A Variation on Ellsberg*

Kfir Eliaz[†] and Pietro Ortoleva[‡]

February 13, 2011

Abstract

Ellsberg's experiment involved a gamble with no ambiguity (N) and a gamble where the prize that could be won is objectively known, but the winning probability depends on the (ambiguous) urn's composition (P). We extend this by including a gamble where the winning probability is objectively known, but the prize depends on the urn's composition (C), and also gambles where both the probability and the prize depend on the urn's composition, and can either be correlated positively (D) or negatively (M). Among transitive subjects who prefer N to P, 40% prefer D to N, 74% prefer D to P, 97% prefer D to M, and the modal ranking (about 39%) satisfies D$\succcurlyeq$N$\succcurlyeq$P,C. We show that this behavior is compatible with the Max-Min Expected Utility model if every prior in the set of priors has a high enough variance, a property that we call 'skeptical pessimism.'

JEL Classification: C91, D11, D81.
Keywords: Ellsberg Paradox, Uncertainty Aversion, Ambiguity Aversion, MaxMin Expected Utility.

*A previous version of this paper circulated under the title "Uncertain Probabilities vs. Uncertain Outcomes: An Experimental Study of Attitudes Towards Ambiguity." We thank Marina Agranov, Colin Camerer, Simone Cerreia-Vioglio, Mark Dean, Federico Echenique, Paolo Ghirardato, Yoram Halevy, Fabio Maccheroni, Massimo Marinacci, Jawwad Noor, Efe Ok, Antonio Rangel, Marciano Siniscalchi, Tomasz Strzalecki, Leeat Yariv, and the participants at ESA 2010 for helpful comments. We also thank Jim Campbell and Salvatore Nunnari for their research assistance. Financial assistance from Brown University's Richard B. Salomon Faculty Research Award is gratefully acknowledged.

[†]Department of Economics, Brown University
[‡]Division of the Humanities and Social Sciences, California Institute of Technology

1 Introduction: Extending Ellsberg

In the seminal experiment suggested by Ellsberg (1961), subjects are presented with an urn containing a fixed number of balls, say 60 and are told that 20 of these balls are black, while the remaining ones are either red or green. However, they are not informed of their relative proportions – only that they could be between 0 and 40 red and green balls. A single ball will be randomly drawn from the urn, and the subjects are asked to choose between three gambles which pay some prize, say \$20, if the ball drawn is of a particular color, and nothing otherwise. In one gamble the "winning" color is black, in another it is red, and in the third it is green. The typical finding is that a majority of the subjects choose the gamble in which the winning color is black, i.e. they prefer the option for which the number of balls is known, the so-called unambiguous option.[1] It is well known that if we interpret this choice as a strict preference to bet on black rather than to bet on any of the other colors, then this behavior is *not* consistent with expected utility maximization using a belief compatible with the information provided.[2] This behavior is usually referred to as ambiguity aversion.

In the Ellsberg experiment when a subject chooses to bet on 'red' or 'green', her probability of winning is uncertain – because she does not know the number of red or green balls – while the amount that she could win is known to her – \$20. Put differently, the Ellsberg experiment tests the choice between options in which the *probability* of winning depends on the color of a ball extracted from an ambiguous urn, but the *outcome* - the amount of the prize - is known. In this paper we extend the original Ellsberg experiment by introducing gambles in which the composition of the ambiguous urn affects either the likelihood of winning (like Ellsberg), or the amount that may be won, or both.

We have three goals in mind. The first goal is empirical in nature: do individuals approach these various forms of dependence on the urn composition in the same way? Suppose an individual prefers a gamble with no ambiguity, where neither the odds of winning nor the amount of the prize depend on the composition of the urn, to a gamble where only the odds depend on the urn's composition. Would this individual also choose the non-ambiguous gamble over one where *both* the odds of winning and the prize depend on the urn's composition? We address these questions in a laboratory setting, and we show that a large fraction of subjects does *not*.

Our second goal is to use the behavior that we observe in this richer domain to derive additional restrictions for the theoretical models used for choice under uncer-

[1] See, amongst many, the survey in Camerer (1995), the recent experimental investigation of Halevy (2007), and the references therein.

[2] To see why, assume instead that the agent is an Expected Utility maximizer with a prior π over which ball will be extracted. Since must have $\pi(\text{black}) = \frac{1}{3}$, and since the agent strictly prefers betting on black than on red or green, then it must be the case that both $\pi(\text{red})$ and $\pi(\text{green})$ are smaller than $\frac{1}{3}$. But this means that $\pi(\text{black}) + \pi(\text{red}) + \pi(\text{green}) < 1$, which means that π is cannot be a probability distribution over $\{\text{black, red, green}\}$.

tainty. This is motivated by the observation that, while many of these models are general enough to allow for a wide range of behaviors, the cost of this generality is that they have a limited predicting power. To see why, consider one of the most well-known of these models, the MaxMin Expected Utility Model (MMEU) of Gilboa and Schmeidler (1989). According to it, the agent has a *set* Π of prior beliefs over the state space, and evaluates each option using the expected utility computed with the most pessimistic belief in the set Π. The model contains no prescriptions on the properties of this set of beliefs (except that it must be compact and convex), offering therefore limited restrictions on the predicted behavior. We will argue that the empirical evidence that we collect in our experiment could be used to derive novel restrictions about which priors should belong to this set, and which priors should not. That is, we do *not* use our experimental data to suggest that we should generalize the model. Rather, we use our data to suggest that we should focus on a *special case* of it. This would naturally increase its predictive power.

