in particular, if and how ROCL may fail. Consequently, the analysis did not look for a unique theory to explain the average decision maker, but rather tried to infer from the sample what the common choice patterns are in the population.

The results, as is evident from Table I, revealed a tight association between ambiguity neutrality and reduction of compound lotteries, consistent with the subjective expected utility model.

Further analysis clarified that the structure of the association between non-neutral attitudes to ambiguity and violations of ROCL is not uniform in the population of subjects: two different (although even in frequency) choice patterns emerged. One pattern corresponded to the theoretical predictions of the RNEU model suggested by Segal (1987, 1990), while the second may be generated by behavioral rules consistent with an environment of bundled risks (Halevy and Feltkamp (2005)) and can be represented by the REU model in which a decision maker does not reduce compound *objective* lotteries.

We conclude that a descriptive theory that accounts for ambiguity aversion should account, at the same time, for a violation of reduction of compound

TABLE I

THE ASSOCIATION BETWEEN ATTITUDES TO AMBIGUITY AND COMPOUND OBJECTIVE LOTTERIES

| | | ROCL | | |
|---|---|---|---|---|
| **Ambiguity Neutral** | | No | Yes | Total |
| No | Count | 113 | 1 | 114 |
| | Expected | 95.5 | 18.5 | |
| Yes | Count | 6 | 22 | 28 |
| | Expected | 23.5 | 4.5 | |
| Total | | 119 | 23 | 142 |

*Note*: This table combines the two samples reported in this study (see Table IV). Ambiguity neutrality is defined as $V1 = V2$ and ROCL is defined as $V1 = V3 = V4$. Fisher's exact test: Exact sig. (2-sided) 2.3E−19.

*objective* lotteries. Furthermore, a general theory of ambiguity aversion should account for different ways in which ROCL may fail.

## 2. METHOD

Two controlled experiments were conducted: an original experiment with moderate stakes and a robustness experiment in which all payoffs were scaled by a factor of 10. In each experiment, the subjects were asked to state their reservation prices for four different lotteries through an incentive compatible elicitation mechanism.

### 2.1. *The Lotteries*

Subjects were presented with four lotteries: the first two are the standard (two colors) Ellsberg urns, used to test ambiguity attitude. The latter two (together with the first one) tested whether behavior satisfies the ROCL assumption for objective lotteries. A graphical presentation of the lotteries is available in the online Supplement (Halevy (2007)). The description of the lotteries as presented to the subjects follows:

There are 4 urns,[2] each containing 10 balls, which can be either red or black. The composition of balls in the urns is as follows:

- *Urn 1*: Contains 5 red balls and 5 black balls.
- *Urn 2*: The number of red and black balls is unknown, it could be any number between 0 red balls (and 10 black balls) to 10 red balls (and 0 black balls).
- *Urn 3*: The number of red and black balls is determined as follows: one ticket is drawn from a bag containing 11 tickets with the numbers 0 to 10 written on them. The number written on the drawn ticket will determine the number of red balls in the third urn. For example, if the ticket drawn is 3, then there will be 3 red balls and 7 black balls.
- *Urn 4*: The color composition of balls in this urn is determined in a similar way to box 3. The difference is that instead of 11 tickets in the bag, there are 2, with the numbers 0 and 10 written on them. Therefore, the urn may contain either 0 red balls (and 10 black balls) or 10 red balls (and 0 black balls).

Each participant was asked to place a bet on the color of the ball drawn from each urn (Red or Black), eliminating problems that arise from potentially asymmetric information (e.g., Morris (1997)). If a bet on a specific urn was correct, the subject could have won \$2 Canadian. If a bet was incorrect, the subject lost nothing. The total money that could have been earned is \$8 (plus \$2 paid for participation). Before balls were drawn from each urn (and before the tickets were drawn from the bags for urns 3 and 4), the subject had the

[2]In the experiment, the word "box" was used, to minimize confusion among subjects who were not familiar with the word "urn."

option to "sell" each one of her bets. The Becker–DeGroot–Marschak (1964) mechanism (henceforth, BDM) was used to elicit (an approximation to) the certainty equivalent of each bet: the subject was asked to state four minimal prices at which she was willing to sell each one of the bets (reservation prices). The subject set the selling prices by moving a lever on a scale between \$0 and \$2. For each urn, a random number between \$0 and \$2 was generated by the computer. The four random numbers were the "buying prices" for each one of the bets. If the buying price for an urn was higher than the reservation price that the subject stated for that urn, she was paid the buying price (and her payoff did not depend on the outcome of her bet). However, if the buying price for the urn was lower than the minimal selling price reported for that urn, the actual payment depended on the outcome of her bet. The random numbers (which were generated by the computer program) and the outcomes of the draws were not revealed until all four reservation prices were set. The relevant data collected from each participant were the reservation prices she stated for each urn, as well as some personal information (all available in the online Supplement).

The BDM (1964) mechanism has been used extensively in the "preference reversal" literature (e.g., Grether and Plott (1979)).[3] Several researchers (Holt (1986), Karni and Safra (1987), Segal (1988)) have pointed out that when preferences do not satisfy the axioms that underlie expected utility theory, and in particular independence (Holt (1986), Karni and Safra (1987)) and reduction of compound lotteries (Segal (1988)), the BDM mechanism may not elicit valuations accurately. Holt's (1986) reservation—which applies to a situation in which several valuations are elicited, but the subject is paid the outcome of only one—has been fully accommodated in the current study because all outcomes are actually paid. Karni and Safra (1987) showed that the "certainty equivalent" of a lottery elicited utilizing the BDM mechanism respects the preference ordering if and only if preferences satisfy the independence axiom. Furthermore, they showed that there exists no incentive compatible mechanism that elicits the certainty equivalent and does not depend on the independence axiom. In the current experiment, the independence axiom does not play any role in the evaluation of urns 1, 3, and 4: if the subject multiplies probabilities correctly (reduces compound lotteries), then she would view all of them as the same lottery. Segal (1988) provided an example in which violation of ROCL (together with nonexpected utility) results in a preference reversal.[4] The latter limitation of the BDM mechanism is important in the current experiment

[3]The main focus of this literature is the relative ranking of two lotteries: one with high probability of winning a moderate prize ($P$ bet) and one with a low probability of winning a high prize (a $p$ bet). Many agents choose the $P$ bet over the $p$ bet, but the valuation of the $P$ bet is lower than the valuation of the $p$ bet. The valuations were elicited using a BDM mechanism.

[4]Furthermore, Keler, Segal, and Wang (1993) showed that under this latter interpretation, the certainty equivalent and the value elicited using BDM may lie on different sides of the expected value of the lottery. Therefore, risk attitude would be impossible to infer from the elicited value.

that focuses on the relationship between ambiguity aversion and violations of ROCL. This problem does not exist for theories that satisfy both reduction of compound lotteries and the independence axiom when only objective probabilities are involved. Even in the REU model (when ROCL is violated) the value elicited equals the certainty equivalent of the lottery. Furthermore, in the RNEU, although the value elicited does not have to be equal to the certainty equivalent, the value elicited for urns 1 and 4 should be equal (consistent with the theoretical predictions in (6) in Section 3).

An important practical problem is that the BDM mechanism is complicated, and if subjects fail to understand it, the elicited values might reflect their confusion and not their evaluation. To minimize the confusion effect, before the experiment the subjects received an extensive explanation of the BDM mechanism and all experienced it in a trial round[5] before the actual experiment. Furthermore, even if the BDM mechanism elicits the valuation with some noise, the patterns of responses in the data are extremely robust and consistent with some of the theoretical predictions. As a result, the concerns raised in the preceding text do not render the data useless.

### 2.2. *Robustness Test*

The original experiment as presented in Section 2.1 may be subject to several imperfections: the price of $2 may seem too small to give the subject sufficient incentives to think seriously of the problems at hand; although the research assistant tried to confirm that the participants understood the BDM mechanism, there was no objective measure of his success; the subjects were asked to behave as "sellers"—a framework that might have influenced their reservation prices; the recruitment of the subjects was based on sign-up sheets. This convenience sampling technique might have introduced biases into the experiment's results. To counter these reservations it might be argued that none of the arguments is systematic, and if it introduces biases, there is no a priori reason to believe they have a differential effect on the reservation prices set for the four urns. In particular, because the focus of the current study is on the relative reservation prices, the conjecture is that similar results would hold in an altered experiment in which these deficiencies would be corrected. To test this conjecture, a robustness experiment was conducted. The prizes were scaled from $2 to $20 per urn. The efficiency of the sampling was increased, thanks to the use of proportional sampling within cohorts, as described subsequently. In addition, although in the first round the subjects were informed of the range of possible payoffs ("earn up to $10"), no definite amount was disclosed in the robustness round before the experiment itself (most subjects did not expect the payoffs to be as high).

[5]The second (ambiguous) urn was a "new" urn, eliminating the possibility of learning from the trial round. This information was conveyed to the participants.

The subjects received two paid opportunities to experience the operation of the BDM mechanism, before the trial round. They were given a toonie ($2 Canadian coin) and were asked to set their minimal reservation price for it. This task was used to guide them how to "find" their minimal reservation price: they were prompted to consider if they would accept five cents less than their stated reservation price. If they accepted, the process was repeated until the minimal reservation price was achieved. Because in this task there is an objectively correct answer, the subjects learned which reservation prices were "too low" and which were "too high." Next, the subjects were given a pen (with a retail value of $2.50 at UBC's bookstore, which was not revealed to the subjects) and were asked to set their minimal reservation price for it using the same mechanism. Only then were they offered a trial round with the four urns. Throughout the experiment the subjects were reminded how to find their minimal reservation price. Furthermore, in the instructions to the experiment, the terms "selling/buying price" and "true valuation" were not used.

To prevent any possibility the subjects might have suspected they were being tricked, the implementation of the experiment was altered somewhat. The lotteries were physical and not computerized: there were four pouches that contained beads that could be red or black. The composition of pouches 3 and 4 was determined by choosing at random a numbered token (1 out of 11 or 1 out of 2, respectively), and the composition of pouch 1 could have been verified by the subject. The random numbers (between $0 and $20) were generated by a computer before the experiment and they were organized in a matrix. The subjects chose at the beginning of the experiment 10 different coordinates,[6] which were revealed sequentially after they set reservation prices for the different urns. In addition, an order treatment was implemented: the subjects were randomly allocated to different orders of urns: $(1, 2, 3, 4)$ (as in the original experiment), $(2, 3, 4, 1)$, $(3, 4, 1, 2)$, and $(4, 1, 2, 3)$. The goal of the random ordering was to test whether the reservation prices were influenced by alternative ordering schemes.[7]

### 2.3. *Recruitment*

One hundred and four subjects were recruited for the original experiment using ads posted at different locations on the University of British Columbia (UBC) campus. The subjects signed up for sequential time slots. Thirty-eight

[6]The first two coordinates were used to teach the BDM elicitation mechanism (range of 0–4), while the latter 8 were used for the trial round and the paid experiment.

