"Don't Put All Your Eggs in One Basket!": An Experimental Study of False Diversification*

Kfir Eliaz — Brown University

Guillaume R. Fréchette — New York University

November 2008

Abstract

We present experiment evidence of a systematic bias towards "false diversification" in decision-making under risk. I.e., the majority of our subjects were willing to pay to switch from a lottery that pays a prize in only one particular state of nature to a lottery that pays a prize in more than one state, *even though the overall distribution over prizes remains the same.* Further evidence suggests that this willingness-to-pay increases with the number of states. Additional treatments provide evidence against potential explanations such as a demand induced effect, regret, a failure to reduce compound lotteries, and an explanation based on bundled risk. The only explanation which seems to be consistent with some of the behavior observed is a model that introduces preferences over the source of risk. However, this model does not explain why one source would be preferred to another, or why the willingness-to-pay increases with the number of states. We conjecture that the bias we document may be a consequence of case-based decision-making.

*We would like to thank the following people: Dan Burghart, Gary Charness, David Levine, Bart Lipman, Steve Salant and Georg Weizsäcker for their helpful comments, Ryan Jaffari for programing the experiment and Chloe Tergiman for superb research assistance. Fréchette gratefully acknowledges the support of NSF via grants SES-0519045 and SES-0721111 as well as support from the Center for Experimental Social Science, and the C.V. Starr Center.

1 Introduction

Managing risk is an important component of our everyday life. We face this task in many decisions we make, such as savings, investments, insurance and retirement. Some of these decisions are quite complex, and as a result, individuals sometimes tend to employ heuristics in an effort to simplify their choice problem. For example, participants in defined contribution pension plans tend to invest an equal fraction either in all funds offered in the plan (what Benartzi and Thaler (01,07) refer to as the "$\frac{1}{n}$ heuristic"), or in a small subset of all funds (what Huberman and Jiang (06) refer to as the "conditional $\frac{1}{n}$ heuristic"). Similarly, financial advisors often warn individual investors from falling into the common trap of "false diversification" - the tendency to invest in a variety of funds, which appear to be well diversified but in actuality consist of highly correlated assets.[1]

Similar allocation rules have also been observed in experimental studies. In particular, Loomes (91) and Rubinstein (02) show that when subjects are asked to choose among a set of actions, each having some probability of yielding a given prize, subjects tend to violate first-order stochastic dominance by using mixed strategies that mimic the distribution of winning probabilities. Also related is the behavior observed in experiments where subjects are asked to bet on which color ball will be randomly drawn from an urn. Most subjects weakly prefer a bet over an urn that contains balls of each color (e.g., five red balls and five black balls) to a bet over an urn containing balls of only one, randomly determined color (e.g., either 10 red balls or 10 black balls - see Halevy (2007)). A similar bias towards "excessive diversification" was reported in Read and Loewenstein (1995) where subjects who chose in advance a set of n items to be consumed sequentially at n future dates (one item per period) picked a greater variety of items than subjects who chose n items sequentially, one item at each date of consumption.

The above behavior has the distinctive feature of spreading a resource - wealth or probability - across a variety of random variables (the returns generated by funds or the prizes generated by different actions in an experiment). Why do people tend to use decision rules that have this feature? Since this behavioral pattern appears in a variety of contexts, it is difficult to dismiss it as a simple mistake resulting from confusion. If it were a mistake, then why should it consistently lead to "excess" diversification rather

[1] see for example, the chapter on mutual funds in N.Y. Times bestseller book, "The Road to Wealth" by Suze Orman (Orman (2008)) as well as http://seekingalpha.com/article/23256-false-diversification-may-prove-costly-in-2007, http://abnormalreturns.com/2008/10/24/false-diversification-and-the-rise-of-novel-asset-classes/,

than excess "concentration"? In addition, in important decisions such as investment for retirement, people who feel confused or unsure would presumably seek advice or spend a long time thinking about the problem. Is this bias towards diversification just another manifestation of well-documented departures from expected utility theory? Or is there a different source for this behavior that cannot be captured by existing models in decision theory?

These are relevant questions since the decision of whether or not to diversify often comes up in many of the decisions we face. It is therefore important to understand whether the bias towards diversification can be rationalized with some preference relation (perhaps a non standard one), or whether the source of this bias is cognitive. If the former is true, then it is difficult to argue that individuals are behaving suboptimally. But if the latter is true, and individuals have a tendency to be biased towards diversification, then their decisions may be susceptible to manipulation. In addition, there may be situations – as in the case of choosing pension funds – where this bias would lead individuals to forgo potential earnings.

This paper attempts to address the above questions in an experimental setting. Specifically, our objective is to document a systematic bias towards what may be perceived to be diversification of risk, and examine whether this bias can be explained by existing models in decision theory. The main innovation of the paper is the experimental design, which is not only capable of generating sharp evidence on the bias, but also accommodates a variety of treatments that allow us to test several different models in decision-theory. The experimental design attempts to capture in a simple way a situation, which is frequently encountered in real life, where an individual is faced with a new decision problem he has not encountered before, and hence, has no opportunity to learn how to solve it. Our experimental design was not meant to mimic a real life dilemma, but rather to create a simple (albeit artificial) setting in which it is clear that the above heuristic is ineffective. The idea is that if subjects tend to employ some heuristic in this simple setting, in which probabilities and outcomes are explicitly spelled out, then this suggests that individuals are inclined to use it almost instinctively.

Our data reveals a systematic bias towards "false diversification". By this we mean that the majority of our subjects were willing to pay to switch from a lottery that pays a prize in only one particular state of nature to a lottery that pays a prize in more than one state, *even though the overall distribution over prizes remains the same.* We also present evidence suggesting that the likelihood of paying increases with the number of states (see Section 7).

More specifically, in our baseline treatment, subjects are told that they will participate in a lottery. It is described to them as drawing a ball from an urn. They are told that there will be three balls in the urn, each ball is marked by one letter, either X, Y or Z. Each letter represents an amount of dollars: one letter represents the amount \$30, while the other two letters represents the amount \$0. The matching between the letters and the two amounts of money will be randomly chosen by the computer. I.e., it is equally likely that $X = \$30$, $Y = \$0$ and $Z = \$0$, or that $X = \$0$, $Y = \$30$ and $Z = \$0$, or that $X = \$0$, $Y = \$0$ and $Z = \$30$. Subjects will not be told which letter corresponds to which amount of money.

At the beginning of the experiment, they are randomly assigned an urn containing three identical balls (the same letter). At the end, a ball will be randomly drawn from their urn, and they will either win \$30 or nothing. Our main finding (Sections 2 and 3) is that a significant proportion of our subjects were willing to pay to replace their starting urn with one that contained at least two distinct letters. However, very few subjects were willing to pay to have some chance of getting a different "homogeneous" urn (i.e., another urn with identical balls but with a different letter). In addition, the fraction of people willing to pay to change the urn they were endowed with is much higher than if subjects are endowed with an urn containing multiple letters and offered to replace it with a more "homogeneous" urn.

The decision-theoretic literature offers a wide variety of models that can potentially explain consistent departures from standard expected utility (for a recent survey of these works, see Starmer (2000)). To test whether some of these models can address the anomalous behavior of our subjects, we examined several variations of our baseline treatment. In particular, we investigate whether the results of the baseline treatment (described in Section 2) may be explained by a failure to reduce compound lotteries (Section 3), an attempt to avoid anticipated regret (Section 4), a preference over sources of risk as proposed by Tversky and Wakker (1995) (what they referred to as "source dependence") and Ergin and Gul (2002) (what they referred to as "issue preferences") (Section 5) and a Bayesian approach to uncertainty aversion as suggested by Halevy and Feltkamp (2005) (Section 6).

Among the theories we test, only a model of preferences over the sources of risk has some potential of addressing our results. However, this model does not explain why individuals may prefer one source of uncertainty over another, nor does it explain why the tendency of individuals to exhibit these type of preferences should depend on the number of states.

Since the above generalizations of expected utility theory do not seem to explain our subjects' behavior, the question remains, what *does* explain it? We *conjecture* that a possible explanation for the diversification bias of our subjects - as well as the naive diversification strategies reported elsewhere - may be the following. Some heuristics may grow out of strategies that were successful in a particular type of environment (one that is frequently encountered), but then become so rooted in our behavior that they turn into a "knee-jerk" reaction, which people employ inappropriately in other environments where the original strategy is not as successful (Gilovitch, Griffen and Kahneman (02)). This suggests that our subjects' behavior (and perhaps, the bias towards excessive/false diversification in general) may stem from the misapplication of the well known rule-of-thumb, "Don't put all your eggs in one basket!".