Finally, we will use our empirical findings to discuss some of possible *interpretations* of ambiguity-aversion or Ellsberg-type behavior that are at times suggested.

In our experiment subjects face the three-color Ellsberg urn described above (60 balls, 20 of which are Black, the others are Red or Green). The typical Ellsberg experiment asks subjects to rank the following two options:

N : If a black ball is drawn, the prize is \$20. Otherwise, the prize is \$0.

P : If a red (green) ball is drawn, the prize is \$20. Otherwise, the prize is \$0.

Option N (for ‘No ambiguity’) is the typical gamble without ambiguity, while P (for ‘Probabilities’) is gamble in which the probability of winning is ambiguous. In addition to the comparison between N and P, in our experiment subjects are asked questions involving also the following gambles:

C : If a black ball is drawn, the prize equals the number of red (green) balls. Otherwise, the prize is \$0.

D : If a red (green) ball is drawn, the prize equals the number of red (green) balls. Otherwise, the prize is \$0.

M : If a red (green) ball is drawn, the prize equals the number of green (red) balls. Otherwise, the prize is \$0.

In option C (for ‘Composition’) the amount of the prize depends directly on the composition of the urn, while the probability of winning does not; in D (for ‘Double dependence’) *both* outcomes and probabilities depend on the urn’s composition, and both values are (perfectly) *positively* correlated; in M (for ‘Mixed dependence’) *both* the outcomes and probabilities depend on the urn’s composition, but here both values are (perfectly) *negatively* correlated. In each question, each type of gamble with some form of ambiguity (P,C,D,M) is presented to the subjects always in two variants, one in which the 'winning' color is red, and the symmetric one in which the 'winning' color is green. Since our treatment is entirely symmetric, and since in our data subjects do not seem to treat red and green differently, in our discussion we refer to the generic option. Subjects are asked all pairwise comparisons between N, P, C, and D, and are also asked to compare M and D, and M and N. In addition, subjects answer questions meant to elicit their attitudes toward risk and to examine the consistency of their answers.

We recruited 108 subjects and obtained the following results. First, 80 subjects (74%) express a transitive ranking, and our analysis focuses on them.[3] Second, 84% (67 out of 80) exhibit the standard Ellsberg behavior, i.e. they rank N above P. An almost identical proportion is found if we look at the ranking of N and C: 85% prefer N to C (68 out of 80). That is, we find that subjects exhibit a similar attitude towards having either the winning probability or the amount of the prize (but not both) depend on the unobserved composition of the urn. However, when we also consider D things change considerably: only 50% of the subjects choose N over D (40 out of 80). That is, half of the subjects are *not* averse to having both the odds of winning and the amount of winnings depend on the composition of an ambiguous urn.

To further investigate the attitude towards D, because we elicit all binary comparisons between the options above, if we focus on transitive subjects we can classify subjects who are ambiguity averse in the standard sense ($N \succcurlyeq P$) into three types:

1) those who rank D at the bottom ($N \succcurlyeq P \succcurlyeq D$);
2) those who rank D at the top, i.e. subjects who are ambiguity averse in the standard sense but prefer the option in which both prizes and odds depend on the urn composition to the one with no ambiguity ($D \succcurlyeq N \succcurlyeq P$); and
3) subjects who choose N over D, but who prefer to have both the prize and the probability depend on the urn's composition than to have only the probability depend on it ($N \succcurlyeq D \succcurlyeq P$).

In our dataset, amongst the subjects who prefer N to P, only 26% rank D at the bottom (18 out of 68). Instead, 40% rank D at the top, while 34% rank it in the middle. This means that 40% of subjects who exhibit standard Ellsberg-behavior also rank D above N, and a total of 74% rank D at least as high as P.[4] Furthermore, if we partition our population into types according to their preference between D, N, C, and P, then the most frequent types are the subjects who rank D as the best option,

[3] Since our analysis is not able to identify indifferences, non-transitive answers are not necessarily 'irrational', but could be simply due to the fact that subjects break indifferences in a way which is not consistent across the questions.

[4] While in the comparison between N and P the risk attitude of the agents does not matter, this is no longer true for the ranking between N and D. But even if we focus only on (weakly) risk averse subjects, our results remain unchanged: 40% rank D at the top, 34% in the middle, and 26% at the bottom. That is, these results do not seem to be connected with the agent's risk attitude.

then N, and then C or P (they represent 33% of the total population, 39% of the ambiguity averse one). We should also emphasize that, since the questionnaire includes all pairwise comparisons and since we are focusing only on transitive answers, then a subject who ranks D at the top must have chosen D against N, P, C, and M. That is, to rank D on top they must choose D not only once, but consistently across questions.

Coherently with the results above, subjects seem to uniformly *dislike* M: amongst the transitive ones, 90% rank N above M (72 out of 80), and of the subset of the subjects who are asked to compare D and M, 97% prefers D (34 out of 35). That is, almost all subjects prefer the option in which the two unknown values (winning probability and prize) are (perfectly) positively correlated (D), to the option in which they are (perfectly) negatively correlated (M).

Finally, subjects tend to compare P and C with N in a similar way, and to be almost evenly distributed in their preferences between the two: 53% of the subject prefer P to C (42 out of 80). We also find significant gender effects: the attraction towards D seems to be much stronger for men rather than women.