[7]Harrison, Johnson, McInnes, and Rutström (2005) found a significant order effect (in addition to scale effect) in Holt and Laury's (2002) study of the effect of higher scale of *real* incentives on risk aversion. In a followup study, Holt and Laury (2005) showed that the magnitude of the scale effect is robust to the elimination of the order effect.

subjects participated in the robustness experiment. The recruitment of subjects to the second round was based on proportional sampling within each cohort of undergraduate students in the faculties of Arts and Science at UBC (about 12,500 students): 100 students were sent e-mail invitations and about half of them responded, out of which 38 students participated in the experiment (mainly due to scheduling conflicts).

### 2.4. *Administration*

During each experiment, only one subject was present in the room. Following her arrival, each subject signed a consent form (available in the online Supplement) that explained the experiment. In the first experiment, a research assistant was always in the room to answer any concerns and to make sure the subject knew how to run the computer program that simulated the lotteries. The computer program was written by CASSEL's[8] staff. The second experiment was supervised by the author and all the lotteries were executed by physical randomization devices.

### 2.5. *Related Experimental Literature*

Two previous experimental studies added an objective two-stage lottery to the classic two color Ellsberg example. Yates and Zukowski (1976) tested the "range hypothesis"[9] by offering urns similar to the first three in the current study. Each subject was allowed to choose one urn out of the possible three pairs of urns. The value of the chosen lottery was elicited using the BDM mechanism. Yates and Zukowski (1976) found evidence that urn 1 was weakly preferred to urn 3, which was weakly preferred to urn 2. Yates and Zukowski's evidence should be treated with care, because they averaged over different subjects who were offered different choice sets.

Chow and Sarin (2002) tested the distinction between known (risk), unknown, and unknowable uncertainties using urns 1, 2, and 3, respectively. They found that unknowable uncertainty is intermediate to the known and the unknown forms of uncertainties. They related their findings to Fox and Tversky's (1995) "comparative ignorance hypothesis," in which the availability of an informed agent (experimenter) decreases the attractiveness of a lottery.

[8]California Social Science Experimental Lab, which is a joint project of UCLA, California Institute of Technology, and the NSF. The basic version of the software can be downloaded from: http://multistage.ssel.caltech.edu/extensions/Individual_Decisions/multistage_decision/multistage_decision.html.

[9]The range hypothesis claims that the range of the second-order distribution is the critical element in accounting for the attractiveness of an "ambiguous" lottery. Hence, urn 3 has the largest range, and should be (weakly) inferior to the second (ambiguous) urn.

## 3. THEORETICAL PREDICTIONS

This section describes how different theories of choice under uncertainty predict individual choices in the experiment. The theories that were focused on during the research include subjective expected utility (SEU), maxmin expected utility (within the Anscombe–Aumann framework), RNEU, REU, and the "bundling" rationale to ambiguity aversion.

A state in the experiment is the quadruple $(s_1, s_2, s_3, s_4)$, where $s_i \in \{r, b\}$ $(i = 1, \dots, 4)$, representing that color $s_i$ was drawn from urn $i$. Therefore, the experiment's state space has 16 states, and a bet on urn $i$ $(i = 1, \dots, 4)$ has two payoff-relevant events: a red ball is drawn from urn $i$ ($\{s_i = r\}$, denoted by $ir$) and a black ball is drawn from urn $i$ ($\{s_i = b\}$, denoted by $ib$).

### 3.1. *Subjective Expected Utility*

According to Savage's (1954) theory, the decision maker assigns subjective probability $\pi(ij)$ to the event in which a ball of color $j \in \{r, b\}$ is drawn from urn $i$. The (maximal) subjective expected utility of a bet on urn $i$ ($L_i$) is therefore

$$U_{\mathrm{SEU}}(L_i) = \max_{J \in \{R,B\}} \sum_{s \in \{ir, ib\}} \pi(s) u(JL_i(s)).$$

In the following discussion, it is assumed that when probabilities are given objectively, the subjective and objective probabilities coincide. For example, the objective probability of drawing red from urn 1 is one-half, and it is assumed that $\pi(1r) = 0.5$. Furthermore, although strictly speaking Savage's axioms are not stated in a dynamic framework, many works have shown that reduction of compound objective lotteries is a necessary part of expected utility theory when the latter's domain includes two-stage lotteries (e.g., Segal (1990, Theorem 3), Anscombe and Aumann (1963)). Therefore, such a decision maker states the same reservation price for urns 1, 3, and 4. It is possible that the decision maker believes that urn 2 has more red or more black balls. That is, $\pi(2r)$ may be different from one-half. If this is the case, she may choose to bet on the more probable color (in which a higher subjective expected utility is attained) and set a higher reservation price for it. Denote by $Vi$ the reservation price for urn $i$ $(i = 1, \dots, 4)$. Then

$$(1) \qquad V1 = V3 = V4 \leq V2.$$

If the decision maker is an expected value maximizer, then $\mathbb{E}(L_1) = V1 = V3 = V4$.

### 3.2. *Maxmin Expected Utility*

A decision maker whose preferences are described by MEU (Gilboa and Schmeidler (1989)) has a set of prior beliefs (core of belief) and her utility of an act is the minimal expected utility on this set. The utility of betting on urn $i$ is therefore

$$U_{\mathrm{MEU}}(L_i) = \max_{J\in\{B,R\}} \min_{\pi\in\mathrm{core}} \sum_{s\in\{ir,ib\}} \pi(s)u(JL_i(s)).$$

Because MEU is a generalization of expected utility, it allows for a pattern of reservation prices as in (1). If the core is not degenerate to a unique prior, it can accommodate the typical choice pattern suggested by Ellsberg: $RL_2 \sim BL_2 \prec BL_1 \sim RL_1$. For example, suppose that the core contains the two "pessimistic" nonsymmetrical beliefs that all balls are red and that all balls are black.[10] Then (if the prize is \$$x$) $U_{\mathrm{MEU}}(L_2) = u(0) < 0.5u(0) + 0.5u(x) = U_{\mathrm{MEU}}(L_1)$. Within the Anscombe and Aumann (1963) framework used by Gilboa and Schmeidler (1989), MEU reduces to expected utility in the realm of objective probabilities (urns 1, 3, and 4). Therefore, the decision maker reduces compound objective lotteries and is indifferent between bets on urns 1, 3, and 4. That is, if the core of belief includes a probability measure $\pi$, such that $\pi(2r) = 0.5$, then

(2) $\quad V2 \leq V1 = V3 = V4,$

where strict inequality follows if the core includes other measures that assign probability that is different from one-half to the event that a red ball is drawn from urn 2. Note that within the Choquet expected utility model (Schmeidler (1989)), which also uses the Anscombe and Aumann (1963) framework, reduction of compound objective lotteries holds. The case of convex capacity (which corresponds to ambiguity aversion) is a special case of the MEU model.

It should be noted that when the MEU or the CEU is derived in a purely subjective world (as in Casadesus-Masanell, Klibanoff, and Ozdenoren (2000), Gilboa (1987)) they make no prediction as to how the decision maker evaluates compound objective lotteries. Therefore, it is perfectly conceivable that if she is ambiguity averse, then she will not reduce compound objective lotteries. The structure of preferences on compound objective lotteries that are consistent with MEU and CEU preferences in a purely subjective world may give an important theoretical perspective on the findings of the present study, and is left for future work.

### 3.3. *Recursive Nonexpected Utility*

Segal (1987, 1990) relaxed the ROCL axiom and applied Quiggin's (1982) rank dependent utility (RDU; also known as anticipated utility) to evaluate

[10] This is done for expositional purposes only. Nothing in Gilboa and Schmeidler's (1989) axioms that underlie the MEU representation forces these extreme priors to be elements of the core.

the first-and second-stage lotteries.[11] To better understand Segal's theory, let $x_1 \leq x_2 \leq \cdots \leq x_n$. The RDU of the lottery that pays $x_i$ with probability $p_i$ $(i = 1, \ldots, n)$ is

$$U_{\mathrm{RDU}}(x_1, p_1; x_2, p_2; \ldots; x_n, p_n) = u(x_1) + \sum_{i=2}^{n} [u(x_i) - u(x_{i-1})] f\left(\sum_{j=i}^{n} p_j\right), \tag{3}$$

where $f : [0, 1] \to [0, 1]$, and $f(0) = 0$ and $f(1) = 1$. The RDU of the simple lottery that gives a prize of $\$x$ with probability $p$ and $\$0$ with probability $1 - p$ (after normalizing $u(0)$ to 0) is therefore

$$U(x, p; 0, 1 - p) = u(x) f(p) \tag{4}$$

and its certainty equivalent is

$$\mathrm{CE}(x, p; 0, 1 - p) = u^{-1}(u(x) f(p)). \tag{5}$$

To demonstrate Segal's approach, assume that the decision maker's model of the ambiguous urn ($L_2$) is that with probability $\alpha$, it contains 10 red balls; with probability $\alpha$, it contains 0 red balls; and with probability $(1 - 2\alpha)$, it contains 5 red balls. That is, the decision maker has a second-order subjective belief over the possible probability distributions over the states. If the agent bets on red from urn 2, then she first evaluates the second-stage lotteries ($\$x, 1; \$0, 0$), ($\$x, 0; \$0, 1$), and ($\$x, 0.5; \$0, 0.5$) using (4). Then, she evaluates the ambiguous lottery by substituting the certainty equivalents (calculated from (5)) as the prizes in (3):

$$\begin{aligned}
&U_{\mathrm{RNEU}}(RL_2) \\
&\quad = u\big[u^{-1}(v(0))\big] + \big[u\big(u^{-1}(u(x) f(0.5))\big) - u\big(u^{-1}(u(0))\big)\big] f(1 - \alpha) \\
&\qquad + \big[u\big(u^{-1}(u(x))\big) - u\big(u^{-1}(u(x) f(0.5))\big)\big] f(\alpha) \\
&\quad = 0 + [u(x) f(0.5) - 0] f(1 - \alpha) + [u(x) - u(x) f(0.5)] f(\alpha) \\
&\quad = u(x)\big[f(0.5) f(1 - \alpha) + (1 - f(0.5)) f(\alpha)\big] \\
&\quad < u(x) f(0.5) \\
&\quad = U(RL_1),
\end{aligned}$$

[11]The RNEU model is not restricted to RDU: other models of decision making under risk, as weighted utility, could be applied and similar predictions would be attained. The critical assumptions are Segal's (1990) "time neutrality" and "compound independence." It should be noted, however, that although the term "nonexpected utility" is commonly used to indicate a generalization of expected utility theory, the model suggested by Segal, imposes different (not weaker) restrictions on the data than the recursive expected utility model (see subsequent). Therefore a reader may wish to think of this model as recursive rank dependent utility (RRDU).