This suggests to us that perhaps a theory of case-based decisions (Gilboa and Schmeidler (1995, 2001)) may explain the behavior of our subjects as well as the behavioral patterns described above. When individuals face a decision such as choosing a retirement portfolio, guessing where a prize is hidden (as in Rubinstein's experiments) or deciding on which urn to bet (as in Halevy's experiments), oftentimes they have no prior experience, and hence, rely on their experience in other decision problems, which appear to be similar. However, individuals may judge two problems to be similar even though the optimal solutions to these problems may be very different. Consequently, they may end up applying decision rules, which performed well in past problems, to new problems in which they perform poorly. However, we do not have any evidence in support of this argument, and it is unclear to us how one could test for Case-Based Decision Theory.

Our findings hint at an additional source of concern regarding the validity of *some* experiments in which subjects are asked to make a series of choices, where only one of them is randomly selected to be played out for real. Holt's (1986) original criticism against this methodology was driven by the possibility that subjects may fail to reduce compound lotteries. According to Holt, a subject may treat the series of decisions as one big compound lottery. If such a subjects is *not* indifferent between compound lotteries that induce the same distribution over outcomes, then the choice he makes in a single decision may not reveal his preferences over the possible alternatives in that decision. Starmer and Sugden (1991), have argued that this is not a concern if individuals exhibit "narrow bracketing": the tendency to solve multiple decision problems in isolation without regards to the final payoff from all the decisions (for recent evidence on this phenomenon, see Rabin and Weizsäcker (in press)). Thus, according to this argument, if individuals violate expected utility theory *twice*, by failing to reduce compound lotteries *and* by failing to consider the distribution over their final wealth, then Holt's critique can be safely ignored.

However, the experimental results reported here suggest that in fact, individuals may exhibit "broad bracketing" in situations where they should be "narrow bracketers". The reason for this is, that in some experiments, broad bracketing may give a false sense of diversification. For example, consider experiments in which subjects are asked to answer multiple questions, which are meant to elicit the value of some preference parameter such as risk-aversion, ambiguity-aversion or discount-factor. Even if subjects are explicitly told that only one of the questions will be randomly drawn to be played out for real, they may approach this questionnaire in a similar way that they approach the urn in our experiment. I.e., they may view the answers to the questionnaire as "balls in an urn", and even though only one answer ("ball") will be drawn, they may be tempted to give a variety of answers in a way that gives them a false sense of diversification; for example, by picking the riskier alternative in one decision and the safer alternative in another. Furthermore, our findings suggest that the source for this type of behavior may not be the failure to reduce compound lotteries, as Holt suggested. It follows that one should exercise extra caution in interpreting subjects' choices in *some* experiments where only one of the choices is actually carried out.[2]

A common criticism that is often raised against experimental evidence of anomalous behavior is that such behavior would be eliminated outside the lab where individuals are able to communicate and discuss difficult decision problems. We argue that this critique overlooks several important points. The simplicity of a laboratory design is meant to serve as an abstract model that allows the researcher to cleanly identify a systematic pattern of behavior and be able to isolate its. Because of the simplicity of the design, it may be easy to detect suboptimal behavior in the lab - and hence, easily corrected in that particular context - it may be less obvious in real life decision problems that give rise to similar anomalies. The question is not whether we can eliminate the anomaly in the lab by guiding the subjects, but whether the anomalous behavior will resurface in other, more complicated situations (that have similar features) in the absence of guidance or intervention.[3]

[2] Here we qualify the statement with *some* as these types of concerns may very well be irrelevant for some types of experiments. For instance in multilateral bargaining games, varying payment methods have been used (paying for one round or multiple rounds) without any noticeable impact (Fréchette, Kagel and Lehrer 2003; Fréchette, Kagel and Morelli 2005). We speculate that the simpler is the task to be performed, the greater the concern because it is easier for the subject to consider the experiment in its entirety when each part is very simple, and thus individual decision making experiments which repeat the same task, especially biary choices over simple urns with varying amounts of risk might be particularly sensitive.

[3] As an example, consider the simple experiments on "narrow bracketing" by Tversky, and Kahneman (1981) (and more recently, Rabin and Weizsäcker (in press)). It may not be difficult to design

In addition, it is not necessarily true that outside the lab, communication between individuals would eliminate behavioral anomalies. First, if the majority of individuals with whom a decision-maker communicates suffer from the same bias, he may not arrive at the conclusion that his initial inclination is wrong. Second, when the individual's bias stems from applying some well-known heuristic (albeit in the wrong environment), the decision-maker may not feel the need to consult with others, as he may believe that he has solved his decision problem.

2 The bias

Consider an urn that can be filled with balls, each marked by one of the three letters, X, Y, Z. One - and only one - of the letters is randomly chosen to represents a prize of \$30. The remainder letters represent \$0. You are not told which letter was chosen to represent the prize. You are handed an urn with three X balls. If you keep this urn, one of the balls will be randomly drawn and you will be paid either \$30 or \$0 depending on whether or not $X = \$30$. Alternatively, you may pay \$1 and replace one X ball with either a Y ball or a Z ball, or you may pay \$2 and replace two X balls with a pair of Y and Z ball. Again, only one ball will be drawn from the new urn and you will be paid according to which ball was drawn and which letter was chosen to represent \$30.

What will you do?

This question was posed to subjects in a series of computerized experiments that were conducted in the laboratory of the Center for Experimental Social Science at New York University. Subjects were recruited from the undergraduate population at New York University. A total of 74 subjects participated in the baseline treatment, which consisted of the decision problem described above. In each session, subjects were handed written instructions, which they first read on their own and were later read out loud by the experimenter (these instructions appear in the Appendix). In the instructions subjects were informed that about a third of them would start off with an urn having three X balls, another third would be assigned three Y balls, and the remaining subjects would have an urn with three Z balls. Each subject received a show up fee of \$8, which they could keep regardless of the decisions they made in the experiment. The experiment lasted around 45 minutes and the average payment

treatments in which subjects are primed to realize that they are violating first-order stochastic dominance. However, the point of these experiments is that outside the lab, in the absence of guidance, individuals are highly susceptible to manipulations that exploit their tendency to exhibit narrow bracketing.

Baseline	subjects	paid $1	paid $2	total paid	%
session 1	16	4	3	7	44%
session 2	16	2	9	11	69%
session 3	16	9	5	14	67%
session 4	21	8	5	13	62%
Total:	**74**	**23**	**22**	**45**	
%		**31%**	**30%**	**61%**	

Table 1: Choices in the Baseline Treatment

(including the show-up fee) was $21.69. A post experiment questionnaire was administered after the experiment was over asking subjects to explain their actions. Table 1 below summarizes the results of the baseline treatment.

As evident from the table, more than 60% of the subjects chose to pay for an urn with at least two types of balls. Slightly more than half of these subjects paid $1, while the remainder paid $2.

To better judge the magnitude of these percentages, it is helpful to compare them to percentages reported in previous experimental works on anomalies. The most well known anomaly in decision-making under risk is the Allais Paradox. In one of the earliest experimental studies of this paradox, MacCrimmon (65) reported that 40% of his subjects exhibited this paradox. Later studies by Slovik and Tversky (1974) and Conslik (89) reported 60% and 50% of occurrences, respectively. Two other well known "paradoxes" are the "Common Consequence Effect" and the "Preference Reversal" phenomenon. Starmer (92) reported 30%-40% occurrences of the first paradox, while Tversky, Slovic and Kahneman (90) reported that 46% of their subjects exhibited the second anomaly. Note, however, that in contrast to our study, the above papers on the Allais paradox and the common consequence effect did not investigate intensities of preferences, but only examined pair-wise choices of lotteries.