We then turn to investigate the consequences of our findings in light of well known models of ambiguity aversion. Focusing on the MaxMin Expected Utility (MMEU) model, we show that the behavior we document *is compatible* with this model, but it suggests restrictions on the set of agents' priors Π. To illustrate this, consider first of all the typical 'classroom explanation' of the Ellsberg behavior with the MMEU model: Suppose the decision-maker holds a *set* of priors that includes the belief that there are only 15 red balls in the urn, as well as the belief that there are only 15 green balls. Then, if the agent evaluates each option using the worst possible prior in the set, she should strictly prefer betting on red or betting on green to betting on black. She would then rank N above P. However, for the same reason she should also rank N above D, and even P above D – in contrast with our findings. That is, the set of priors of the typical 'example' of MMEU preferences is not compatible with our observations.

At the same time, the MMEU model can generate the ranking D$\succcurlyeq$N$\succcurlyeq$P. However, this happens if, and only if, *every* prior in Π has a 'high' variance on the number of red balls.[5] Intuitively, D is an option in which the probability of winning and the amount that is won are positively correlated. In particular, the expected gains from D increase with the number of red (green) balls in a *convex* fashion. This means that, if the agent's belief on the number of red balls has a high variance, then the expected utility of D computed for each of these priors is relatively high. We can characterize this condition precisely: a risk-neutral MMEU subject exhibits D$\succcurlyeq$ N$\succcurlyeq$P if, and only if,

$$\mathrm{VAR}_\pi \geq 20^2 - \mathbb{E}_\pi[x]^2 \quad \forall \pi \in \Pi \tag{1}$$

(where $\mathbb{E}_\pi[x]$ and VAR_π are, respectively, the expectation and the variance of the

[5] That is, if we look at the marginal belief that each prior has on the number of red balls, it must have a relatively 'high' variance.

marginal of π over the number of red balls). That is, each prior in the set of priors must have a variance above a threshold, where this threshold is *higher* the *lower* the expected value of the prior.[6] In particular, if the agent has a 'pessimistic prior,' with an expected value below 20, as ambiguity aversion would imply, then this prior must also has a relatively 'high' variance – the higher the further below 20 the expected value is.

Condition (1) may be interpreted as follows. On the one hand, a decision-maker who is concerned about the lack of objective information when evaluating an act may contemplate the "worst case scenario," i.e. an extremely pessimistic prior belief. On the other hand, even when she is pessimistic, this decision-maker should consider the fact that 'she doesn't really know,' and she should also considers the possibility that 'she is wrong in being pessimistic.' In fact, Condition (1) implies that there is *no* π in Π with $\mathbb{E}_\pi < 20$ and $\pi(\{21, \ldots, 40\}) = 0$: even when the agent acts as pessimist ($\mathbb{E}_\pi < 20$), she should always allow for the possibility that there are *more* than 20 red balls in the urn ($\pi(\{21, \ldots, 40\}) > 0$). This is a decision maker who is never sure of her pessimism, and who acknowledges her lack of precise information by including some variance in her beliefs. We call this behavior "skeptical pessimism."[7] We also discuss similar requirements for another well-known model of ambiguity aversion, Second Order Expected-Utility: in this case the decision-maker would express her "skepticism" by putting low weight on pessimistic beliefs with low variance.

We conclude this introduction with a small discussion on the compatibility of our empirical findings with some of the interpretations of ambiguity-aversion, and of the typical Ellsberg behavior, that are sometimes informally suggested. In fact, while the behavior of our subjects is compatible with existing representations, it raises questions as to in which sense it represents a true 'aversion to ambiguity:' at the end of the day, a large fraction of our subjects do prefer a gamble with a fair amount of ambiguity, D, to a gamble with no ambiguity, N. In fact, ambiguity-aversion is at times informally associated with the naïve idea of aversion to incomplete information. That is, consider an uncertain prospect that pays a prize in each state. This prospect can be described by a table that for every state gives information about the likelihood of that state and the amount of the prize. Each piece of information can either be precise - a number - or ambiguous - an interval. This interpretation suggests that if we compare two tables, where one is obtained from the other by changing some of the precise numbers into intervals (containing the original precise numbers), then the decision-maker would prefer the prospect given by the original table. This interpretation appears to be incompatible with D$\succcurlyeq$N$\succcurlyeq$P, the modal ranking in our experiment. In fact, in some sense N depends on black just like D depends on red, but while on black we have a precise information, on red we do not, and the gamble D could be seen derived from the gamble N after we replace both the likelihood of winning, 20/60, and the prize

[6] We also derive the corresponding conditions for D $\succcurlyeq$ P and D $\succcurlyeq$ M, both of which impose minimal bounds on the variance of the priors in Π and are strictly milder requirement that (1).

[7] We thank Paolo Ghirardato for suggesting this term.

that could be won, 20$, with an interval that represents the number of red balls. This does not seem to be compatible with the interpretation above, although, as we have seen, it *is* compatible with some of the most well-known models of ambiguity aversion.[8]

The remainder of the paper is organized as follows. Section 2 describes the experimental design and results. Section 3 presents the restrictions that the observed data impose on the MaxMin Expected Utility model, and briefly discuss similar restrictions to other theoretical models. Section 4 concludes.