where the last inequality follows from the convexity of $f$, which in this theory is a necessary condition for risk aversion (Chew, Karni, and Safra (1987)), and reasonable properties of the transformation function $f(\cdot)$.[12]

Segal's (1987) novel interpretation of the Ellsberg paradox identifies ambiguity with a compound lottery, which the decision maker might fail to reduce. The critical feature of this model for the current experiment is that the certainty equivalent of a compound lottery is not monotone in the dispersion of the first-stage lottery. In particular, the decision maker is indifferent between bets on urns 1 and 4 (this is the time neutrality assumption in Segal (1990)). As before, the decision maker may believe that there are more red or more black balls in the second urn and prefer to bet on the (subjectively) more probable color. Hence, the prediction of Segal's theory is that the decision maker will be indifferent between urns 1 and 4, and prefer them (under the conditions specified in footnote 12) to a bet on urn 3. Indifference between the three objective urns results if $f$ is the identity function (in which case RDU reduces to EU) and the reduction of compound lotteries holds. Hence, in terms of elicited valuations the theory's predictions are

$$\text{(6)} \qquad (V1 = V4), \quad (V1 > V2 \Rightarrow V1 > V3), \quad \text{and}$$
$$(V3 > V1 \Rightarrow V2 > V1).$$

That is, the recursive nonexpected utility model predicts a negative correlation between $V21$ $(= V2 - V1)$ and $V43$ $(= V4 - V3)$, because ambiguity aversion $(V1 > V2)$ implies that $(V4 > V3)$, and $(V4 < V3)$ implies ambiguity seeking $(V1 < V2)$.

### 3.4. *Recursive Expected Utility*

Klibanoff, Marinacci, and Mukerji (2005; henceforth KMM) studied the preferences of a decision maker who has an expected utility on first- and second-stage lotteries, but whose ambiguity attitude is determined by the relative concavities of the two utility functions. To understand KMM's (2005) model, consider a bet on urn 2. The payoff-relevant partition of the state space is $\{2r, 2b\}$. To simplify notation, denote by $\pi$ a marginal probability distribution over the preceding partition (derived from an underlying belief over the finer state space). For each $\pi$, the decision maker calculates its certainty equivalent according to a von Neumann–Morgenstern utility index $u$. The decision maker has a subjective prior $\mu$ over the possible $\pi$ and evaluates an act using subjective expected utility according to the utility index $v$ with respect to $\mu$,

[12] Segal (1987) proved in Theorem 4.2 that if, in addition to convexity, $f$ has nondecreasing elasticity and $\bar{f} = 1 - f(1 - p)$ has nonincreasing elasticity, the decision maker will prefer a degenerate compound lottery (like $L_1$) to a compound lottery like $L_3$ or a subjective compound lottery like the foregoing $L_2$.

substituting the certainty equivalents (calculated from $u$) of the objective lotteries for every $\pi$.[13]

For example, suppose the support of $\mu$, the set of possible objective probabilities, is composed of $\pi_1 = (1, 0)$—no blacks in urn 2, and $\pi_2 = (0, 1)$—no reds in urn 2. The decision maker evaluates a bet on red from the ambiguous urn using the subjective prior $\mu = (\pi_1, \frac{1}{2}; \pi_2, \frac{1}{2})$. That is, the *subjective* probability that urn 2 has only red balls ($\pi_1$) is equal to the subjective probability that it has all blacks ($\pi_2$), which is equal to $\frac{1}{2}$ (similar to the *objective* urn 4). Given this belief, the decision maker's evaluation of a bet on either color from the ambiguous urn is

$$U_{\mathrm{REU}}(JL_2) = \tfrac{1}{2}v\big(u^{-1}(u(x))\big) + \tfrac{1}{2}v\big(u^{-1}(u(0))\big).$$

Let $\phi = v \circ u^{-1}$. Then

$$U_{\mathrm{REU}}(JL_2) = \tfrac{1}{2}\phi(u(x)) + \tfrac{1}{2}\phi(u(0)).$$

Klibanoff, Marinacci, and Mukerji (2005) generalized this representation and showed that the utility of an act $f$ is given by the functional $U_{\mathrm{REU}}(\cdot)$,

$$U_{\mathrm{REU}}(f) = \sum_{\pi \in \Delta} \phi\left(\sum_{s \in S} u(f(s)) \Pr(s|\pi)\right)\mu(\pi), \tag{7}$$

where $\Delta$ is the set of all possible first-order (second-stage) objective lotteries. In addition, KMM (2005) defined "smooth ambiguity aversion" and showed that it is equivalent to $\phi$ being concave. Therefore, it is equivalent to aversion to mean preserving spreads of the expected utility values induced by the second-order (first-stage) subjective probability ($\mu$) and the act $f$. However, when $\mu$ is given objectively, there is no behavioral reason to expect the decision maker to have differential risk attitudes when evaluating lotteries ($u$) and second-order acts ($v$), which induce *identical* objective probability distributions over outcomes. In this case, $v$ would be an affine transformation of $u$ and ROCL would apply. As a result, a decision maker whose preferences are described by (7) will be indifferent between urns 1, 3, and 4:

$$V1 = V3 = V4. \tag{8}$$

However, being strictly formal, lotteries and second-order acts (even when the second-order distribution is objective) are different mathematical concepts. Hence, it is possible that $v(\cdot)$ would be more concave than $u(\cdot)$ even when

[13]Note that $\pi$ is sometimes called second-stage (or first-order) objective lottery, while $\mu$ is first-stage (or second-order) subjective lottery.

$\mu$ is objective. If this is the case, the decision maker will evaluate urns 1, 3, and 4 in the manner

$$U_{\mathrm{REU}}(JL_1) = \phi(0.5u(x) + 0.5u(0)),$$
$$U_{\mathrm{REU}}(JL_3) = \frac{1}{11}\sum_{r=0}^{10}\phi\left(\frac{r}{10}u(x) + \left(\frac{10-r}{10}\right)u(0)\right),$$
$$U_{\mathrm{REU}}(JL_4) = 0.5\phi(u(x)) + 0.5\phi(u(0)),$$

and the reservation prices will satisfy[14]

$$(V1 \geq V3 \geq V4 \Rightarrow V2 \geq V4),$$
$$(V1 \geq V2 \Rightarrow V1 \geq V3 \geq V4), \tag{9}$$
$$(V1 \leq V3 \leq V4 \Rightarrow V1 \leq V2).$$

That is, if the subjective prior belief over the composition of the second urn is symmetric around 0.5 and nondegenerate, we would expect a positive correlation between $V43$ $(= V4 - V3)$ and $V21$ $(= V2 - V1)$, because then $V4 \leq V3$ if and only if $V1 \geq V2$.

It is important to note that such an interpretation requires a behavioral argument as to why the decision maker should be more averse to second-order acts (where all the uncertainty is resolved in the first stage) than to lotteries. The bundling model (Halevy and Feltkamp (2005)) presented in the following section suggests one possible source for such divergence.

Ergin and Gul (2004) suggested that ambiguity aversion is related to "issue preference."[15] That is, an agent may prefer an act that depends on one issue (risk) over an act that depends on another issue (ambiguity). Ergin and Gul provided an axiomatic foundation for "second-order probabilistically sophisticated" preferences—being able to assign subjective probabilities to the two issues, but allowing strict preference of a bet that depends on one issue over another. They showed that if the agent's preferences satisfy the sure thing principle or a comonotonic sure thing principle, then ambiguity aversion (in the sense of Schmeidler (1989)) is equivalent to "second-order risk aversion,"

[14] Note that if the subjective prior belief over the composition of the ambiguous (second) urn—$\mu(\pi(2r))$—is not symmetric around 0.5 (5 red balls and 5 black balls), the decision maker may prefer to bet on the ambiguous urn over all the risky urns (1, 3, and 4). If the subjective prior belief over the composition of the second (ambiguous) urn is symmetric around 5 red and 5 black balls, then a decision maker who is smooth ambiguity averse will rank urns 1, 2, and 4 in the following way: $V1 \geq V2 \geq V4$. If $\mu$ is degenerate on 5 red and 5 black balls then $V1 = V2$, whereas if the subjective prior is extreme (as the objective fourth urn), then $V2 = V4$.

[15] Similar preference has been previous named source preference (Tversky and Wakker (1995)).

which is aversion to mean preserving spreads in the subjective belief. The representation derived is identical (in the case of expected utility) to (7). Formally, Ergin and Gul (2004) could allow for issue preference even when the two issues were generated objectively, much in the same vein that Savage's approach could formally allow disparities between subjective and objective probabilities when the state is generated objectively. Therefore, although Ergin and Gul's (2004) model is formally consistent with (9), justifying this pattern might be problematic: if both issues are objective, it is not clear why a decision maker would/should prefer one over the other. The bundling model (Halevy and Feltkamp (2005)) described in the next section offers one possible explanation for such a pattern of preferences.

To facilitate understanding of Ergin and Gul's (2004) idea, consider a decision maker who has to choose a bet on a ball drawn from one of two urns: urn 1 has one red and one black ball, and urn 4 has either two red balls or two black balls. The payoff-relevant state space is $\{(1r, 4r), (1r, 4b), (1b, 4r), (1b, 4b)\}$. If, for example, the decision maker decides to bet on $1R$ (red from urn 1), and the state $(1r, 4b)$ obtains, she will win a prize of $\$x$. Ergin and Gul's innovation is to decompose the states into two issues, as follows:

$$\begin{array}{ccl} (1r, 4r) & (1r, 4b) & R \text{ from urn } 1 \\ (1b, 4r) & (1b, 4b) & B \text{ from urn } 1 \\ rr & bb & \end{array}$$

The right column describes "risk" associated with which ball is drawn from urn 1 (issue $a$), while the bottom row describes "uncertainty" regarding the composition of urn 4 (issue $b$). Assume that the decision maker considers every possible resolution of issue $a$ equally likely (given objectively) and that she considers every resolution of issue $b$ equally likely (given either objectively or subjectively). Furthermore, the two issues are statistically independent. A bet on urn 1 is a bet that depends only on issue $a$, and a bet on urn 4 is a bet that depends only on issue $b$. A decision maker may have issue preference and, as a result, prefers a bet that depends on issue $a$ (e.g., $1R$) to a bet that depends on issue $b$ (e.g., $4R$). To see how this relates to aversion to mean preserving spreads in the second-order distribution, note that the prior is uniform. That is, the probability of each state is 0.25. Let $\pi_\alpha$ be the lottery that yields $\$x$ with probability $\alpha$ and $\$0$ otherwise. Let $\mu_f$ be the compound lottery associated with an act $f$. For example, a bet on urn 4 (e.g., $4R$) induces with probability $\frac{1}{2}$ a lottery $\pi_1$ (if the state of urn 4 is $rr$) and with probability $\frac{1}{2}$ a lottery $\pi_0$ (if the state of urn 4 is $bb$). A bet on urn 1 induces with probability 1 the lottery $\pi_{0.5}$. Thus, the issue preference translates to what Ergin and Gul (2004) call second-order risk aversion (for further discussion, see Section 3 in Ergin and Gul (2004)).