Some insights into what made subjects pay for a diversified urn may be obtained by examining their answers to the post-experiment questionnaire. Of course, not all of the answers provided are easily interpreted and some may not truly reflect the subjects' state of mind at the point of decision. Still, a general message that comes out of the subset of coherent answers is that on the one hand, some subjects were not confused about the probability of winning, but on the other hand they cared about having the "winning ball" in their urn. Some representative answers to the question, "why did you pay?" that fit this class are the following:

- "to try to increase my chances of getting the money ball"

- "The only guarantee that one could make in this experiment was to make sure that the ball that was worth 30 dollars was present in the pot when the computer made the drawing"

- "because given the probability $\frac{2}{3}$ of receiving a \$0 urn, it seemed best to diversify urn as much as possible in order to increase probability of urn containing \$30 ball"

- "Yes, I did. If I didn't pay, there was an all or nothing chance that I had or didn't have a single ball worth \$30 in my urn. I would rather have paid and had an equal chance that one of my balls is worth \$30 than to risk the chance of not having any worth \$30"

We relate our experimental design to the "don't put all your eggs in one basket" heuristic by interpreting the set of distinct balls as the set of available "baskets". Thus, a subject who is willing to pay to diversify his urn is viewed as a subject who prefers to spread his eggs among more than one basket. In what follows we investigate whether the bias our subjects exhibited towards diversified urns may be explained by some of the decision-theory models discussed in the literature.[4]

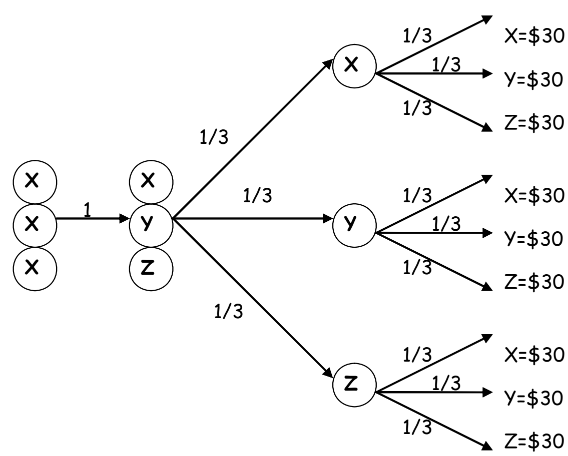
3 Did subjects fail to reduce compound lotteries?

While the *probability of winning* a prize is independent of the composition of the urn, different urns induce different *compound lotteries*. Therefore, if subjects fail to reduce compound lotteries (as previous experimental studies have argued, e.g. Halevy (07) and the references therein), then some may strictly prefer the lotteries induced by urns with at least two distinct balls. To investigate this possibility, we conducted the following variation of the original treatment (we refer to this treatment as the reduction of compound lotteries, or ROCL, treatment).

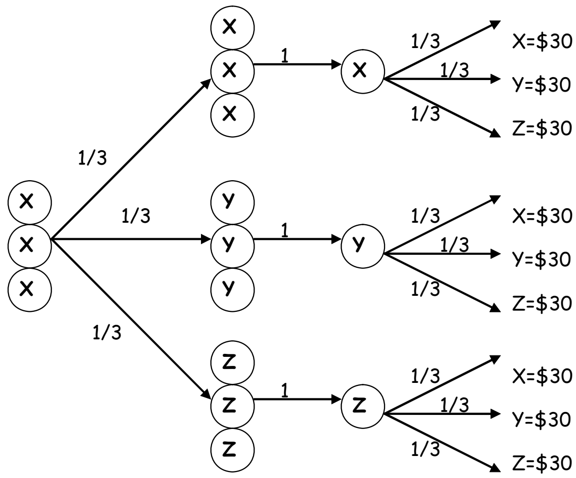
As in the baseline, subjects begin the experiment with an urn having three identical balls. But now, instead of exchanging balls, subjects have the following options. They may pay a \$1 to have a $\frac{1}{3}$ chance of replacing their assigned urn (say (X, X, X)) with an urn having three identical balls of their choice (i.e., either (Y, Y, Y) or (Z, Z, Z)). Alternatively, they may pay \$2 to have an equal chance of either keeping their assigned urn, replacing it with (Y, Y, Y) or replacing it with (Z, Z, Z).

[4] The instructions of these treatments are similar to those of the baseline treatment and are available upon request.

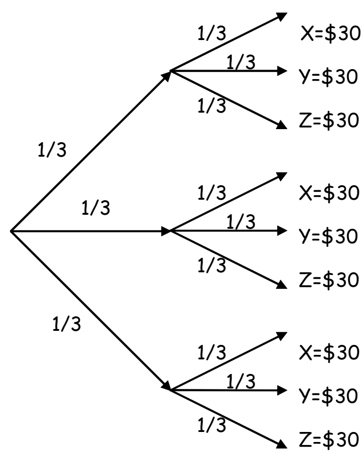
Note that subjects may hold different beliefs about the order in which the different random variables are determined. For example, they may believe that first, it is determined which letter represents the prize, second it is determined which urn they have, and only then a ball is drawn from their urn. Alternatively, they may believe that first it is determined which urn they have, then a ball is drawn from their urn, and only then it is decided which letter represents the prize. However, it seems natural that the stage in which a ball is drawn would come only after the stage in which the urn was determined. This leaves three possible orderings of the stages: urn → ball → letter, urn → letter → ball, and letter → urn → ball. Fix any such order and assume subjects reduce degenerate lotteries (or degenerate "branches on a lottery tree"), then the compound lottery induced by the original treatment is *identical* to the one induced by this treatment. This is displayed in Figures 1-3, which address the case in which subjects believe in the following sequence of realizations: urn → ball → letter.



Paying $2 in Baseline treatment
Figure 1



Paying $2 in ROCL treatment
Figure 2



Reducing degenerate lotteries
Figure 3

However, subjects may fail to reduce even degenerate stages in a compound lottery. Consequently, such subjects may view the compound lotteries induced by paying in Treatments 1 and 2 as two different lotteries. Models of "recursive expected utility" (e.g., Kreps and Porteus (1978), Segal (1990), Klibanoff, Marinacci and Mukerji (2005) and Ergin and Gul (2004)) allow for such failures. The various representation theorems in this literature provide the conditions under which there exist three real valued functions, u, v and w such that an individuals' valuation of a compound lottery may be described as follows. At the end of the third stage, the individual evaluates each of the final outcomes according to the utility index u. Thus, any third stage lottery may be viewed as a lottery over different values of u. The value of each third-stage lottery is computed by substituting the expected value of u into v. Similarly, the value of each

	urn → ball → letter
Not paying	$w\{v[\frac{1}{3}u(30)+\frac{2}{3}u(0)]\}$
Paying \$2 in T1	$w\{v[\frac{1}{3}u(30)+\frac{2}{3}u(0)]\}$
Paying \$2 in T2	$w\{v[\frac{1}{3}u(30)+\frac{2}{3}u(0)]\}$

Table 2a

	urn → letter → ball
Not paying	$w\{\frac{1}{3}v[u(30)]+\frac{2}{3}v[u(0)]\}$
Paying \$2 in T1	$w\{v[\frac{1}{3}u(30)+\frac{2}{3}u(0)]\}$
Paying \$2 in T2	$w\{\frac{1}{3}v[u(30)]+\frac{2}{3}v[u(0)]\}$

Table 2b

	letter → urn → ball
Not paying	$\frac{1}{3}w\{v[u(30)]\}+\frac{2}{3}w\{v[u(0)]\}$
Paying \$2 in T1	$w\{v[\frac{1}{3}u(30)+\frac{2}{3}u(0)]\}$
Paying \$2 in T2	$w\{\frac{1}{3}v[u(30)]+\frac{2}{3}v[u(0)]\}$

Table 2c

Table 2: Valuations of options according to Recursive Expected Utility

second-stage lottery is computed by inserting the expected value of v that it generates into the function w. The overall value of the three-stage lottery is then computed as the expected value of w.

To derive the predictions of recursive expected utility (REU henceforth) in our set-up, we evaluate the different options a subject faces in Treatments 1 and 2 for each possible sequence of stages. Tables 2a-2c display the results for the compound lotteries induced by not paying (i.e., drawing a ball from a homogeneous urn with three identical balls), paying \$2 in Treatment 1 (i.e., drawing a ball from an urn with all three types of balls) and paying \$2 in the current treatment (i.e., drawing a homogeneous urn at random and then drawing a ball from that urn). The rankings induced by these valuations will be the same if we were to compare not-paying to paying-only-a-Dollar in each treatment.

According to Tables 2a-2c, there are two predictions we can make without imposing additional assumptions on the functions u, v and w. First, a REU subject who believes in the sequence, urn → letter → ball, would *not* pay in Treatment 2. Second, a REU subject who believes in the sequence, urn → ball → letter, would *not* pay in *both* treatments, since the value of the lottery he obtains is the same whether he pays or not. Further predictions would necessitate some assumptions on the relative curvatures of v and u.