2 The laboratory experiment

2.1 Design

The experiments were conducted in the Social Science Experimental Laboratory (SSEL) at the California Institute of Technology. Subjects were undergraduate students at Caltech, recruited from a pool of volunteer subjects, maintained by SSEL. There were a total of four sessions, with a total of 108 subjects. No subject participated in more than one session. Each of the sessions were conducted using the following procedure. The subjects were handed two packets, one containing instructions and another containing a list of questions. (Appendix A contains a copy of the instructions, and Appendix B contains one example of a questionnaire.) The experimenter stood in front of the subjects with a non-see-through bag that contained 60 poker chips. The subjects were told that 20 of the chips are black, r chips are red, and g chips are green, where $r + g = 40$. It was emphasized to the subjects that they were not told how the values of r and g were determined. Subjects were also told that they could inspect the content of the bag at the end of the experiment.[9] In all four sessions, the questionnaire included questions in which subjects was asked to *choose* a single lottery from a list containing two or more lotteries.[10] All questions

[8]Another well-known interpretation for ambiguity aversion relates to issue preferences. As suggested in Ergin and Gul (2009), subjects might prefer gambles that depend only on one uncertain issue against others that depend on two uncertain issues. Indeed, in our case the gamble D depends on two issues (the color of the ball extracted and the composition of the urn), while P depends only on one (the color of the ball extracted), and yet most subjects rank D$\not\succcurlyeq$N$\not\succcurlyeq$P. At the same time, however, in our case the two issues are not independent, but *perfectly correlated* – rendering the intuition very different from the standard case discussed in the literature.

[9]In three of the four sessions this was in fact requested by one or more subjects.

[10]In sessions 1 and 2 the questionnaire also included questions where subjects were asked to assign a monetary value to lotteries, using the Becker-DeGroot-Marchak (BDM) mechanism. However, this procedure proved to be complicated and many subjects seemed not to understand it. (For example, many subjects gave a monetary evaluation of \$20 to the gamble N, which pays \$20 if a black ball is extracted, and \$0 otherwise.) Moreover, it is well-known that the BDM mechanism might not elicit the true ranking in situations of choice under uncertainty (see Karni and Safra (1987)). Finally, pricing objects (and especially artificial lotteries) is a new and foreign task for most of our subjects, while they make choices on a regular basis. This suggests that responses to choice problems may

were handed at the same time, and subjects were allowed to modify their answers. Moreover, the experimenter explicitly advised them to review their choices after they completed the questionnaire, before turning it in. Once all participants have finished answering and reviewing their questionnaires, the experimenter proceeded as follows. Standing in front of the subjects, he first randomly drew a chip from the cloth-bag. Then, using a pair of die, the experimenter drew one question number from the questionnaire. Subjects were then paid according to the amount specified by the lottery they have chosen for the drawn question, plus a show-up fee of $7. (Note that this sometimes involved counting the number of red and green chips in the bag.)

In the questionnaire itself, each of these lotteries was framed as follows: "If BLACK, then $_, if RED then $_ and if GREEN then $_". (The instructions explained to subjects what we mean by the term "lottery.") The ordering of the options in each question was of two kinds. For approximately half of the subjects, the options were listed in a descending order of the amount of objective information provided, while this ordering was reversed for the remaining subjects.[11] In each question subjects are asked to choose only one option, but are not allowed to express the strength of such preference. We are therefore unable to separate the cases of indifferences from those of strict preference using only our data.[12] At the same time, however, the relatively large fraction of transitive answers suggest that subjects were not simply indifferent between all options and gave random answers.

Subjects answered two types of questions: questions meant to elicit a ranking between the gambles in {N,P,C,D,M}, described above, and questions meant to test for basic properties of their preferences such as risk-aversion and various forms of coherence across choices. In the questions of the first kind, subjects were asked all pairwise comparisons between {N,P,C,D}, to choose between M and N, and, for sessions 3 and 4, also between M and D.

As we mentioned before, each of these options except N were presented to subjects in two forms: with winning color red, or winning color green. For example, in the

be more reliable than responses to pricing questions. For these reasons, our analysis will ignore the answers to the BDM questions. In order to use the choice data from these sections, we test whether the presence of the BDM questions had any significant impact on the answers given, and found that it did not.

[11]That is, for half the subjects, whenever N was available it was listed first, and if P or C were available, each would be listed before D and M. In a question containing both P and C, P appeared first in an odd numbered session and second in an even-numbered session. Similarly, a lottery where a prize is won if a red ball is drawn appeared in an odd-numbered session before a lottery where a prize is won if a green ball is drawn, while the opposite ordering took place in the other sections.

[12]The issue of separating strict and weak preferences in a framework with ambiguity is well known, and to our knowledge no easy solutions have been suggested. In particular, it is well-known that the usual technique according to which subjects choose both options when indifferent, and then the experimenter randomizes, would not work in the case of uncertainty – being ambiguity averse, subjects have a strict preference for hedging, and would value the external randomization even when they are not indifferent. Also, as discussed in footnote 6, eliciting monetary valuations via the BDM method is problematic.

question that asks subjects to compare N with P, subjects are confronted with three options:

Question 1. Which lottery do you prefer?

If BLACK then \$20, if RED then \$0 and if GREEN then \$0	
If BLACK then \$0, if RED then \$20 and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$20	

We adopted a framework of absolute symmetry between red and green in order to avoid the risk that subjects might make an inference about the composition of the urn from the questions that were asked. We then classify subjects as ranking N above P if they choose the first alternative, while P above N if they choose the second or the third. That is, because each treatment is entirely symmetric with respect to red and green, we disregard any difference between red and green in the answers. As discussed below, our data seem to support the fact that subjects were in fact treating red and green indifferently (see Section 2.2.1).