The Kreps and Porteus (1987) model of decision making over temporal (objective) lotteries does not concern ambiguity, but has exactly the same two-stage recursive structure, with expected utility at each stage, as in (7). Smooth ambiguity aversion (as in KMM (2005)) or second-order risk aversion (as in Ergin and Gul (2004)) correspond to preference for late resolution in the Kreps and Porteus framework.[16] Segal (1990) provided perspective on the relationship between the RNEU and REU models: whereas the Kreps–Porteus REU model is derived by relaxing the time neutrality and ROCL axioms (maintaining mixture independence and compound independence), the RNEU is derived by relaxing the mixture independence and ROCL axioms (maintaining time neutrality and compound independence).

### 3.5. *Bundling and "Rule Rationality"*

A complementary "behavioral" perspective on ambiguity aversion was suggested by Halevy and Feltkamp (2005): if more than a single ball (bundle) may be drawn from each urn and the prize is determined as the sum (or average) of the correct bets, a decision maker who is averse to mean preserving spreads (Rothchild and Stiglitz (1970)) will prefer a bet on the risky (first) urn to a bet on the ambiguous (second) urn.[17] Halevy and Feltkamp (2005) suggested that the behavior observed in the actual experiment (in which only one ball is drawn from each urn) may be the result of *rule rationality*: the criterion that prefers risk to ambiguity is appropriate in the environment of bundled risks, and because it is hard wired into the decision making process, it is applied to the standard experiment (in which the decision maker is actually indifferent between the urns).

To demonstrate the bundling rationale to ambiguity aversion, let the decision maker hold a second-order prior belief (objective or subjective) over the composition of each urn. Each composition induces a first-order probability distribution over outcomes. Assume two draws with replacement are taken from each urn. The payoff distributions from betting on either red or black

[16] Four other recent works generalize this recursive structure. Nau (2006) allowed for state dependent preferences; Chew and Sagi (2003) studied the possibility of maintaining probabilistic sophistication on separate domains while distinguishing between different sources of uncertainty, hence not being globally probabilistically sophisticated. Ahn (2003) did not impose an exogenous state space, and did not distinguish between subjective and objective uncertainty. As a result, he presented an axiomatic foundation for a representation similar to (7), where the interpretation of differentiating between sources of objective uncertainty emerges naturally. Seo (2006) used the original Anscombe and Aumann (1963) framework (with preferences defined over objective lotteries over act lotteries.) Using a novel axiomatic approach, he derived (7) when the decision maker is ambiguity averse if and only if she violates ROCL.

[17] An alternative interpretation is of a decision maker who has to choose between two possible sequences of random outcomes—risky and ambiguous—and is constrained to decide ex ante on a unique color to bet on in each sequence (that is, always has to bet on the same color).

in urns 1, 3, and 4 are given by

| | $L_{1(2)}$ | $L_{3(2)}$ | $L_{4(2)}$ |
|---|---|---|---|
| \$2$x$ | 0.25 | 0.35 | 0.5 |
| \$$x$ | 0.5 | 0.3 | 0 |
| \$0 | 0.25 | 0.35 | 0.5 |

where $L_{i(2)}$ represents the random variable generated by two draws with replacement from urn $i$. These distributions are the result of averaging binomial distributions, using the second-order (first-stage) objective probabilities. To be more specific, let $k$ denote the number of red balls in urn $i$, corresponding (in the case of a bet on red) to a first-order lottery of $(x, \frac{k}{10}; 0, \frac{10-k}{10})$. Then the probability of drawing **two** red balls when betting on red is $(\frac{k}{10})^2$. Averaging over $k = 0, 1, \ldots, 10$ using the respective second-order probabilities for the different urns results in

$$\Pr\{L_{1(2)} = 2x\} = 1 \cdot \left(\frac{5}{10}\right)^2 = 0.25,$$

$$\Pr\{L_{3(2)} = 2x\} = \sum_{k=0}^{10} \frac{1}{11}\left(\frac{k}{10}\right)^2 = 0.35,$$

$$\Pr\{L_{4(2)} = 2x\} = \frac{1}{2} \cdot 1^2 + \frac{1}{2} \cdot 0^2 = 0.5.$$

The decision maker may have any second-order belief over the composition of the second (ambiguous) urn. As long as this subjective belief is symmetric and nondegenerate around the event that urn 2 contains 5 red balls, she will exhibit ambiguity aversion. That is, she will prefer to bet on the first urn rather than the second urn. Furthermore, if the decision maker is averse to mean preserving spreads, then for any second-order belief over the composition of the second urn, she will weakly prefer a bet on the second urn to a bet on the fourth urn. As a result, the predictions of the bundling theory coincide with an interpretation of the REU model (9) in which reduction of compound *objective* probabilities is violated. It is important to emphasize that the bundling rationale to ambiguity aversion does not depend on a distinction between objective and subjective second-order probabilities (similarly to Segal (1987), Seo (2006)), and could explain why a decision maker exhibits smooth ambiguity aversion (KMM (2005)) or issue (source) preference/second-order risk aversion (Ergin and Gul (2004)) with second-order objective probabilities.

## 4. RESULTS

The 104 subjects who participated in the first round of the experiment were paid a total of \$613 Canadian, which is about \$5.9 on average. The 38 subjects in the robustness experiment were paid a total of $1,948 Canadian (about $51 on average). The analysis will concentrate on the reservation prices $(V1, V2, V3, V4)$ set by the subjects.[18] Descriptive statistics are reported in Table II.[19]

The table reveals the anticipated pattern in the aggregate on both rounds: the average reservation price set by the subjects for urn 1 is higher than those set for the other urns, and the average price set for the ambiguous urn (urn 2) is the lowest. Moreover, in both samples, the distribution of reservation prices for the first risky urn $(V1)$ first order stochastically dominates (FOSD) the distribution of reservation prices for the ambiguous urn $(V2)$. In both samples, $V1$ does not FOSD $V3$ or $V4$, and the latter two do not FOSD $V2$.[20] To test statistically whether the valuations of the urns are different, a Friedman test[21] was performed to test the null hypothesis that the four valuations came from the same distribution. In the first experiment, the hypothesis that the assignment of the valuations is random was rejected (Friedman test is $\chi^2(3, N = 104) = 29.55,\ p < 0.001$). A similar test that is performed on the valuations of urns 2, 3, and 4 (the ambiguous and the nondegenerate compound lotteries) could not reject the null hypothesis that the valuations of the three urns came from the same distribution ($\chi^2(2, N = 104) = 0.33$, $p = 0.847$).[22] In the robustness test, the nonparametric Friedman test rejected the null hypothesis that the four reservation prices came from the same distribution ($\chi^2(3, N = 38) = 13.7,\ p < 0.0033$). When comparing urns 2, 3, and 4 in the second round, the Friedman test was inconclusive ($\chi^2(2, N = 38) = 7.43$, $p < 0.0245$).[23] To sum up both rounds: the reservation prices set for urn 1 are

[18] Although data on age, gender, exposure to mathematics and economics courses, and years of study were collected, none (except one that will be discussed subsequently) of these variables seems to be related to the reservation prices, in general, and measures of ambiguity aversion, in particular.

[19] The term $Vi$ is the reservation price set for urn $i$; AVE, STD, MAX, and MIN are the average, standard deviation, maximum, and minimum, respectively, in the four urns; $Vij = Vi - Vj$.

[20] If one studies the average reservation prices, a possible interpretation could be that the ranking is based on perceived simplicity of the lotteries (according to the order 1, 4, 3, 2). However, as one studies response patterns (for example, in Table V), it is clear that this perception is not universal, and the patterns correspond closely to some of the theories tested. Moreover, it could be that "simplicity" is a measure complementary to the concepts of compound lotteries and ambiguity.

[21] A nonparametric test that compares several paired groups. The Friedman test first ranks the valuation for each subject from low to high (separately). It then sums the ranks for each urn. If the sums are very different, the test will tend to reject the null hypothesis that the valuations of different urns came from the same distribution.

[22] A parametric test, such as the repeated measures analysis of variance (ANOVA), that allows for heterogeneity between subjects could not reject the null hypothesis that the mean valuations of urns 2, 3, and 4 are equal at the 10% significance level.

[23] However, the parametric repeated measures ANOVA cannot reject this latter hypothesis at a significance level of 10%.

TABLE II

DESCRIPTIVE STATISTICS

| | $V1$ | $V2$ | $V3$ | $V4$ | AVE($V1, \ldots, V4$) | STD($V1, \ldots, V4$) | Max($V1, \ldots, V4$) | Min($V1, \ldots, V4$) | $V21$ | $V43$ | $V41$ | $V31$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| First Round ($2 prize, 104 subjects) | | | | | | | | | | | | |
| Mean | 1.06 | 0.88 | 0.93 | 0.95 | 0.95 | 0.19 | 1.17 | 0.75 | −0.18 | 0.02 | −0.11 | −0.13 |
| Median | 1 | 0.99 | 0.99 | 1 | 0.99 | 0.16 | 1.05 | 0.8 | −0.1 | 0 | 0 | −0.03 |
| Mode | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| SD | 0.34 | 0.32 | 0.30 | 0.38 | 0.26 | 0.18 | 0.32 | 0.32 | 0.33 | 0.42 | 0.34 | 0.31 |
| Minimum | 0.1 | 0 | 0.13 | 0.06 | 0.2 | 0 | 0.4 | 0 | −1.26 | −1.39 | −1 | −1 |
| Maximum | 2 | 1.83 | 2 | 2 | 1.73 | 0.74 | 2 | 1.5 | 1.2 | 1.4 | 1.6 | 1.02 |
| Robustness Round ($20 prize, 38 subjects) | | | | | | | | | | | | |
| Mean | 8.37 | 6.66 | 7.25 | 7.74 | 7.51 | 2.00 | 9.89 | 5.62 | −1.70 | 0.50 | −0.63 | −1.12 |
| Median | 10 | 6 | 6 | 9 | 7.5 | 1.95 | 10 | 5 | −2 | 0 | 0 | −1 |
| Mode | 10 | 5 | 5 | 10 | 7.5 | 0 | 10 | 5 | −5[a] | 0 | 0 | 0 |
| SD | 3.02 | 3.00 | 3.13 | 3.54 | 2.51 | 1.24 | 2.89 | 2.69 | 3.08 | 3.56 | 3.42 | 3.47 |
| Minimum | 2 | 1 | 2 | 2 | 2.75 | 0 | 4 | 1 | −7 | −8 | −10 | −10 |
| Maximum | 15 | 13 | 15 | 17 | 12 | 4.72 | 17 | 10 | 8 | 9 | 8 | 7 |

[a]Multiple modes exist. The smallest value is shown.

significantly higher than the rest of the urns, and even FOSD the reservation prices for urn 2.