To understand the implications of such assumptions it is helpful to relate our first two treatments to the experimental design that was recently used by Halevy (2007) to link preferences over compound lotteries to ambiguity aversion. Subjects in his experiment were faced with four urns, each containing 10 balls, where each ball was either red or black. Subjects were asked to guess which color ball will be drawn from each urn and also to specify a minimum selling price for each urn. We shall focus on three of these urns: *urn 1*, which contained five red balls and five black balls, *urn 2*, which contained an unknown distribution of red and black balls and *urn 4*, which had a 50-50 chance of containing either 10 red balls or 10 black balls. Subjects who ranked urn 1 above urn 2 and urn 4 were identified as subjects whose ambiguity aversion could be accommodated by an REU model with "second-order risk aversion". Put differently, the utility function these subjects apply to the final-stage lottery (i.e., u) is *less* concave than the function they apply to the lottery in the preceding stage (v).

Consider modifying Halevy's design by having nature, rather than the subject himself, guess a color at random. The subject is then paid a prize whenever the color of the drawn ball matches nature's guess. This modification allows us to add a new urn, say *urn 5*, to the two urns described above, one which contains 10 red balls with certainty. Our first two treatments are essentially analogous, in this modified set-up, to endowing subjects with urn 5 and asking them if they would pay \$2 to switch to urn 1 in the baseline treatment and to urn 4 in the ROCL (to be more precise, we also add a third possible color and five additional balls to each urn).

We say that a subject with REU preferences exhibits "higher-order risk-aversion" if w is more concave than v, and v is more concave than u. Note that with two-stage lotteries, this definition reduces to second-order risk-aversion, which has been linked to ambiguity aversion. From Tables 2a-2c it follows that whenever the letter is believed to be chosen first, an REU subject with higher-order risk aversion strictly prefers an urn with all three types of balls to an urn with three identical balls. But such a subject also prefers to draw at random an urn with three identical balls than to be given one of these urns with certainty. This suggests that such a subject would not pay in *both* the baseline treatment and the ROCL treatment. Tables 2a-2c also indicate that if a subject believes that the letter is not chosen in the first stage, then the relative concavities of w and v do not affect his willingness to pay. However, if v is more concave than u, and a subject believes that the letter is chosen *second*, then he may pay in the baseline treatment (since he prefers a diversified urn to a homogeneous one), but he would not pay in the ROCL treatment (since he is indifferent between a homogeneous urn with certainty and a lottery over homogenous urns).

Table 3 displays the results of the ROCL treatment. A total of 55 subjects partic-

ROCL	subjects	paid \$1	paid \$2	total paid	%
session 1	9	1	1	2	22%
session 2	13	2	1	3	23%
session 3	19	5	3	8	42%
session 4	14	3	2	5	36%
Total:	55	**11**	**7**	**18**	
%		**20%**	**13%**	**33%**	

Table 3: Choices in the ROCL Treatment

ipated in this treatment, which also lasted about 45 minutes, and where the average payment (including the show-up fee) was $13.

Pooling all sessions together yields that about a third of the subjects paid at least $1 (compared to 61% in the baseline) where the ratio of subjects who paid $2 to those who paid $1 was roughly 2 : 1. Examining the pooled data in each of the first two treatments, we find that a Fisher's exact test rejects (at the 1% level) the hypothesis that the same number of subjects paid at least 1$ in both treatments. We find a similar conclusion when we compare the number of subjects who paid $0, $1 and $2.

The key result in this treatment is that the fraction of subjects willing to pay in the baseline is higher than in the ROCL treatment. This means that if a subject pays in the baseline treatment only because he fails to reduce compound lotteries, then such a subject also fails to reduce *degenerate* stage-lotteries. Furthermore, our subjects' behavior is consistent with REU preferences and higher-order risk aversion, only if they believe that the winning letter is determined in the *second* stage. This seems to us a rather unnatural sequencing of events, and we conjecture that the subjects who paid did not believe that the random variables were determined in this order. While a more general form of REU can, in principle, be consistent with our results, most of the literature has focused on either the case of higher-order risk neutrality (which is inconsistent with our results) or higher-order risk aversion. In light of this, we interpret the combined results of the first two treatments as evidence against an explanation that relies solely on the failure to reduce compound lotteries.

The results of the current treatment also provide some evidence against obvious explanations of our baseline data. First, it does not seem that subjects in our baseline treatment paid simply because they were confused. Such an explanation cannot account for the fact that more subjects pay in the baseline than in the ROCL treatment since both treatments follow nearly identical procedures. If subjects were simply confused, one would expect to find the same fraction of subjects paying in both cases.[5] Second,

[5] An additional concern might be that because subjects were confused about what they were sup-

it is difficult to attribute our subjects' behavior to a mere demand induced effect.[6] Again, a demand induced effect cannot account for the fact that the frequency of subjects paying is lower in the ROCL treatment than in the baseline treatment since in both cases subjects were asked to pay for an alternative option. Finally, one could argue that in the baseline, mistakes can only go against standard expected utility. However, the comparison between baseline and ROCL could go either way. Finally, the current treatment suggests that our subjects' behavior may not stem from an "illusion of control". This is typically defined as the tendency to falsely believe that our actions can affect the realization of a state of nature (see Langer (1975) and the recent experiment by Charnass and Gneezy (in press)). This definition suggests that subjects in the baseline treatment paid because they falsely believe that the mere act of paying would tilt the outcome in their favor. But this also suggests that they should keep paying in the ROCL treatment, in contrast to what we find.

4 Are subjects trying to avoid anticipated regret?

Consider a subject who started out with an (X, X, X) urn and decided not to pay. Suppose he finds out at the end of the experiment that he won nothing. This subject infers that $X = \$0$, and therefore concludes that by his decision not to pay, he guaranteed himself zero chances of winning. Such a subject may then regret not having paid at least \$1 as this would have allowed him (in retrospect) at least some chances of winning. Moreover, given that either Y or Z correspond to the prize, if he had paid \$2, he would have had a "winning" ball for sure. This suggest that subjects' willingness to pay may stem from a conscious effort to avoid regret.[7]

Regret-avoidance is one explanation for people's tendency to use naive diversification strategies. In fact, even Harry Markowitz, the pioneer of modern portfolio theory, admitted that regret was the reason why he used such strategies: "My intention was to minimize future regret. so I split my contribution fifty-fifty between bonds and equities." (Benartzi and Thaler (2001)).

To test for whether regret-avoidance is the source of our subjects' willingness-to-pay

posed to do, they viewed an urn with more than one letter, as a "safety net". This suggests that when presented with an urn that already has two distinct letters, a substantially lower fraction of subjects would pay for a third letter. However, this is not the case, as we show in Section 6.

[6] We take demand effects to be a tendency to do something simply because the option is offered. In our experiments, this would be to pay to change the urn, independently of the original urn or the resulting urn.

[7] For formal models of anticipated regret, see Bell (1982), Loomes and Sugden (1982), Sugden (1993), Hayashi (2007) and Sarver (2008).

we conducted the following treatment. It is identical to the baseline treatment except for the following change: one of the three letters now represented a 30% chance of winning $100 (so that the expected prize is $30), while the other two letters represented zero chances of winning. Hence, a subject who does not win the prize can never learn whether he lost because he did not have the "winning" ball in his urn, or whether he lost because he drew the "winning" ball but ended with a bad realization.

The baseline treatment and the regret treatment can both be viewed as two special cases of a treatment where the "winning ball" is actually a lottery ticket, or the right to participate in a lottery where the probability of winning a prize is p. Suppose you started off with (X, X, X), you decided *not* to pay, and you did *not* win. Feelings of regret can be described in the following form. You say to yourself, "Suppose I could go back in time to the moment of decision (to pay or not to pay) 'knowing what I know now'. Could I have done better by choosing to pay?"

The question is, what does one mean by "knowing what I know now" and "could I have done better"? This is not very well defined in this hypothetical time-travel thought exercise. Here are some possible interpretations of this statement:

1. "Am I certain that by paying I would have had better chances?" If not, I don't feel regret.

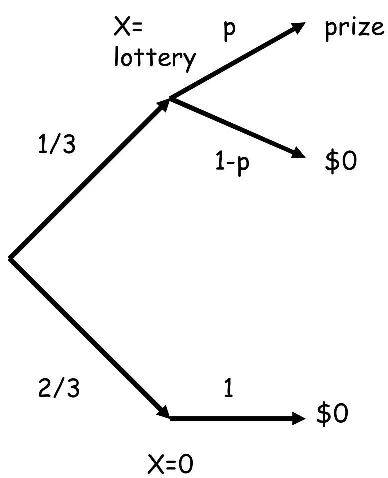
2. You believe that when you go back in time in your thought exercise, the lottery-ticket ball will be the same *and also* the outcome of the lottery will be the same. This means that if X *is* the lottery ticket, it promises zero dollars. Hence, whether or not X is the lottery ticket ball, an X ball guarantees zero and so you regret not having paid.