Evidence for the subject's risk attitude was elicited by a question testing whether the subject was risk-loving or not (question 7 on the questionnaire):[13]

Question 7. Which lottery do you prefer?

If BLACK then \$40, if RED then \$0 and if GREEN then \$0	
If BLACK then \$0, if RED then \$20 and if GREEN then \$20	

Since most models of decision-making under uncertainty impose some variant of the sure-thing principle, we also test whether subjects satisfy two forms of coherence using the following three questions (Question 2, 9 and 10 respectively):

Question 2. Which lottery do you prefer?

If BLACK then \$20, if RED then \$0 and if GREEN then \$0.	
If BLACK then \$r, if RED then \$0 and if GREEN then \$0.	
If BLACK then \$g, if RED then \$0 and if GREEN then \$0.	

[13]This question tests for *weak* risk aversion: it separates subjects who are risk neutral or strictly risk averse, from subjects who are strictly risk loving. We are mostly interested in these subjects since in this setup strict risk loving would induce a different behavior, and because risk neutrality plays a special role in our theory analysis, and therefore we would like to allow for it. At the same time, as a robustness check, about half of our subject pool (sessions 3, 4) was also asked a question meant to elicit whether they are *strictly* risk averse: they are asked to choose between a gamble that pays \$40 if a black chip is extracted (\$0 otherwise), versus a gamble that pays \$13 (<40/3) no matter which ball is extracted (Question 8 in the questionnaire). Subjects behaved coherently between the two questions – only 7 subjects gave the inconsistent answer of showing strict risk aversion in one question and strict risk loving in the other. At the same time, however, only about 30% of subjects showed strict risk aversion, and the data suggest that our focus on weak instead of strict risk aversion is inconsequential: the analysis in the paper would be essentially identical if we used the answers to Question 8 instead of Question 7 when we test for the role of the risk attitude. For this reason, in what follows we focus only on the answers to Question 7.

Question 9. Which lottery do you prefer?

If BLACK then \$20, if RED then \$20 and if GREEN then \$20.	
If BLACK then \$r, if RED then \$r and if GREEN then \$r.	
If BLACK then \$g, if RED then \$g and if GREEN then \$g.	

Question 10. Which lottery do you prefer?

If BLACK then \$20, if RED then \$20 and if GREEN then \$20.	
If BLACK then \$r, if RED then \$20 and if GREEN then \$20.	
If BLACK then \$g, if RED then \$20 and if GREEN then \$20.	

A coherence in the answers to the questions above is related to Savage's P2 (Question 2 and 9) and P3 (Questions 2 and 10).[14] At the same time, however, such coherence is naturally much weaker than P2 and P3, and it is in fact satisfied by most of the models that study ambiguity aversion, which violate Savage's postulates.[15] For later discussion, let us also point out that if we were to make the (additional) assumption that the likelihood of a black ball being extracted is considered objectively equal to 1/3, then these two forms of coherence would correspond to the Certainty-Independence condition of Gilboa and Schmeidler (1989): in fact, each of the options in Question 2 could be seen as a mixture between the corresponding option in Question 9 and \$0, with weight $\frac{20}{60}$; and each option in Question 10 can be seen as the same mixture between the corresponding option in Question 9 and \$20, again with weight $\frac{20}{60}$.[16] A subject is said to be "regular" if she gives coherent answers in Questions 2 and 9. A subject is said to be "regular*" if she gives coherent answers in Questions 2, 9, and 10.

2.2 Analysis

2.2.1 Basic Facts

Before we conduct a thorough analysis of the data, we need to obtain some evidence on the extent to which subjects gave consistent answers. We begin by checking how many subjects were transitive. Out of the 108 subjects, 80 (74%) do not violate transitivity. (We refer to the this group as Transitive in the subsequent analysis.) We should emphasize that transitivity is a relatively demanding requirement in this

[14] By 'coherent' we understand that if subjects choose the first alternative in Question 2, they should also choose the first alternative in the latter questions. Conversely, if they choose the second or the third alternatives in Question 2, they should choose again either the second or the third option in the other questions.

[15] For example, both forms of coherence would be compatible with $\bar{S}$-act-independence Axiom of Casadesus-Masanell et al. (2000), and would in fact be satisfied by their representation.

[16] In turns, the coherence between the answers in Questions 2 and 10 would correspond to the Weak Certainty-Independence condition of Maccheroni et al. (2006).

context, as it involves a coherent answer to all the pairwise comparisons above.[17] One potential explanation for the relatively high fraction of transitive subjects may be that subjects answered questions in pen and paper, and could - and were explicitly advised to - review their answers after they were done.[18]

As a further test of consistency, we examined how many subjects are regular and regular*. We find that 86 (80%) of the 108 subjects are regular, of which 71 (89% of the transitive subjects; 66% of the total pool) are also transitive. 55 subjects are regular* (69% of transitive subjects; 51% of the total pool), and all of them are also transitive.

We also test both for effects of the order in which options appear in each question, and for the effects of the presence of different questions in the last two sessions using the Kolmogorov-Smirnov (K-S) test.[19] We find no such effect both for the general population and for transitive subjects.[20] To gain some preliminary evidence on the extent to which subjects actually deliberated on the questions, we test whether in each question the distribution of answers is significantly different from the uniform distribution (over possible answers), and find that it is. (The K-S test yielded p values of 0.0000 for all questions.) Together with the relative high fraction of transitive subjects, we understand this as a signal that subjects were not answering randomly.