The only important difference between the first round and the robustness round is the fact that the average reservation price for urn 1 is lower than the expected value of the lottery—implying risk aversion. This change reflects not only increase in risk aversion as the stakes have increased (Holt and Laury (2002)), but the more careful design of the experiment: the sampling was more efficient (attracting fewer individuals who just wanted to earn the $10 or enjoyed the gambling aspect of the experiment); there was no framing in terms of "selling price" (that might cause a subject to state a higher reservation price)[24]; the subjects were repeatedly prompted to find their *minimal* reservation price; and the operation of the BDM elicitation mechanism was better demonstrated. The effect was uniform across the urns (the average reservation price for all of the urns increased by about eightfold); hence, the effect on the relative attractiveness of the urns was minimal. To sum up, even if $V1 > \mathbb{E}(L_1)$ (a frequent observation in the first round), it does not necessarily imply the subject is risk seeking.[25] Moreover, the factors that may have inflated $V1$ in the first round had similar effects on the other urns; hence, the results that follow, which focus on the differences in the elicited values, continue to hold.[26]

The focus of the analysis in this work is to identify (possibly heterogeneous) patterns of choice in the subjects' population. Table III shows the relatively high positive correlation between the reservation prices set for the four urns in both rounds. This positive correlation could be related to Ariely, Loewenstein, and Prelec's (2003) "coherent arbitrariness": a subject may find it difficult to evaluate each urn separately, but easier to compare two or more lotteries. Fox and Tversky (1995) compared valuations of risky (urn 1) and ambiguous (urn 2) lotteries, and found that when the subjects were not comparing the lotteries, the valuations where not significantly different. However, as argued previously, this may be exactly the reflection of the "arbitrariness." Ellsberg type behavior exists especially when the decision maker compares an ambiguous lottery to a risky one. The environment in the current study is comparative and is enriched by the existence of objective compound lotteries.

The difference between the prices set by the subjects for urn 2 (ambiguous) and urn 1 (one stage risky)—the (negative of) ambiguity premium—is used as

[24]The endowment effect may be responsible in part for the high reservation prices in the first round, although a recent study by Plott and Zeiler (2005) found it to be insignificant, when sufficient controls were introduced (which may be the case in the second round).

[25]Furthermore, Keller, Segal, and Wang (1993) showed that even theoretically, when subject's preferences do not satisfy ROCL, the true certainty equivalent and the elicited value may lie on opposite sides of the expected value.

[26]The random order treatment implemented in the robustness round showed that the higher reservation price for urn 1 in the original sample was not a consequence of this urn being the first task the subject confronted in the original round. Details of the order treatment are reported in Appendix A.

TABLE III

CORRELATION MATRICES FOR RERESERVATION PRICES

| | $V1$ | $V2$ | $V3$ | $V4$ |
|---|---|---|---|---|
| | | First Round | | |
| $V1$ | 1 | | | |
| $V2$ | 0.5011 | 1 | | |
| $V3$ | 0.54026 | 0.449509 | 1 | |
| $V4$ | 0.557401 | 0.369145 | 0.266104 | 1 |
| | | Robustness Round | | |
| $V1$ | 1 | | | |
| $V2$ | 0.4787 | 1 | | |
| $V3$ | 0.3637 | 0.711 | 1 | |
| $V4$ | 0.4645 | 0.5548 | 0.4353 | 1 |

a measure of ambiguity aversion. Therefore, if this variable is negative (positive), it implies ambiguity aversion (seeking). Similarly, the difference in the reservation prices set for urns 1, 3, and 4 (separately) measures the subject's attitude to (objective) second-order risk. Because all of $(V2-V1)$, $(V3-V1)$, and $(V4-V1)$ are defined relative to $V1$, they will always be positively correlated. This observation does not apply to $(V4-V3)$, and, therefore, this will be the main variable on which the test of theories that do not abide by ROCL will be based.

### 4.1. *Ambiguity Neutrality and Reduction*

One of the main characteristics of the population of subjects is the strong association between ambiguity neutrality ($V1 = V2$) and reduction of compound objective lotteries ($V1 = V3 = V4$). This behavior is evident in both rounds, as described in Table IV.

In the original sample, 18 subjects set $V1 = V3 = V4$, and more than 94% of them (17 subjects) asked for no ambiguity premium (set $V2 = V1$). This is more than four times the expected frequency under a null hypothesis of independence. Out of the 86 subjects who did not abide by ROCL, only 6 were ambiguity neutral (less than one-third of the expected frequency under the null hypothesis of independence). In the scaled sample: 5 subjects set $V1 = V3 = V4$ and *all* 5 of them set $V2 = V1$ (more than seven times the expected frequency under independence). Out of 33 subjects who did not abide by ROCL, *none* was ambiguity neutral (compared to expected frequency of 4.34 under independence). In the first round, there was a group of 13 subjects who set prices of \$1 for all four urns, and in the second round this group (reservation prices of \$10) consisted of four subjects. These subjects are responsible for a substantial part of the association. We can only speculate what these subjects' preferences are: it could be that \$1 (\$10 in the second round) is

TABLE IV

THE ASSOCIATION BETWEEN AMBIGUITY NEUTRALITY AND ROCL

| $V1 = V2$ | | $V1 = V3 = V4$ No | $V1 = V3 = V4$ Yes | Total |
|---|---|---|---|---|
| | | First Round* | | |
| No | Count | 80 | 1 | 81 |
| | Expected | 66.98 | 14.02 | |
| Yes | Count | 6 | 17 | 23 |
| | Expected | 19.02 | 3.98 | |
| Total | | 86 | 18 | 104 |
| | | Robustness Round** | | |
| No | Count | 33 | 0 | 33 |
| | Expected | 28.66 | 4.34 | |
| Yes | Count | 0 | 5 | 5 |
| | Expected | 4.34 | 0.66 | |
| Total | | 33 | 5 | 38 |

*Fisher's exact test: Exact sig. (2-sided) 1.24E−13.

**Fisher's exact test: Exact sig. (2-sided) 1.99E−06.

a focal point. Alternatively, it could be that these subjects are expected value maximizers. Some indication is given by the fact that taking at least one advanced (second year or higher) mathematics course increased the probability of choosing $(1, 1, 1, 1)$ from 10% to 21% in the first round. Similarly, the conditional probability in the second round of being an expected value maximizer increased from 6% to 40% when controlling for taking an advanced mathematics course. Note, however, that even after removing these subjects, the strong association between ROCL and ambiguity neutrality is retained ($p < 0.001$ in the first sample and $p < 0.03$ in the second sample).

The conclusion derived from Table IV is that there is a very tight association between ambiguity neutrality $\{V2 = V1\}$ and reduction of compound lotteries $\{V1 = V3 = V4\}$. Therefore, a descriptive theory that accounts for ambiguity aversion should account—at the same time—for violation of reduction of compound *objective* lotteries.

### 4.2. *Attitude Toward Mean Preserving Spreads in Probabilities and Ambiguity*

As discussed in Section 3, alternative theories that can account for a non-neutral attitude toward ambiguity by relaxing ROCL in objective probabilities (recursive nonexpected utility and bundling/possible interpretation of recursive expected utility) have different predictions on the relative attractiveness of urns 1, 3, and 4. Furthermore, their predictions on the sign of the correlation between ambiguity premium (as measured by $V21 = V2 - V1$) and the premium to dispersion in the second-order distribution (as measured by $V43 = V4 - V3$) differ.

If one looks at the "average" subject, who does not satisfy reduction of compound lotteries,[27] it seems that this latter correlation is very weak.[28] However, the data clearly exhibit different patterns of reservation prices for urns 1, 3 and 4 that correspond to the two alternative models. Therefore, the absence of significant correlation between ambiguity premium ($V21$) and the premium to dispersion in the second-order objective probability ($V43$) may be a result of averaging the two sets of subjects (that conform to bundling/REU and RNEU), which exhibit approximately opposite correlations between the two variables. To test whether the data are consistent with this interpretation, it is necessary to first classify the subjects; second, to test whether the classification is internally consistent;[29] third, to test whether ambiguity aversion, as measured by the ambiguity premium, could be accounted for using these theories. The objectivity of this methodology relies crucially on the fact that the first two stages do not use any information that contains $V2$, but rely solely on subjects' preferences over mean preserving spreads of the objective second-order distribution (that is, urns 1, 3, and 4 only). The third step is a test of whether the classification produces the correlations between $V21$ and $V43$ predicted by the theories.

Ranking of urns 1, 3, and 4 may exhibit 13 possible ordinal ranking schemes. The classification is based on the following criteria:

- If $V1 = V3 = V4$, the subject reduces compound objective lotteries and therefore is consistent with SEU (Savage (1954)) or MEU (Gilboa and Schmeidler (1989))/CEU (Schmeidler (1989)) in the Anscombe and Aumann (1963) framework.
- If $V1 \geq V3 \geq V4$ or $V1 \leq V3 \leq V4$ (where at least one of the inequalities is strict), the preferences are consistent with the bundling rationale (Halevy and Feltkamp (2005)) or the REU model (KMM (2005), Ergin and Gul (2004), Nau (2006), Ahn (2003), Seo (2006)) when the decision maker does not reduce compound *objective* lotteries.
- If $V1 = V4 \neq V3$, then the subject is consistent with the RNEU model (Segal (1987)).

This classification leaves four ordinal rankings[30] that are not consistent with the foregoing theories. Acknowledging possible noise/error/randomness in as-

[27]That is, the subset of 83 subjects in the original experiment and 33 subjects in the robustness experiment who do not conform to SEU or MEU (3 subjects who set reservation prices within 2¢ of the expected value predictions are included in the set of 21 subjects who satisfy the ROCL in the original sample).

[28]In the first sample, the Pearson correlation is $-0.1$ ($p = 0.35$), and Spearman's $\rho$ is $-0.07$ and insignificantly different from zero. Similar "average" behavior is exhibited in the scaled sample: the Pearson correlation is $-0.2$ ($p = 0.24$) and the Spearman $\rho$ is $-0.23$ ($p = 0.18$).

[29]Especially in the case of RNEU (Segal (1987)), does $\mathbb{E}(V4|V1) = V1$ hold?