3. You believe that when you go back in time in your thought exercise, the lottery-ticket ball will be the same *but* the lottery will be executed *again*. The idea behind this notion of regret is that the letter corresponding to the prize is an underlying state of nature, which is a fixed feature of the environment, while all other random draws can be made again and again. In this case, you calculate the probability that $X = 0$, conditional on having won zero when X was chosen. You then use the updated probabilities when making the decision again (assuming that if $X = 0$, it is equally likely that either Y or Z is the lottery ticket).

The first and second notions of regret are in some sense extreme. In the first, feelings of regret are completely eliminated when there is even the slightest doubt on whether one could have done better by choosing a different action. In the second notion of regret, a person believes that if there was a possibility to redo history, he would be able to change his decision, but *all* future realizations of nature would be exactly the same. Our regret treatment completely eliminates the first notion of regret, but cannot address the second notion.

If one is willing to entertain a less extreme form of regret - specifically, the third notion described above - then this notion *is* addressed by our regret treatment. In particular, we argue that our regret treatment significantly diminishes the scope of this form of regret. Hence, if this is the sole source of our subjects' willingness to pay, then we should expect a much lower occurrence of payments in the present treatment, compared with our baseline.

To see this, consider the following history: first, nature chooses which letter corresponds to the lottery ticket, second, a ball is drawn from an (X, X, X) urn, and third, an outcome of \$0 is observed. Conditional on this history, and assuming the letter corresponding to the lottery remains fixed, the probability that $X = \$0$ is $\frac{2}{3-p}$ (see Figure 4)



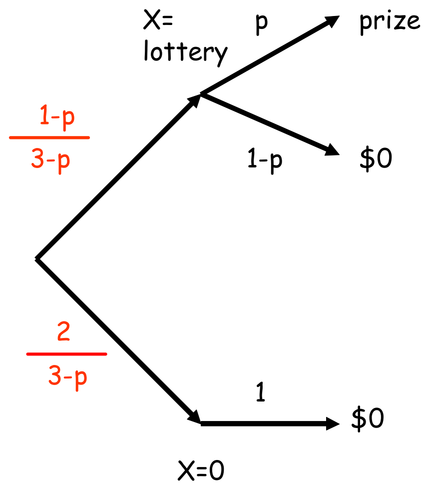
keeping (X,X,X)
Figure 4

Consider now a subject who needs to decide whether to draw a ball from an urn U, where

$$U \in \{(X, X, X), (X, X, Y), (X, X, Z), (X, Y, Z)\}$$

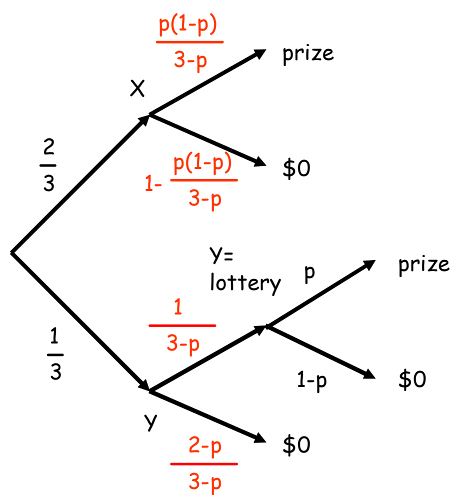
assuming that if he draws the lottery ticket, the lottery will be played again. Let $q_U(p)$ denote the probability of winning the prize with urn U, given that the probability that

$X = \$0$ is $\frac{2}{3-p}$ (see Figures $5-7$)



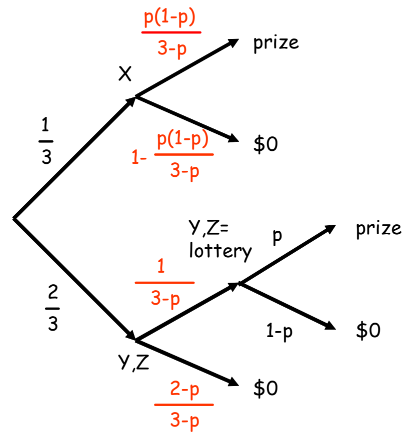
keeping (X,X,X) again

Figure 5



Switching to (X,X,Y)

Figure 6



Switching to (X,Y,Z)
Figure 7

Note that

$$q_{(X,Y,Z)}(p) > q_{(X,X,Y)}(p) = q_{(X,X,Z)}(p) > q_{(X,X,X)}(p)$$

Define

$$\begin{aligned} R_1(p) &\equiv q_{(X,X,Y)}(p) - q_{(X,X,X)}(p) = q_{(X,X,Z)}(p) - q_{(X,X,X)}(p) \\ R_2(p) &\equiv q_{(X,Y,Z)}(p) - q_{(X,X,X)}(p) \end{aligned}$$

and assume that a person's feelings of regret for not having paid t dollars, $t \in \{1, 2\}$, are strictly increasing in $R_t(p)$. Note that for $p \in [0, 1]$, $R_t(p)$ is convex and increasing in p. This means that compared with the baseline treatment, in the regret treatment, we significantly decrease $R_1(p)$ and $R_2(p)$ (which may be interpreted as significantly decreasing the "scope" of regret). To see this, notice that in our baseline treatment, $R_1(1) = \frac{1}{6}$ and $R_2(1) = \frac{1}{3}$, while in our regret treatment, $R_1(0.3) = 0.001$ and $R_2(0.3) = 0.002$. However, our findings (presented in Table 3) suggest that even though we significantly lowered the scope of regret, the willingness-to-pay remained almost the same.

A total of 84 subjects participated in the regret treatment (see Table 4), where the average payment was approximately \$8. Pooling the subjects in all the sessions of this treatment and in the baseline treatment, we find that a Fisher's exact test cannot reject (at the 10% level) the hypothesis that the same number of subjects paid at least 1\$ in both treatments. We find a similar conclusion when we compare the number of subjects who paid \$0, \$1 and \$2.

The conclusion we may draw from our regret treatment is that the subjects' willing-

Regret	subjects	paid \$1	paid \$2	total paid	%
session 1	21	8	4	12	57%
session 2	21	6	7	13	62%
session 3	21	7	4	11	52%
session 4	21	8	5	13	62%
Total:	84	**29**	**20**	**49**	
%		**35%**	**20%**	**49%**	

Table 4: Choices in the Regret Treatment

ness to pay, (*i*) *cannot* be due to the **first** form of regret, (*ii*) *is not likely* to be caused by the **third** form of regret, but (*iii*) *could* be due to the **second** form of regret.[8]

5 Are subjects exhibiting "preferences over issues"?

Tversky and Wakker (1995) and Ergin and Gul (2002) propose models of choice under risk and uncertainty that accommodate "issue-preference" or "source dependence" (henceforth we shall use the former term): agents may not be indifferent among gambles that yield the same probability distribution if they depend on different sources of randomness ("issues"). In our context there are three issues: (*i*) which letter represents the prize (issue l), (*ii*) what is the urn (issue u), and (*iii*) which ball is drawn from the urn (issue b). Consider three equivalent lotteries, where one depends both on issue l and issue b (lottery (l,b)) a second depends on issue l alone (lottery (l)) and a third depends both on issue l and issue u (lottery (l,u)). Our data from the first two treatments suggest that for the majority of our subjects,

$$(l,b) \succ (l) \succ (l,u)$$

The reason for this is that not paying results in a lottery that depends only on issue l, paying in the baseline treatment results in a lottery that depends on both issue l and b (the urn is known), and finally, paying in the ROCL treatment results in a lottery

[8]As we remarked above, we view the second form of regret as somewhat extreme, and moreover, we cannot think of a treatment that is capable of controlling for it. In addition, a subject who anticipates this form of regret may actually prefer not to pay \$2. To see why, imagine a subject who started with three X balls, paid \$2, and did not win. This can happen for two reasons: (*i*) the ball that was drawn represented zero chances of winning, or (*ii*) the ball representing 30% of winning was drawn but the outcome of the lottery was no prize. If the ball drawn was X, then even if the subject had not paid, he would not have won, but at least he would have saved \$2. If the ball drawn was different than X, then the subject might have actually won had he remained with only X balls. Since more than 40% of the subjects, who paid in the experiment, paid \$2, we conjecture this was not driven by the above form of regret.