As discussed above, our analysis ignores the distinction between the two types of lotteries in which D,P,C and M are presented – the ones involving red and those involving green. The implicit assumption is that subjects are indifferent between having the prize, or the likelihood, or both, depend on the number of red balls or on the number of green balls, as the description of the problem would suggest. To verify the validity of this assumption, we test whether there are significant differences between the proportion of subjects who chose to bet on red and the proportion of subjects who chose to bet on green in each of the relevant questions (i.e. excluding the questions where there is no distinction between red or green), and we find that there are not (the p-values of the K-S tests are 0.0000 for all questions).

[17]Moreover, as we mentioned in the introduction, a violation of transitivity need not represent a violation of 'rationality,' but could simply be due the fact that subjects break indifferences in different way depending on the questions.

[18]This could be naturally problematic if were we were aiming to test transitivity. But of course our goal here is the opposite: assuming the existence of a underlying transitive ranking between the available options, we are trying to find the procedure that best elicits it.

[19]Recall that in Sessions 1 and 2 we ask subjects for their monetary evaluations using the BDM mechanism. We need to test that the presence of these questions did not affect the results.

[20]More precisely, we ran a K-S test to compare the distribution of responses, both on all subjects and also on the restricted sample of transitive subjects. In both cases, the test yielded that at the 1% level we cannot reject the null hypothesis that the two distributions are the same. The one exception is the question in which subjects were asked to choose between P and C. For that question, the null hypothesis of equal distributions is rejected at the 10% but not at the 5% level.

2.2.2 Distribution of Rankings

We start by classifying subjects according to their attitude towards risk and towards standard ambiguity (N vs. P), and compare these proportions with the typical findings in the literature. A subject is said to be *weakly Risk Averse* (wRA) if she choses the gamble that pays $20 if red or green ball is drawn over the gamble that pays $40 if black is drawn (Question 7 above). We say that a subject is *weakly Ambiguity Averse* (wAA) if she exhibits the classical Ellsberg behavior of ranking N above P. Table 1 below depicts the distribution of subjects according to attitude towards risk and uncertainty using the definitions above.

Table 1: Percentage of Subjects Ambiguity Averse in Probabilities and Risk Averse

	Total (108)		*Transitive* (80)		*Regular* (71)		*Regular** (55)	
wRA	72	67%	55	69%	49	69%	39	71%
wAA	85	79%	67	84%	60	85%	47	85%
wRA & wAA	58	54%	47	59%	43	61%	37	67%

The data above seem to be in line with previous studies. First, a large majority of subjects prefer N to P: between 79% and 85% depending on the subject pool we look at. Second, a slightly smaller number, but still a majority, exhibit (weak) risk aversion: between 67% and 71%. Third, more than half the subjects (54%-67%) show both features at the same time. That is, in line with previous studies, both tendencies seem to be widespread in the subject pool and to coexist in many subjects.

We now turn to analyze the ranking that transitive subjects give to the options N, D, P, and C. We can analyze the distribution across all possible 24 rankings of these 4 alternatives. (The relative ranking of M will be discussed later, since we do not elicit it against all possible options.)

We start by focusing on subjects who rank N above P – the standard Ellsberg-behavior. There are 12 possible rankings with N better than P: Table 2 presents the relative distribution, according to different consistency requirements: transitivity, regularity and regularity*. (Notice that the table represents only the types that were encountered in the data.)

A few observations stand out:

- The largest type, in each of the subgroups, is D$\succcurlyeq$N$\succcurlyeq$P$\succcurlyeq$C. In each subgroup this type makes up about a quarter of the subjects.
- About 40% of the subjects who exhibit the typical Ellsberg behavior of choosing N over P, also choose D over N. That is, many subjects seem to dislike having the probability of winning depend on the composition of the urn (P), but have

Table 2: Distribution of Rankings for Ambiguity Averse subjects

Type	*Transitive*		*Regular*		*Regular**	
$N \succcurlyeq P \succcurlyeq C \succcurlyeq D$	5	7%	5	8%	3	6%
$N \succcurlyeq C \succcurlyeq P \succcurlyeq D$	9	13%	9	15%	7	15%
$N \succcurlyeq D \succcurlyeq C \succcurlyeq P$	6	9%	6	10%	4	9%
$N \succcurlyeq D \succcurlyeq P \succcurlyeq C$	11	16%	9	15%	8	17%
$N \succcurlyeq C \succcurlyeq D \succcurlyeq P$	6	9%	6	10%	5	11%
$N \succcurlyeq P \succcurlyeq D \succcurlyeq C$	2	3%	2	3%	2	4%
$D \succcurlyeq N \succcurlyeq P \succcurlyeq C$	16	24%	15	25%	12	26%
$D \succcurlyeq N \succcurlyeq C \succcurlyeq P$	10	15%	7	12%	6	13%
$D \succcurlyeq C \succcurlyeq N \succcurlyeq P$	1	1%	1	2%	0	0%
$C \succcurlyeq N \succcurlyeq P \succcurlyeq D$	1	1%	0	0%	0	0%
Total:	67	100%	60	100%	47	100%

the opposite attitude when the urn composition affects both the probability and the amount won (D).