[30]Including 36 and 11 subjects in the original and scaled samples, respectively, that have one of the following patterns of reservation prices: $V3 < V4 < V1$, $V1 < V4 < V3$, $V3 < V1 < V4$, and $V4 < V1 < V3$.

signing reservation prices allows us to classify these subjects. It captures human error in assigning reservation prices using the graphical interface of the experiment or may result from the use of the BDM mechanism, from difficulties understanding the mechanism, or from other sources. A possible avenue to model this randomness may be by providing a random utility model as in Gul and Pesendorfer (2006),[31] who provided an axiomatic foundation that allows a representation of choices by a random expected utility function. Here, however, the space of one-stage lotteries is replaced with compound lotteries, and the axioms (especially linearity, which is comparable to the standard independence axiom) have to be modified. This is an important and challenging task, which is beyond the scope of the current paper. An intermediate solution can be adopted from the logistic choice literature (see Anderson, Goeree, and Holt (2005)). For an urn $i$ ($L_i$) with possible prize of \$$x$, let the element of choice be the reservation price—a number in the interval $[0, x]$. For every $t \in [0, x]$, let $U_k(L_i, t)$ be the utility of the decision maker whose preference corresponds to theory $k$ (one of those discussed in Section 3) of choosing reservation price $t$ for urn $i = 1, \dots, 4$. Unlike in Gul and Pesendorfer (2006), this approach assumes a possible error in comparing the utility of different reservation prices, that influences the probability of choosing the number (reservation price) with the highest utility. The density of choosing a reservation price $t$ for urn $i$ if the decision maker's preferences are described by theory $k$ is

$$f_k(L_i, t) = \frac{\exp(U_k(L_i, t)/\xi)}{\int_0^x \exp(U_k(L_i, s)/\xi)\, ds},$$

where $\xi$ is an error parameter that determines the importance of an error term (which in this case is logistic). A small $\xi$ implies that the choices of reservation prices are close to the one predicted by the respective theory, $k$. For example, low $\xi$ for a decision maker described by the RNEU model (Segal (1987)) implies that $V1$ would be relatively close to $V4$, while if the decision maker's preferences originate in the bundling rationale (or are described by the REU model without ROCL), the difference between the two would be relatively large. The goal in classifying the remaining subjects is to choose, for every subject, the theory that is consistent with the lowest $\xi$. To better understand how allowing for noise allows us to classify the subjects, consider an observation of $V3 < V1 < V4$: it may belong to a decision maker described by Segal (1987), who, without error, had a ranking of $V3 < V1 = V4$, or it may belong to an agent described by the bundling rationale (whose preferences may be

[31]The decision maker's behavior of choosing from finite menus is described by a random choice rule, which assigns to each possible menu a probability distribution over feasible choices. A random utility function is a probability measure on some set of utility functions. The random choice rule maximizes the random utility function if, for every possible menu, the random choice rule coincides with the probability that the random utility function attains its maximum on the corresponding alternatives.

represented by the REU) whose "before noise" ranking is $V1 < V3 < V4$. To distinguish between the alternative explanations, attention is focused on the relative reservation prices of urns 1 and 4 for low values of $\xi$: under Segal's theory a decision maker is indifferent between urns 1 and 4, hence, we would expect a small cardinal difference between the valuations of the two urns. Under the bundling rationale and the REU, the cardinal difference between the evaluations of urns 1 and 4 should be relatively larger. The size of the cardinal differentiation is measured by the absolute difference between $V1$ and $V4$, normalized by the standard deviation of the reservation prices determined in urns 1, 3, and 4.

Following the foregoing principle, subjects with the patterns $V3 < V4 < V1$, $V1 < V4 < V3$, and $V3 < V1 < V4$, who uniformly exhibited (in the original sample) low values of $|V41|/\mathrm{STD}_{V1,V3,V4}$, are classified as recursive nonexpected utility (Segal (1987)). The group of subjects with the pattern $V4 < V1 < V3$ exhibited higher variability with respect to this measure: therefore, three of them are classified as RNEU while the rest are classified as bundling (Halevy and Feltkamp (2005))/REU (KMM (2005), Nau (2006), Ergin and Gul (2004), Ahn (2003), Seo (2006)) subjects.[32] It is important to note that the classification method used does not employ any information on the reservation prices for the ambiguous (second) urn. The partition is reported in Table V.

The partition results in 21 (5 in the scaled sample) subjects who reduced compound lotteries, 20 (all 5) of whom were consistent with the SEU model,[33] 1 (no) subject who corresponded to the MEU (in the Anscombe–Aumann framework that assumes ROCL) predictions, 41 (17) subjects who corresponded the RNEU model (Segal (1987)), and 42 (16) subjects who corresponded to the bundling rationale to ambiguity aversion (Halevy and Feltkamp (2005)) and to the interpretation of the REU model (KMM (2005), Ergin and Gul (2004), Nau (2006), Ahn (2003), Seo (2006)) in which the decision maker does not reduce compound objective lotteries.

### 4.2.1. *Cardinal analysis*

The current section tests whether the classification method presented in Table V is internally consistent, and whether it accounts for ambiguity attitudes in a way that is consistent with the theories.

[32]Similar partition was performed in the scaled sample: eight subjects were classified as RNEU ($U3 < U4 < U1$, $U1 < U4 < U3$, and two observations with $U4 < U1 < U3$) and three subjects were classified as bundling/REU (a single observation with $U3 < U1 < U4$ and two observations with $U4 < U1 < U3$).

[33]This includes 13 subjects who behaved as expected value maximizers, 3 subjects within 2¢ of the theoretical prediction of the expected value model, 2 subjects who set the four reservation prices higher than \$1 (\$1.3 and \$1.5), and 2 subjects who set the four reservation prices lower than \$1 (\$0.8 and \$0.99). In the scaled sample, 4 subjects were expected value maximizers and 1 subject set all reservation prices to \$5.

TABLE V

CLASSIFICATION SUMMARY

| | $2 Sample | | | | | | $20 Sample | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Theory | # of Obs | $V1$ | $V2$ | $V3$ | $V4$ | corr($V21, V43$) | # of Obs | $V1$ | $V2$ | $V3$ | $V4$ | corr($V21, V43$) |
| Rule/REU | 42 | 1.04 | 0.78 | 0.96 | 0.79 | 0.447 | 16 | 7.66 | 5.34 | 6.36 | 6.47 | 0.592 |
| Consistent[a] | **31** | **1.01** | **0.84** | **0.90** | **0.73** | **0.644** | **12** | **7.88** | **5.33** | **5.83** | **5.50** | **0.571** |
| Optimist[b] | 1 | 1.10 | 1.20 | 1.00 | 1.00 | | 0 | | | | | |
| MREU[c] | 7 | 1.19 | 0.62 | 1.23 | 0.89 | | 1 | 13 | 9.5 | 14 | 10 | |
| Inconsistent[d] | 3 | 0.89 | 0.39 | 0.92 | 1.14 | | 3 | 5 | 4 | 5.92 | 9.17 | |
| RNEU | 41 | 1.11 | 0.91 | 0.85 | 1.07 | −0.620 | 17 | 8.85 | 7.22 | 7.57 | 8.57 | −0.771 |
| Consistent[a] | **32** | **1.17** | **0.86** | **0.85** | **1.12** | **−0.735** | **15** | **9.37** | **7.12** | **7.58** | **8.98** | **−0.886** |
| Optimist | 7 (for 6[b]) | 0.89 | 1.17 | 0.75 | 0.87 | | 1 | 2 | 10 | 5 | 3 | |
| Inconsistent[e] | 2 | 0.92 | 0.77 | 1.27 | 0.92 | | 1 | 8 | 6 | 10 | 8 | |
| SEU | **20** | **1.03** | **1.03** | **1.03** | **1.03** | | **5** | 9 | 9 | 9 | 9 | |
| Expected value | 13 | 1 | 1 | 1 | 1 | | 4 | 10 | 10 | 10 | 10 | |
| EV with noise[f] | 3 | 0.99 | 1.00 | 0.99 | 0.99 | | 0 | | | | | |
| Risk averse[g] | 2 | 0.90 | 0.90 | 0.90 | 0.90 | | 1 | 5 | 5 | 5 | 5 | |
| Risk seeking[h] | 2 | 1.40 | 1.40 | 1.40 | 1.40 | | 0 | | | | | |
| MEU-AA[i] | **1** | **0.90** | **0.80** | **0.90** | **0.90** | | **0** | | | | | |
| *Total* | 104 | 1.06 | 0.88 | 0.93 | 0.95 | −0.104 | 38 | 8.37 | 6.66 | 7.25 | 7.74 | −0.211 |

[a] Includes subjects averse or seeking MPS in the second-order objective distribution.

[b] Averse to MPS in the second-order distribution but $V2 > V1$.

[c] Maxmin REU: $V2 < V4 < V1$.

[d] REU inconsistent: $V2 < V1 < V3 < V4$.

[e] RNEU inconsistent: $V2 < V1 = V4 < V3$.

[f] Within 1–2¢ of the E$V$ predictions.

[g] $Vi < \mathrm{E}V$ for $i = 1, \ldots, 4$.

[h] $Vi > \mathrm{E}V$ for $i = 1, \ldots, 4$.

[i] $V2 < V1 = V3 = V4$.

*Recursive Nonexpected Utility*: The average reservation prices for urns 1 and 4 in this subsample of subjects are very close (\$1.11 and \$1.07, respectively, in the original sample, and \$8.85 and \$8.57 in the scaled sample). Nonparametric (Friedman) and parametric (repeated measures ANOVA) tests cannot reject the null hypothesis that the two series came from the same underlying distribution.[34] However, the theory imposes a stricter restriction on the series: differences between $V1$ and $V4$ should be due only to nonsystematic noise/error. In other words, the expected value of $V4$ conditional on $V1$ should be $V1$. The model estimated is

$$V4 = \alpha + \beta V1 + \varepsilon$$

and the null composite hypothesis tested is, therefore, $\alpha = 0$ and $\beta = 1$. The value of the $F(2, 39)$ statistic is 1.02 ($p = 0.37$); hence, the hypothesis cannot be rejected and this group is internally consistent with the theoretical predictions of the RNEU model (Segal (1987)). The group of RNEU subjects in the robustness sample exhibits similar consistency ($F(2, 15) = 0.83,\ p = 0.45$).