2nd Phase / 1st Phase	**Start with Lottery**		**Start with (X,Y,Z)**	
	Paid for (X,Y,Z)	Did not pay	Paid for lottery	Did not pay
Paid for (X,Y,Z)	**Yes**	**No**	**No**	**Yes**
Did not pay	**No**	**Yes**	**No**	**Yes**

Table 5: Behavioral patterns that are consistent with the TWEG model

that depends on both issues l and u (given the urn, there is no uncertainty over which ball will be drawn).

We propose an indirect test, which is capable of providing some evidence on whether or not issue-preference is at the heart of our bias. The basic idea of this test is to ask whether subjects who exhibit $(l, b) \succ (l) \succ (l, u)$ also exhibit $(b) \succ (u)$ when issue l is resolved. In the new treatment (henceforth, the Tversky-Wakker-Ergin-Gul, or TWEG, treatment) subjects first participate in the baseline treatment. We refer to this part of the treatment as the "first phase". After they make their decision but *before* they are informed of the outcome, they are asked whether they want to continue with the first phase, or whether they would like to abort that phase (not be told the outcome and not be paid for that phase) and participate in a brand new treatment (the "second phase"), which is independent of the first, and where the average earnings are twice as high.[9]

In the new treatment, $X = \$60$ and $Y = Z = \$0$. Half of the subjects are told that the computer will randomly select for them an urn with three identical balls such that they will have an equal chance of getting either (X, X, X), (Y, Y, Y) or (Z, Z, Z). However, they may pay \$2 and receive an urn with all three balls (X, Y, Z) and have a ball randomly drawn from this new urn. The remaining subjects will be given an urn with (X, Y, Z), from which a ball will be randomly drawn. If they pay \$2, then they will get either the urn (X, X, X), (Y, Y, Y) or (Z, Z, Z), each with equal probability.

Table 5 displays the different patterns of behavior that may be observed in the TWEG treatment. Each cell in the table depicts a possible scenario describing what the subject did in the first phase, whether he started with a diversified urn or a lottery over homogeneous urns in the second phase, and his decision in the second phase. A cell marked by "yes" represents a pattern consistent with issue-preferences, while a cell marked by "no" represents an inconsistent pattern.

Table 6 displays the results of the TWEG treatment. A total of 79 subjects partic-

[9]The goal was to have a design with the advantages of a standard within-subject format but such that (*i*) there is no feedback between the two phases, (*ii*) the outcome of the first phase does not affect the second phase, and (*iii*) the knowledge of the second phase does not affect decisions in the first phase.

2nd Phase / 1st Phase	**Start with Lottery**		**Start with (XYZ)**	
	Paid for (XYZ)	Did not pay	Paid for lottery	Did not pay
Paid for (XYZ)	65%	35%	25%	75%
Did not pay	14%	86%	0%	100%
Total	**37%**	**63%**	**16%**	**84%**

Table 6: Choices in the TWEG Treatment

ipated in this treatment, where the average payment was about $22.

The main message from this table is that the behavior of the majority of our subjects is roughly consistent with the TWEG model.[10] Consider first the group of subjects who paid to switch from lottery (l) to lottery (l, b) in the initial phase. About 65% of those who started out with (u) in the second phase, paid to switch to lottery (b), and 25% of those who started with (b) paid to switch. Consider next the group of subjects who did not pay in the initial phase. None of those subjects who started out with (b) paid to switch to (u), and 14% of the subjects who started out with (u) also paid to switch.

In light of these findings, the following questions remain unanswered: (i) why do subjects prefer the ball issue to the urn issue? and (ii) how do we explain the behavior of those subjects who paid to switch from lottery (l) to lottery (l, b), but when endowed with lottery (u), did not pay to switch to lottery (b)?

Our findings in this treatment, together with our findings in the ROCL treatment, hint at an interesting relation between issue-preferences and preferences between one-stage and two-stage lotteries. In both the baseline and the ROCL treatments subjects were offered the opportunity to switch from a one-stage lottery to a two-stage lottery. They were willing to pay for the switch in the baseline treatment but not in the ROCL treatment. The two-stage lotteries in each treatment differed in the source of risk in the second stage. In the baseline treatment this source was the random draw of a *ball* from an urn with three distinct balls, while in the ROCL treatment, the source of risk was the random draw of an *urn* with three identical balls. This suggests an interesting framing effect: individuals' preferences between one-stage and two-stage may be sensitive to the sources of risk at each stage. To the best of our knowledge, this framing effect has yet to be explored in the literature.

The first phase of the TWEG treatment also provided additional data points for the baseline treatment. Indeed, a Fisher's exact test cannot reject (at the 10% level)

[10] 96% of the subjects who participated in this treatment decided not to abort their decisions in the initial phase of the experiment (which was identical to the baseline), and continue to the second phase (where issue l was resolved).

Treatment	Paid $1	Paid $2	Total paid
Baseline	30%	27%	57%
ROCL	20%	13%	33%
Regret	34%	24%	58%

Table 7: Choice frequencies adding the first part of the TWEG treatment

the hypothesis that the same number of subjects paid at least 1$ in both treatments (TWEG Phase 1 and Baseline). We find a similar conclusion when we compare the number of subjects who paid $0, $1 and $2. Pooling the results of the original baseline treatment and of the first phase of the TWEG treatment (a total of 153 subjects), we obtain that about 57% of the subjects paid at least $1 for a heterogeneous urn. Table 7 presents the percentage of paying subjects across the various treatments (with the baseline treatment pooled together with the first phase of the TWEG treatment) and across the two possible fees.

6 Do subjects misperceive the situation as one of "bundled risk"?

Halevy and Feltkamp (2005) argue that in many real life decisions, individuals face "bundles" of correlated random variables such that the final payoff is the sum of these variables (one example they give is the repair cost of a car, which consists of the sum of the repairs of different components). Such "bundles of risks" may be thought of as drawing more than one ball *with replacement* from an urn, and receiving a prize, which equals the sum of prizes associated with the balls (they call this "bundled risks"). Because of this, subjects sometimes are "programmed" to "instinctively" think of a situation as one in which there is some small probability that more than one ball will be drawn *with replacement*.

This implies that if under standard expected utility, a risk-averse individual will be indifferent between two gambles when only one ball is drawn, he may not be indifferent if he gives some small chance, say ε, to the event that two (or more) balls will be drawn (and the prize will equal the sum). Halevy and Feltkamp use this logic to show that even if you take ε to be arbitrarily small, standard expected utility and risk-aversion can explain the Ellsberg paradox.

This logic may also explain why subjects pay in our baseline and regret treatment, but not in the ROCL treatment. An expected utility maximizer is indifferent about the composition of the urn when only one ball is drawn. However, a risk averse individual strictly prefers the lottery induced by drawing with replacement two balls from (X,Y,Z) versus drawing two balls with replacement from (X,X,X) since the latter lottery is a mean preserving spread of the former.

The Halevy-Feltkamp theory may be interpreted as providing a rationale for having a preference over a ball issue. This theory (which is aimed at providing an expected-utility rationale to the Ellsberg paradox) addresses a situation in which there is only a single draw after the state of nature is determined (in the Ellsberg world, the state is the composition of the urn, whereas in our case, it may be viewed as which letter corresponds to the prize). Since there is only a single draw - a ball from an urn - the "bundled risk" argument applies only to the ball, which means that more than one ball may be drawn.

In our decision problem there are *two* draws: an urn is drawn and then a ball is drawn from the chosen urn. This raises the question, to which draw should we apply the bundled risk argument? Do subjects give some small chance that more than one urn will be drawn (with replacement) and then only one ball will be drawn from each urn? Or maybe more than one urn is drawn and more than one ball is drawn from each urn? Because Halevy and Feltkamp were mainly concerned with the Ellsberg environment, their original model does not provide answers to these questions. If, indeed, more than one urn is drawn and then only a single ball is drawn from each, then bundled risk plus rule-rationality implies that people should pay in the ROCL treatment, in contrast to what our findings indicate.

Suppose we accept the Halevy-Feltkamp theory as is, and assume that only balls may be drawn more than once with replacement. Then the following treatment may provide evidence on whether or not this theory explains our subjects' behavior. Subjects start off with an urn having exactly two distinct balls. In particular, about a third of the subjects are endowed with (X,Y), another third with (X,Z), and the remaining third with (Y,Z). A subject may keep his assigned urn, in which case, a single ball would be randomly drawn and the subject wins either \$30 or \$0 depending on which letter corresponds to the prize. Alternatively, a subject may pay \$1 to add a third distinct ball to his urn, in which case, a single ball will be drawn from the urn (X,Y,Z).