- Only 1 subject (1%) amongst those who rank N above P prefers C to N. That is, subjects seems to treat the choice between N and C the same way they treat the choice between N and P.

- Almost all subjects (more than 90%) prefer D to either C or P. A large majority prefers D to P (74%), or D to *both* C and P (66%).

- About 90% of the subjects prefer C or D to P.

- The preference between C and P is fairly evenly divided: 34 subjects prefer P to C, 33 subjects prefer C to P. (It is 25 and 22 for regular* subjects.)

- The preference between N and D, and between P and C, do not seem to be independent. If we focus on the subjects to prefer C to P (33), then 22 (66% of the 33) prefer N to D, while only 11 prefer D to N. Conversely, if we look at the subjects who prefer D to N (27), 16 (59% of the 27) also prefer P to C, while only 11 prefer C to P.

To summarize, there are two main findings in our data. First, subjects seem to treat gambles where only the prize amount depends on the urn's composition (C) in a similar way in which they treat gambles where only the probability of winning depends on the composition (P): almost no subject who ranks N above P ranks C above N; and the ranking between P and C is evenly distributed in the population. Second, subjects seem to treat 'double, correlated dependence' (D) differently than how they treat either form of dependence in isolation: 40% of subjects who rank N above P and C, also rank D above N; almost 74% rank D above P; and 90% rank D above P *or* C. We should emphasize again that for a transitive subject to rank D at the top, she needs to choose D not only once, but in all questions in which it is compared to N, P, and C.

The fact that many subjects prefer "double dependence" to "prize/probability dependence" and even to "no dependence," raises the question of whether these results are driven by the presence of risk-loving subjects. In fact, while risk attitude plays no role in the comparison between N and P – only the attitude towards ambiguity matters – this is no longer true when the comparison involves D or C: both options could, at least theoretically, return any amount in the $\{\$0, \ldots \$40\}$ range. To address this, we perform an identical analysis focusing on subjects who are (weakly) risk averse, as defined above. The results appear in Table 3.[21]

Table 3: Distributions of Rankings for (Weakly) Risk and Ambiguity Averse subjects

Type	*Transitive & wRA*		*Regular & wRA*		*Regular* & wRA*	
$N \succcurlyeq P \succcurlyeq C \succcurlyeq D$	3	6%	3	7%	2	5%
$N \succcurlyeq C \succcurlyeq P \succcurlyeq D$	6	13%	6	14%	5	14%
$N \succcurlyeq D \succcurlyeq C \succcurlyeq P$	2	4%	2	5%	1	3%
$N \succcurlyeq D \succcurlyeq P \succcurlyeq C$	9	19%	8	19%	8	22%
$N \succcurlyeq C \succcurlyeq D \succcurlyeq P$	5	11%	5	12%	4	11%
$N \succcurlyeq P \succcurlyeq D \succcurlyeq C$	2	4%	2	5%	2	5%
$D \succcurlyeq N \succcurlyeq P \succcurlyeq C$	12	26%	11	26%	9	24%
$D \succcurlyeq N \succcurlyeq C \succcurlyeq P$	7	15%	6	14%	6	16%
$C \succcurlyeq N \succcurlyeq P \succcurlyeq D$	1	2%	0	0%	0	0%
Total:	47	100%	43	100%	37	100%

As evident from Table 3, the proportion of types stays roughly the same when we restrict attention only to the (weakly) risk-averse subjects: for example, the fraction of subjects who rank D above N remains 40%. This suggests that *the main force driving the ranking above is indeed the attitude towards ambiguity rather than the attitude towards risk*.

Next, we analyze the subjects who rank P above N – those who would usually be classified as (weakly) ambiguity-loving. Table 4 displays the distribution of these types.

[21] As we mentioned before, about half of our subjects were also asked a question meant to elicit *strict* risk aversion. The conclusions below would remain identical if we focused on the subjects who showed strict risk aversion, instead of focusing on weakly risk averse ones as we do here.

Table 4: Distributions of Rankings for Ambiguity Loving subjects

Type	*Transitive*		*Regular*		*Regular**	
$P \succcurlyeq D \succcurlyeq N \succcurlyeq C$	1	8%	1	9%	1	13%
$P \succcurlyeq D \succcurlyeq C \succcurlyeq N$	1	8%	1	9%	1	13%
$D \succcurlyeq C \succcurlyeq P \succcurlyeq N$	4	31%	4	36%	4	50%
$C \succcurlyeq D \succcurlyeq P \succcurlyeq N$	1	8%	0	0%	0	0%
$D \succcurlyeq P \succcurlyeq N \succcurlyeq C$	2	15%	2	18%	0	0%
$D \succcurlyeq P \succcurlyeq C \succcurlyeq N$	4	31%	3	27%	2	25%
Total:	13	100%	11	100%	8	100%

First, note that very few subjects in our sample belong to this group: only 13 out of a total of 80 transitive subjects (about 16%). Second, by far the most common types in this group are the subjects who rank D at the top and N at the bottom. Together they represent 62% of the subjects who prefer P to N (75% if we look at regular*). In particular, if we focus on regular and regular* subjects, the most common types are those attracted to gambles where the prize amount depends on the urn's composition, as expressed by the ranking D$\succcurlyeq$C$\succcurlyeq$P$\succcurlyeq$N. Notice also that almost half of the these subjects are not (weakly) risk-averse.