Once internal consistency of the RNEU group is established, the focus shifts to the predictive power of the theory on ambiguity aversion. Out of 41 subjects in the first sample who were classified to this group based on their reservation prices for the three objective urns (see Table V), two subjects are inconsistent with the theoretical predictions of the model concerning ambiguity aversion, because their reservation prices satisfy $V2 < V1 = V4 < V3$. That is, they like mean preserving spreads in the objective second-order (first-stage) distribution, but are ambiguity averse. In the robustness sample, only one subject exhibited this inconsistency. Seven more subjects in the original sample hold "optimistic" beliefs over the composition of the second urn, while six of them exhibited $V3 < \min\{V1, V4\} \leq \max\{V1, V4\} \leq V2$.[35] Although those subjects are consistent with the RNEU model, because the theory does not impose restrictions on the decision maker's beliefs over the ambiguous urn (the prior may not be symmetric around 5 red and 5 black balls), they were removed from the quantitative analysis. Some support for the view that this pattern of reservation price is due to "mistake" or "carelessness" may be found in the fact that when the payoffs were scaled up by a factor of 10, this pattern of optimistic choice almost disappeared.

The prediction of the RNEU model—that the ambiguity premium ($-V21$) and the premium for mean preserving spread in the second-order distribution ($-V43$) are negatively correlated—was tested on the remaining 32 subjects. The correlation between these two sequences is $-0.735$ and is significantly different from zero ($p = 1.62 \times 10^{-6}$), conforming to the theory's explanation of

[34] The value of the Friedman test is 0.34 in the first sample (0.16 in the second sample) and of the repeated measures ANOVA is 0.27 in the first sample (0.31 in the robustness sample).

[35] One more subject (with $V3 > \max\{V1, V4\}$) set a reservation price for the ambiguous urn that was seven times higher than urn 1.

ambiguity aversion within this subset of subjects. In the robustness test this correlation is $-0.886$ and is significantly different from zero ($p < 0.00025$). Section B.1 further quantifies the relationship between ambiguity premium and the premium to mean preserving spread in the second-order probability within this population via a simple regression of $V21$ on $V31$ or $V43$. The results are consistent with the theoretical predictions of the RNEU model (in both the original and the scaled sample): there is a strong association between subjects' attitude to urn 3 (relative to urns 1 or 4) and their ambiguity premium (urn 2 relative to urns 1 or 4).

*Bundling* (*Rule Rationality*)/*Recursive Expected Utility*: The bundling rationale (Halevy and Feltkamp (2005)) and the interpretation of the REU model (KMM (2005), Ergin and Gulb (2004), Nau (2006), Ahn (2003), Seo (2006)), which is consistent with violation of reduction in objective compound lotteries, impose several restrictions (9) on the ordinal ranking of the urns. Out of the 42 subjects in the \$2 sample that belong to this group (based on their preferences over two-stage objective lotteries), 10 subjects do not satisfy these restrictions (see Table V). Three subjects exhibit the pattern of reservation prices $V2 < V1 \leq V3 \leq V4$ (where one of the two right inequalities is strict): that is, although they seem to set monotonically higher reservation prices for urns with higher dispersion in the second-order objective distribution, they dislike ambiguity. Seven other subjects set $V2 < V4 < V1$, violating the restriction that under the bundling rationale/REU, if a subject is averse to mean preserving spread in the second-order distribution, the valuation of the ambiguous urn is bounded from below by the valuation of urn 4. These subjects' valuation of the ambiguous urn has the "pessimistic" flavor of the MEU model or, alternatively their preferences are based on other criteria such as "simplicity." One additional subject in the \$2 sample had an "optimistic" valuation of urn 2: $V2 > \max\{V1, V4\}$. As noted for the RNEU model, this pattern does not contradict (9), because the decision maker may hold beliefs over the composition of the ambiguous urn that are not symmetric around 5 red and 5 black balls.

This leaves 31 subjects (out of the 42 subjects) with a valuation of the ambiguous urn that is consistent with the theoretical predictions of the bundling/REU model. In the robustness sample, 4 out of 16 subjects are not consistent with the theoretical predictions of the bundling/REU model (see Table V). Among the remaining 31 and 12 subjects (in the \$2 and \$20 samples, respectively), the correlation between $-V43$ (the premium to mean preserving spread in the second-order distribution) and $-V21$ (the ambiguity premium) is positive and significantly different from zero. The Pearson correlations are 0.644 and 0.571, respectively, and are statistically different from zero ($p < 0.0001$ and $p = 0.053$, respectively).[36] Section B.2 further quantifies the

[36]The rank-based Spearman $\rho$ is 0.62 for the original sample and 0.514 for the scaled sample ($p = 0.0002$ and $p = 0.08$, respectively).

relationship between reduction of compound objective probabilities and ambiguity aversion, using linear regression of $V21$ on $V31$ and $V43$ (or their sum, $V41$). The ambiguity premium is a positive function of the premium to mean preserving spread in the second-order probability: from urn 1 to urn 3, and from urn 3 to urn 4. It is important to note that even when controlling for the former ($V31$), the latter's ($V43$) effect on the ambiguity premium ($V21$) is positive and significant ($p < 0.0001$ in both samples). The regression results for this group indicate that, similarly to the subjects who belong to the RNEU group, these subjects identify the ambiguous urn with an urn that has nondegenerate second-order distribution. They assign *subjective* second-order belief over the composition of the second urn that could be anything between the first urn (degenerate second-order belief) and the fourth urn (extreme second-order belief). Their ambiguity preferences, therefore, are associated with their preferences over compound lotteries like urns 3 and 4.

## 5. CONCLUSIONS

The experimental design used in this study enables examination the relationship between individuals' ambiguity attitudes and their attitudes toward compound objective lotteries. This design facilitates a clear empirical test of theories that model ambiguity, while assuming reduction of compound objective lotteries (Schmeidler (1989), Gilboa and Schmeidler (1989)). Furthermore, theories that model ambiguity aversion as a phenomenon that is explicitly associated with a violation of reduction of compound lotteries (Segal (1987), Halevy and Feltkamp (2005), Seo (2006)) or that are open for such an interpretation (KMM (2005), Ergin and Gul (2004), Nau (2006), Ahn (2003)) are evaluated empirically based on their predicted pattern of preference among objective lotteries with a varying amount of dispersion in the second-order distribution.

The results reveal a tight association between ambiguity neutrality and reduction of compound *objective* lotteries that is consistent with the SEU model: subjects who reduced compound lotteries were almost always ambiguity neutral, and most subjects who were ambiguity neutral reduced compound lotteries appropriately (15–20% of the subjects). The remainder of the subjects exhibited violations of ROCL and ambiguity aversion, but there is no unique theory that can accommodate the different choice patterns in the population. The population is heterogeneous and two choice patterns, which account for approximately 70% of all subjects, emerge. In particular, about half (35%) exhibit ambiguity aversion (seeking) together with aversion (affinity) to mean preserving spreads in the second-order distribution. These preferences can be traced back to a "rational rule," which originates in an environment of choice among bundles of lotteries (Halevy and Feltkamp (2005)), and are consistent with an interpretation of the REU model (KMM (2005), Ergin and Gul (2004), Nau (2006), Ahn (2003), Seo (2006)) that allows a decision maker not to reduce (or differentiate) different sources of objective risk. The other half (35%) of the subjects exhibit a pattern of preferences consistent with Segal's (1987, 1990) theory of RNEU, where the decision maker evaluates two-stage lotteries (including ambiguous lotteries) using, recursively, rank dependent utility.

The findings indicate that currently there is no unique theoretical model that universally captures ambiguity preferences. In this sense, the current work confirms Epstein's (1999) approach of defining ambiguity aversion as a behavior that is not probabilistically sophisticated, without committing to a specific functional model. The results suggest that failure to reduce compound (objective) lotteries is the underlying factor of the Ellsberg paradox, and call upon decision theory to uncover the theoretical relationship between ambiguity aversion and different forms in which reduction may fail.

*Dept. of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, BC V6T 1Z1, Canada; yhalevy@interchange.ubc.ca; http://www.econ.ubc.ca/halevy.*

*Manuscript received November, 2004; final revision received September, 2006.*

## APPENDIX A: ORDER TREATMENT

The random-order treatment in the robustness round examines whether the higher reservation price for urn 1 in the original sample is a consequence of it being a simple one-stage objective lottery or a consequence of urn 1 being the first task the subject confronted in the original experiment. The subjects were randomly treated with alternative orders of urns: $(1, 2, 3, 4)$, $(2, 3, 4, 1)$, $(3, 4, 1, 2)$, and $(4, 1, 2, 3)$. The only significant order effect found in the sample is that the first task urn received a significantly *lower* reservation price than under alternative orders in which this urn was not the first (Friedman test value of $\chi^2(3, N = 38) = 8.8,\ p = 0.032$). This order effect seems to operate in a direction opposite to that proposed Harrison, Johnson, McInnes, and Rutström (2005), who found that people are more risk averse (lower reservation price) in late tasks. Table VI reports the average reservation price for each urn as a function of its order. For example, $(V1, 3rd)$, $(V2, 4th)$, $(V3, 1st)$,

TABLE VI

VARIATION IN THE RESERVATION PRICES AS A FUNCTION OF ORDER OF URNS

| | 1st | 2nd | 3rd | 4th |
|---|---|---|---|---|
| $V1$ | 7.28 | 8.90 | 7.35 | 10.00 |
| $V2$ | 6.44 | 6.67 | 7.65 | 5.88 |
| $V3$ | 5.55 | 6.77 | 8.56 | 8.20 |
| $V4$ | 6.70 | 8.10 | 8.25 | 8.00 |
| Average | **6.49** | **7.61** | **7.95** | **8.02** |

and $(V4, 2nd)$ correspond to the average reservation prices of the 10 subjects who were treated with the order $(3, 4, 1, 2)$. The conclusion from the order treatment is that the significantly higher reservation price for urn 1 in the original experiment (in which setting the reservation price for urn 1 was always the first task) cannot be attributed to an order effect (indeed, it persisted in the robustness test).