Notice that if *two* balls are randomly drawn (with replacement) from any given urn, then an urn with only *two* distinct balls and an urn with *three* distinct balls, both induce exactly the same probability distribution over outcomes: a $\frac{1}{6}$ chance of winning \$60, a $\frac{1}{3}$ chance of winning \$30, and a $\frac{1}{2}$ chance of winning \$0. This means that an expected-utility maximzer, who is risk-averse, *and* believes that two balls will

HF (phase 2)	subjects	paid	%
session 1	16	9	56%
session 2	11	22	50%
session 3	21	12	57%
session 4	17	6	35%
Total:	76	**38**	
%		**50%**	

Table 8: Choices in the Halevy-Feltkamp Treatment

be randomly drawn (with replacement) from his urn, will be indifferent between an urn with two distinct balls and an urn with three distinct balls.[11] Therefore, the Halevy-Feltkamp theory cannot rationalize the behavior of a subject who pays to add a third distinct ball to his urn. This experiment was also conducted using the cross-over design of the TWEG treatment however, unlike in the TWEG treatment where the first phase is used to increase the power of the test, this time the first phase (which will be described in the next section) was used as a way to acquire more data (as opposed to a within subject control).[12]

As table 8 indicates, on average 50% of subjects elected to pay to add a third ball to their urn. This frequency is not statistically different from what is observed in the baseline (fisher exact test p-value > 0.1). These results seem to suggest that we may not be able to explain the diversification bias of our subjects using the original Halevy-Feltkamp model.

7 Is the magnitude of the bias sensitive to the number of distinct balls?

An interesting question that arises from our findings is whether the bias we identify is sensitive to the number of distinct balls. In particular, is the tendency to pay for a diversified urn stronger when this urn contains more distinct balls? To investigate this question, we conducted both the baseline treatment and the ROCL treatment with only two letters, $\subset X$ and Y, where one and only one of these letter represented a prize of $20 (so that expected value is still $10 as in the corresponding three-ball treatments). Half of the subjects started with (X, X), while the remaining half started with (Y, Y). Subjects could keep their assigned urn or pay a $1 to have an urn with

[11] Actually, this is true independent of the decision makers' risk attitudes and the number of balls drawn.

[12] In this case 94% of subjects elected to move to the second phase.

(X, Y). Our results are summarized in Table 6.

2 Balls	Baseline		ROCL	
(phase 1)	Total	Paid \$1	Total	Paid \$1
session 1	17	5	22	1
session 2	22	9	20	3
Total:	39	14	**42**	4
%		**36%**		10%

Table 9: Choices in the 2 Balls Treatments

Despite the relatively small sample, it is evident from this table that there is still a bias towards the diversified urn, only its magnitude is smaller. The difference between the baseline and ROCL treatment is still statistically significant (Fisher exact test p-value < 0.01). The percentage of paying subjects in each of the two-ball treatments is lower than the percentage in any of the corresponding three-ball sessions. This suggests that individuals may be more inclined to pay for a diversified urn when this urn contain a higher number of distinct balls. In other words, the "don't-put-all-your-eggs-in-one-basket" bias may be stronger the more distinct "baskets" there are.

8 Discussion

8.1 Case-based decisions

Of the various models we examined, only a theory of source-dependence or issue-preference seems most capable to accommodate the bias we document. However, such a model does not explain why one source of risk is preferred to another. In addition, it does not address the relation between the strength of the bias (i.e., the willingness-to-pay for a diversified urn) and the number of states of nature (i.e., how many types of balls there are).

We propose an alternative source for our subjects' behavior. A subject may view the decision task in our experiment as an investment decision, which requires him to choose the optimal portfolio of "balls". Viewed this way, the task may appear similar to a problem of allocating a fixed resource - funds, time, effort - across several random variables. Consequently, a subject may recall past instances in which he faced problems of this nature. These may include past savings and investment decisions and even decisions such as where to apply for college/grad-school/job. In most of these decisions several random variables may yield a favorable return, hence, generally the optimal strategy is to diversify (apply to several schools, apply to several job positions, invest in several funds). In addition, there are many cases where the prior probability that only one particular option would yield the highest reward is smaller the more options there are. Thus, the incentive to diversify often increases with the number of options.[13]

Gilboa and Schmeidler (1995, 2001) formalized a model that captures this idea of case-based decisions. Their model proposes a set of axioms that give rise to a representation that evaluates a decision according to the sum of utility levels that resulted from using this decision in past cases, each weighted by the similarity of the past case to the problem at hand. One complication that arises in directly applying Gilboa and Schmeidler's Case-Based Decision Theory (CBDT) to our set-up is that the precise set of actions available to our subjects - which urn or lottery on urns to choose and how much to pay - is very different than what they encountered in previous cases. A possible solution might be to modify the CBDT framework by introducing a similarity relation over acts or decision rules. This may allow to consider a choice of a diversified urn as being similar to a decision rule of "spreading one's eggs across several baskets", which may have generated relatively high payoffs in similar problems in the past.

The possible relation between the bias we document and CBDT raises the question of whether other well-known biases and departures from expected utility theory may also stem from case-based decisions. Recall, for example, the finding by Halevy (2007) that individuals weakly prefer a bet from an urn having 5 red balls and 5 black balls to a bet from an urn that has either 10 red balls or 10 black balls. As discussed in Section 3, this preference is consistent with an REU model with second-degree risk aversion. CBDT provides an alternative explanation. To see why, note that the act of choosing an urn with balls of only one color may appear to be similar to a strategy of "putting all your eggs in one basket". Hence, the reported preference may result from adopting a decision rule that advises against such a strategy. An interesting direction for future research is to design a methodology for testing these two competing hypotheses.

Admittedly, we do not have any evidence in support of CBDT. Hence, our claim that this may be an explanation remains a conjecture at this stage. It remains a challenge to design a treatment that attempts to test this conjecture.

[13] An alternative heuristic, suggested to us by Steve Salant, is: "when exploring, cast the net broadly" since some of the things you find may later help you to make the right decision (think, for example, of collecting samples from crime scenes or collecting observations for an empirical work). Translated to our context, this heuristic may lead subjects to gather as many potential winning balls as possible. However, this is an inappropriate use of the heuristic since subjects in our experiment do not have the opportunity to learn anything about the identity of the winning ball.

8.2 Eliciting preferences in experiments

As we remarked in the Introduction, our findings raise a concern regarding a common methodology of eliciting preference in laboratory experiments. For a variety of reasons researchers frequently ask subjects to make many decisions, but choose only one of them to be actually carried out. In our experiment, subjects are given an urn with three balls, but only one of them is drawn to determine the payoff (the "Random-Lottery-Incentive-System", or RLIS for short). However, a majority of our subjects exhibited, what may be interpreted as false diversification, by paying for different types of balls. Similarly, when given a questionnaire with different decision problems, subjects may exhibit an analogous form of false diversification by displaying, for example, greater risk/patience in one question and greater caution/impatience in another question. Consequently, one may not be able to interpret the answer in each question as the subject's revealed preference over the possible answers.

While previous studies have already raised concerns about the use of RLIS to elicit preferences, they all based their critique on failure to reduce compound lotteries. As a result, several researchers have suggested various ways in which this critique may not apply. But these suggestions might work as long as the true source of the difficulty with RLIS is indeed failure to reduce compound lotteries. The Case-Based interpretation of our results suggests that in order to keep using this method of preference elicitation, one may need a different form of remedy if, of course, results of this sort extend to how subjects perceive payment schemes in experiments.

References

[1] David E. Bell. 1982. "Regret in Decision Making Under Uncertainty." *Operations Research*, 30(5): 961-981.

[2] Shlomo Benartzi and Richard H. Thaler. 2001. "Naive Diversification Strategies in Retirement Savings Plans." *American Economic Review*, 91(1): 79-98.

[3] Shlomo Benartzi and Richard H. Thaler. 2007. "Heuristics and Biases in Retirement Savings Behavior." *Journal of Economic Perspectives*, 21(3): 81-104.

[4] John Conslik. 1989. "Three Variants on the Allais Example." *American Economic Review*, 91(1): 79-98.

[5] Gary Charnass and Uri Gneezy. In press. "Portfolio Choice and Risk Attitudes: An Experiment." *Economic Inquiry.*

[6] Haluk Ergin and Faruk Gul. 2004. "A Subjective Theory of Compound Lotteries." Mimeo.