## APPENDIX B: REGRESSION RESULTS

### B.1. *Recursive Nonexpected Utility* (*Segal*)

Two alternative variables can be used: $V31$ or $V43$ (using both will result in multicollinearity). Both options are presented in Table VII.[37] The estimated alternative (given that $\mathbb{E}(V4|V1) = V1$) models are

$$V21 = \alpha_{31} + \beta_{31}V31 + \varepsilon,$$
$$V21 = \alpha_{43} + \beta_{43}V43 + \varepsilon'.$$

TABLE VII

THE AMBIGUITY PREMIUM AS A FUNCTION OF ALTERNATIVE ESTIMATES OF AVERSION TO MPS FOR THE RNEU (SEGAL) GROUP

| | \$2 Sample | | \$20 Sample | |
|---|---|---|---|---|
| Statistics | $V31$ | $V43$ | $V31$ | $V43$ |
| Multiple $R$ | 0.788 | 0.735 | 0.948 | 0.886 |
| $R$ squared | 0.621 | 0.541 | 0.899 | 0.785 |
| Adjusted $R$ squared | 0.608 | 0.526 | 0.882 | 0.768 |
| Standard error | 0.197 | 0.216 | 1.032 | 1.444 |
| Observations | 32 | 32 | 15 | 15 |

| Sample | | Coef | SE | $t$ Stat | Value | Lower 95% | Upper 95% |
|---|---|---|---|---|---|---|---|
| \$2 | Intercept | −0.110 | 0.045 | −2.451 | 0.020 | −0.202 | −0.018 |
| | $V31$ | 0.620 | 0.089 | 7.004 | 8.78E−08 | 0.439 | 0.801 |
| \$20 | Intercept | −0.925 | 0.311 | −2.970 | 0.011 | −1.597 | −0.252 |
| | $V31$ | 0.740 | 0.076 | 9.705 | 2.54E−07 | 0.576 | 0.905 |
| \$2 | Intercept | −0.170 | 0.045 | −3.806 | 0.001 | −0.262 | −0.079 |
| | $V43$ | −0.502 | 0.085 | −5.946 | 1.62E−06 | −0.675 | −0.330 |
| \$20 | Intercept | −1.295 | 0.398 | −3.258 | 0.006 | −2.155 | −0.436 |
| | $V43$ | −0.679 | 0.099 | −6.885 | 1.11E−05 | −0.892 | −0.466 |

[37]A test of whether $V1$ has a significant effect beyond $V31$ or $V43$ reveals that it is insignificant at 5%. That is, the ambiguity premium could be explained by agents' attitudes to mean preserving spread in the second-order distribution.

TABLE VIII

AMBIGUITY PREMIUM AS A FUNCTION OF ALTERNATIVE ESTIMATES OF AVERSION TO MPS FOR THE BUNDLING/REU GROUP

| | $2 Sample | | $20 Sample | |
|---|---|---|---|---|
| Statistics | $V31$, $V43$ | $V41$ | $V31$, $V43$ | $V41$ |
| Multiple $R$ | 0.806 | 0.803 | 0.951 | 0.951 |
| $R$ square | 0.650 | 0.644 | 0.905 | 0.904 |
| Adjusted $R$ square | 0.625 | 0.632 | 0.883 | 0.894 |
| Standard error | 0.151 | 0.150 | 1.018 | 0.970 |
| Observations | 31 | 31 | 12 | 12 |

| Sample | | Coef | SE | $t$ Stat | $P$-Value | Lower 95% | Upper 95% |
|---|---|---|---|---|---|---|---|
| $2 | Intercept | −0.057 | 0.032 | −1.821 | 0.079 | −0.122 | 0.007 |
| | $V31$ | 0.454 | 0.105 | 4.341 | 0.0002 | 0.240 | 0.668 |
| | $V43$ | 0.372 | 0.063 | 5.888 | 2.48E–06 | 0.243 | 0.501 |
| $20 | Intercept | −1.178 | 0.338 | −3.484 | 0.007 | −1.942 | −0.413 |
| | $V31$ | 0.579 | 0.078 | 7.388 | 0.00004 | 0.401 | 0.756 |
| | $V43$ | 0.549 | 0.080 | 6.833 | 7.61E–05 | 0.367 | 0.730 |
| $2 | Intercept | −0.060 | 0.031 | −1.950 | 0.061 | −0.124 | 0.003 |
| | $V41$ | 0.393 | 0.054 | 7.249 | 5.55E–08 | 0.282 | 0.504 |
| $20 | Intercept | −1.202 | 0.312 | −3.847 | 0.003 | −1.898 | −0.506 |
| | $V41$ | 0.564 | 0.058 | 9.683 | 2.13E–06 | 0.434 | 0.694 |

### B.2. *Bundling* (*Rule Rationality*)/*Recursive Expected Utility*

Two alternative and equivalent formulations are possible. One possibility is that the variables on the right-hand side are $V31$ and $V43$, which measure the subjects' aversion to mean preserving spreads in the second-order distribution moving from urn 1 to urn 3 and from the latter to urn 4, respectively. An alternative method is to place the sum of the two—$V41$—on the right-hand side of the regression equation.[38] The two alternative models estimated are, therefore,

$$V21 = \theta + \gamma_{41} V41 + \varepsilon,$$

$$V21 = \theta' + \gamma_{31} V31 + \gamma_{43} V43 + \varepsilon'.$$

Table VIII summarizes the results of estimating these models.

## REFERENCES

AHN, D. S. (2003): "Ambiguity Without a State Space," available at http://emlab.berkeley.edu/~dahn/AHN_ambig.pdf. [504,518,525,527,530,531]

[38]The effect of $V1$ on ambiguity premium (beyond its effect on $V31$ and $V41$) is insignificant at 10%.

ANDERSON, S. P., J. K. GOEREE, AND C. A. HOLT (2005): "Logit Equilibrium Models of Anomalous Behavior: What to Do When the Nash Equilibrium Says One Thing and the Data Say Something Else," in *Handbook of Experimental Economic Results*, ed. by C. Plott and V. Smith. New York: Elsevier, forthcoming. [526]

ANSCOMBE, F. J., AND R. J. AUMANN (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34, 199–205. [505,511,512,518,525]

ARIELY, D., G. LOEWENSTEIN, AND D. PRELEC (2003): " 'Coherent Arbitrariness': Stable Demand Curves Without Stable Preferences," *The Quarterly Journal of Economics*, 118, 73–105. [522]

BECKER, G. M., M. H. DEGROOT, AND J. MARSCHAK (1964): "Measuring Utility by a Single Response Sequential Method," *Behavioral Science*, 9, 226–232. [504,507]

CASADESUS-MASANELL, R., P. KLIBANOFF, AND E. OZDENOREN (2000): "Maxmin Expected Utility over Savage Acts with a Set of Priors," *Journal of Economic Theory*, 92, 35–65. [512]

CHEW, S. H., E. KARNI, AND Z. SAFRA (1987): "Risk Aversion in the Theory of Expected Utility with Rank-Dependent Probabilities," *Journal of Economic Theory*, 42, 370–381. [514]

CHEW, S. H., and J. S. SAGI (2003): "Small Worlds: Modeling Attitudes Towards Sources of Uncertainty," available at http://faculty.haas.berkeley.edu/sagi/060625Lambda.pdf. [518]

CHOW, C. C., AND R. SARIN (2002): "Known, Unknown and Unknowable Uncertainties," *Theory and Decision*, 52, 127–138. [510]

ELLSBERG, D. (1961): "Risk, Ambiguity and the Savage Axioms," *The Quarterly Journal of Economics*, 75, 643–669. [503]

EPSTEIN, L. G. (1999): "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66, 579–608. [504,532]

ERGIN, H., AND F. GUL (2004): "A Subjective Theory of Compound Lotteries," available at http://econ-www.mit.edu/faculty/download_pdf.php?id=898. [504,516-519,525,527,530,531]

FOX, C. R., AND A. TVERSKY (1995): "Ambiguity Aversion and Comparative Ignorance," *The Quarterly Journal of Economics*, 110, 585–603. [510,522]

GILBOA, I. (1987): "Expected Utility with Purely Subjective Non-Additive Probabilities," *Journal of Mathematical Economics*, 16, 65–88. [512]

GILBOA, I., AND D. SCHMEIDLER (1989): "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153. [503,511,512,525,531]

GRETHER, D. M., AND C. R. PLOTT (1979): "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review*, 69, 623–638. [507]

GUL, F., AND W. PESENDORFER (2006): "Random Expected Utility," *Econometrica*, 74, 121–146. [526]

HALEVY, Y. (2007): "Supplement to 'Ellsberg Revisited: An Experimental Study'—Instructions and Data," *Econometrica Supplementary Material*, 75, http://www.econometricsociety.org/ecta/supmat/5581data.pdf. [506]

HALEVY, Y., AND V. FELTKAMP (2005): "A Bayesian Approach to Uncertainty Aversion," *The Review of Economic Studies*, 72, 449–466. [504,505,516-518,525,527,530,531]

HARRISON, G. W., E. JOHNSON, M. M. MCINNES, AND E. RUTSTRÖM (2005): "Risk Aversion and Incentive Effects: Comment," *American Economic Review*, 95, 897–901. [509,532]

HOLT, C. A. (1986): "Preference Reversal and the Independence Axiom," *American Economic Review*, 76, 508–515. [507]

HOLT, C. A., AND S. K. LAURY (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92, 1644–1655. [509,522]

——— (2005): "Risk Aversion and Incentive Effects: New Data Without Order Effects," *American Economic Review*, 95, 902–912. [509]

KARNI, E., AND Z. SAFRA (1987): " 'Preference Reversal' and the Observability of Preferences by Experimental Methods," *Econometrica*, 55, 675–685. [507]

KELLER, L. R., U. SEGAL, AND T. WANG (1993): "The Becker–DeGroot–Marschak Mechanism and Generalized Utility Theories: Theoretical Predictions and Empirical Observations," *Theory and Decision*, 34, 83–97. [507,522]

KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73, 1849–1892. [504,514,515,518,519,525,527,530,531]

KREPS, D. M., AND E. L. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200. [517]

MACHINA, M. J., AND D. SCHMEIDLER (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, 745–780. [503,504]

MORRIS, S. (1997): "Risk, Uncertainty and Hidden Information," *Theory and Decision*, 42, 235–269. [506]

NAU, R. F. (2006): "Uncertainty Aversion with Second-Order Utilities and Probabilities," *Management Science*, 52, 136–145. [504,518,525,527,530,531]

PLOTT, C. R., AND K. ZEILER (2005): "The Willingness to Pay/Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions and Experimental Procedures for Eliciting Valuations," *American Economic Review*, 95, 530–545. [522]

QUIGGIN, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343. [504,512]

ROTHSCHILD, M., AND J. E. STIGLITZ (1970): "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 2, 225–243. [518]

SAVAGE, L. J. (1954): *The Foundations of Statistics*. New York: Wiley. Revised and enlarged edition (1972), New York: Dover. [503,511,525]

SCHMEIDLER, D. (1989): "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57, 571–587. [504,512,516,525,531]

SEGAL, U. (1987): "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 28, 175–202. [504,505,512,514,519,525-527,529,531,532]

——— (1988): "Does the Preference Reversal Phenomenon Necessarily Contradict the Independence Axiom," *American Economic Review*, 78, 233–236. [507]

——— (1990): "Two-Stage Lotteries Without the Reduction Axiom," *Econometrica*, 58, 349–377. [505,511-514,518,532]

SEO, K. (2006): "Ambiguity and Second Order Belief," available at http://troi.cc.rochester.edu/~kseo/. [504,518,519,525,527,530,531]

TVERSKY, A., AND P. WAKKER (1995): "Risk Attitudes and Decision Weights," *Econometrica*, 63, 1255–1280. [516]

YATES, F. J., AND L. G. ZUKOWSKI (1976): "Characterization of Ambiguity in Decision Making," *Behavioral Science*, 21, 19–25. [510]