[7] Guillaume R. Fréchette, John H. Kagel and Steven F. Lehrer 2003. "Bargaining in Legislatures: An Experimental Investigation of Open versus Closed Amendment Rules." *American Political Science Review*, 97(2): 221-232.

[8] Guillaume R. Fréchette, John H. Kagel and Massimo Morelli 2005. "Behavioral Identification in Coalitional Bargaining: An Experimental Analysis of Demand Bargaining and Alternating Offers" *Econometrica*, 73(6): 1893-1938.

[9] Thomas Gilovich, Dale Griffin and Daniel Kahneman, ed. 2002. *Heuristics and Biases: The Psychology of Intuitive Judgement.* Cambridge University Press: New York.

[10] Yoram Halevy. 2007. "Ellsberg Revisited: An Experimental Study." *Econometrica*, 75(2):503 536.

[11] Yoram Halevy and Vincent Feltkamp. 2005. "A Bayesian Approach to Uncertainty Aversion." *Review of Economic Studies*, 72(2): 449-466.

[12] Takashi Hayashi. 2008. "Regret Aversion and Opportunity Dependence." *Journal of Economic Theory*, 139(1): 242-268,

[13] Itzhak Gilboa and David Schmeidler. 1995. "Case-Based Decision Theory." *Quarterly Journal of Economics*, 110: 605-639.

[14] Itzhak Gilboa and David Schmeidler. 2001. *A Theory of Case-Based Decisions.* Cambridge University Press.

[15] Charles A. Holt. 1986. "Preference reversals and the Independence Axiom." *American Economic Review.* 76: 508-515.

[16] Gur Huberman and Wei Jiang. 2006. "Offering versus Choice in 401(K) Plans: Equity Exposure and Number of Funds." *Journal of Finance*, 61(2): 763-801.

[17] Graham Loomes. 1991. "Evidence of a New Violation of the Independence Axiom." *Journal of Risk and Uncertainty*, 4(1): 91-108.

[18] Graham Loomes and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty." *Economic Journal*, 92(368): 805-824.

[19] Kenneth MacCrimmon. 1965. "An Experimental Study of the Decision Making Behavior of Business Executives." Mimeo.

[20] Peter Klibanoff, Massimo Marinacci and Sujoy Mukerji. 2005. "A Smooth Model of Decision Making Under Ambiguity." *Econometrica*, 73: 1849-1892.

[21] David M. Kreps and Evan L. Porteus. 1978. "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica*, 46: 185-200.

[22] Matthew Rabin and Georg Weizsäcker. In press. "Narrow bracketing and dominated choices." *American Economic Review.*

[23] Ariel Rubinstein. 2002. "Irrational Diversification in Multiple Decision Problems." *European Economic Review*, 46(8): 1369-1378.

[24] Todd Sarver. 2008. "Anticipating Regret: Why Fewer Options May Be Better." *Econometrica*, 76(2): 263-305.

[25] Uzi Segal. 1987. "Two-Stage Lotteries Without the Reduction Axiom." *Econometrica*, 58: 349-377.

[26] Paul Slovic and Amos Tversky. 1974. "Who Accepts Savage's Axioms." *Behavioral Science*, 19: 9-118.

[27] Chris Starmer. 1992. "Testing New Theories of Choice Under Uncertainty using the Common Consequence Effect." *Review of Economic Studies*, 59(4): 813-830.

[28] Chris Starmer. 2000. "Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk." *Journal of Economic Literature*, 38: 332-82.

[29] Chris Starmer and Robert Sugden. 1991. "Does the random-lottery incentive system elicit true preferences? An experimental investigation." *American Economic Review*. 81(4): 971-978.

[30] Robert Sugden. 1993. "An Axiomatic Foundation of Regret." *Journal of Economic Theory*, 60(1): 159-180.

[31] Amos Tversky, Paul Slovic and Daniel Kahneman. 1990. "The Causes of Preference Reversal." *American Economic Review*, 80(1): 204-217.

[32] Amos Tversky and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science*, 211: 453-458.

[33] Amos Tversky and Peter Wakker. 1995. "Risk Attitudes and Decision Weights." *Econometrica*, 63(6): 1255-1280.

Appendix: Baseline Instructions

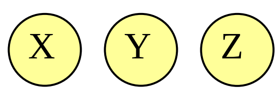
INSTRUCTIONS

Introduction

This is an experiment in decision making. Various research institutes have provided the money for this experiment and you can make a considerable amount of money in a short time if you pay attention.

The decision task

The experiment you will participate in is very simple. Imagine there is an urn with three balls in it. Each ball is marked by one letter, either X, Y or Z:



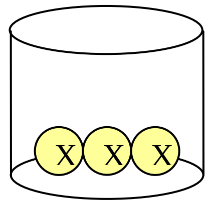
Each letter represents an amount of dollars:
One letter represents the amount **$30**
while the other two letters represents the amount **$0**.

The matching between the letters and the two amounts of money will be randomly chosen by the computer. This means that it is equally likely that X=$30, Y=$0 and Z=$0, or that X=$0, Y=$30 and Z=$0, or that X=$0, Y=$0 and Z=$30. *You will not be told which letter corresponds to which amount of money. This information will be revealed to you only at the end of the experiment.*

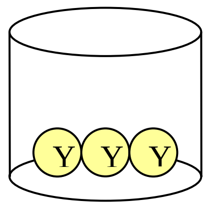
At the beginning of the experiment, you will be assigned an urn containing three balls. At the end of the experiment, the computer will randomly select one of the three balls in your urn (so that each ball has an equal chance of being selected), and you will be paid the value that corresponds to the letter on that ball.

Your assigned urn will have three balls with the same letter.

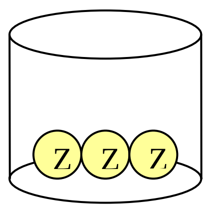
About a third of the participants will be assigned an urn with three X balls:



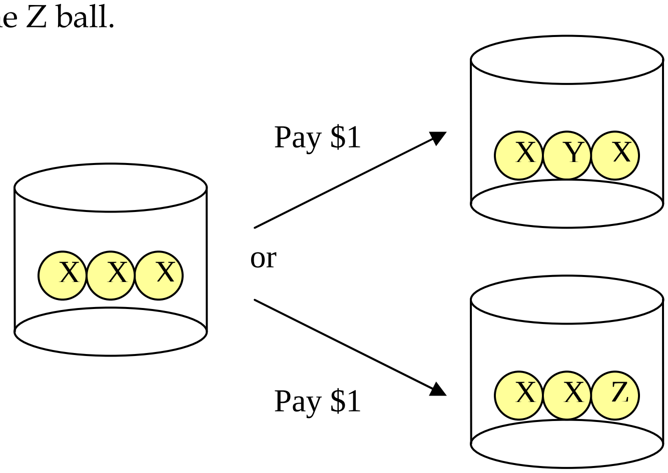
Another third will be assigned an urn with three Y balls:



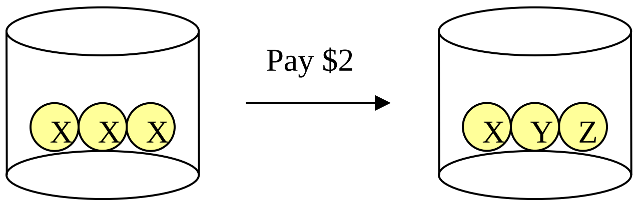
And the remaining third will be assigned an urn with three Z balls:



You may either keep your assigned urn, or you may pay **$1** and exchange one of the balls in your urn for a ball with a different letter of your choice. That is, if you pay **$1** your urn will be changed to include *two balls* of the same letter and *one ball* of a different letter. For example, if you started with three X balls, then you may pay $1 for an urn with two X balls and one Y ball or an urn with two X balls and one Z ball.



You also have the option of paying $2 for exchanging two balls in your original urn such that the resulting urn has one X ball, one Y ball and one Z ball. For example, if you started with three X balls, then you may pay $2 for an urn with one X ball, one Y ball and one Z ball.



After you have made your decision, the computer will randomly select a ball from your urn. At that point, your payoff, including the value of the ball that was drawn, and the cost you incurred, if you decide to switch one or two balls, will be revealed to you.

Payoffs

Your payoff in the experiment will be equal to your show up fee of $8, the amount you obtain from your urn, minus the cost of switching one or two balls, if you have elected to do so. The minimum payoff is $6 and the maximum is $38.

Thanks for your participation!