Finally, we turn to discuss how subjects rank M against the other options. In the experiment we did not elicit the relative ranking of M against every possible alternative. Rather, all subjects were asked to compare M against N, and about half of the subjects (sessions 3 and 4, 49 subjects in total) were also asked to compare M and D. The results are presented in Table 5 and 6.

Table 5: Ranking of M and N

Type	*Total*		*Transitive*		*Regular*		*Regular**	
M $\succcurlyeq$ N	15	14%	8	10%	7	10%	4	7%
N $\succcurlyeq$ M	93	86%	72	90%	64	90%	51	93%
Total:	108	100%	80	100%	71	100%	55	100%

The results are very clear: Almost all subjects prefer N to M and D to M. That is, when asked to choose between two gambles, one in which both the prize amount and the probability of winning it are perfectly positively correlated (D), and the symmetric one in they are perfectly negatively correlated (M), almost all subjects prefer the former.

Table 6: Ranking of M and N

Type	*Total*		*Transitive*		*Regular*		*Regular**	
M $\succcurlyeq$ D	3	6%	1	3%	1	3%	1	4%
D $\succcurlyeq$ M	46	94%	34	97%	29	97%	23	96%
Total:	49	100%	35	100%	30	100%	24	100%

2.3 Gender Effects

We conclude our analysis of the experimental data by examining gender differences in behavior. Our subjects were asked to specify their gender at the beginning of the questionnaire, and our subject pool included 39 women (36%). We start by noticing that we find no significant difference in the behavior of male and female subjects in the standard Ellsberg questions: the proportion of subjects who rank N$\succcurlyeq$P is 77% for women, and 80% for men (85% and 83% if we focus on transitive subjects). Mild differences appear in risk aversion: 72% of women are (weakly) risk averse, against 64% of men, which become 81% and 63% if we focus on transitive subjects.

Gender effect, however, are much more significant if we look at the ranking of the options introduced by our experiment. In particular, let us focus on transitive subjects who rank N$\succcurlyeq$P, and look at the ranking of N, P, and D, and of P and C. Results appear in Tables 7 and 8.

Table 7: Gender Effects on the ranking of N, P, D

Type	*M*		*F*		*Total*	
N $\succcurlyeq$ P $\succcurlyeq$ D	10	22%	7	32%	17	25%
N $\succcurlyeq$ D $\succcurlyeq$ P	14	31%	9	41%	23	34%
D $\succcurlyeq$ N $\succcurlyeq$ P	21	47%	6	27%	27	40%
Total:	45	100%	22	100%	67	100%

Two features stand out. First, men seem to rank D at the top more frequently than women: only 27% of women rank it at the top versus 47% of men. That is, our results on the special role played by D seem stronger when we focus on men, and less strong in women. Also the relative ranking of P and C seems different. While in the aggregate data we showed that subjects are evenly split between the two options, when we divide across genders we find that, in fact, a small majority of men seem to choose C over P, while a (larger) majority of women choose P over C.[22]

[22] The difference is even larger when we focus on (weakly) risk averse subjects: in this case 77%

Table 8: Gender Effects on the ranking of P and C

Type	*M*		*F*		*Total*	
P $\succcurlyeq$ C	19	42%	15	68%	34	51%
C $\succcurlyeq$ P	26	58%	7	32%	33	49%
Total:	45	100%	22	100%	67	100%

3 Theory and Data: Compatibility and Restrictions

3.1 The MaxMin Expected Utility model

We now turn to analyze the results of our experiment in light of one of the most well-known models of choice under ambiguity: the MaxMin Expected Utility (MMEU) model of Gilboa and Schmeidler (1989). According to this model, for every state space Ω and set of consequences X, agents are endowed with a utility function u over X and with a convex and compact set of priors Π over Ω such that each available option (act) is evaluated by ambiguity averse agents using a functional of the form

$$V(f) = \min_{\pi \in \Pi} \int_{\Omega} \int_{X} u(f(w)(x)) \mathrm{d}x \ \mathrm{d}\omega.$$

(The min is replaced by a max in case of ambiguity loving.) The usual interpretation is that ambiguity averse (loving) agents do not have a single belief over the states of the world, but rather a whole *set* of them, and they evaluate each option using the most pessimistic (optimistic) prior in the set. We refer to Gilboa and Schmeidler (1989) for an in-depth discussion.

We focus initially on the MMEU model for three reasons. First, it is one of the most common and most-well known model for choice under ambiguity. Second, it is the most *restrictive* model within a large and well-known class of models, which means that (*i*) if the behavior is compatible with MMEU, it is compatible also with any model in this more general class, and (*ii*) if the observed behavior suggests any further restriction for the MMEU model, then there would be further restrictions also for the more general models.[23] Third, one of the postulates that distinguish MMEU

of women choose P over C versus 43% of men. When we focus on (weakly) risk averse subjects, numbers change slightly also for the gender effects of the ranking of N, P, and D: the proportion of men who rank D at the top remains 47%, while that of women rises to 29%.

[23] That is, say that the observed behavior satisfies also some additional property which is not implied by the MMEU model. (In this sense, the observed behavior suggests further restrictions on the MMEU model.) Then, this additional property should also be satisfied by models that generalize MMEU, suggesting restrictions for these models as well.

from other models is C-independence. While we do not test this property directly, as we argued above we do test a property reminiscent of it, and we find that it is satisfied by 69% of the transitive subjects, whom we called regular*. As we have seen, our results seem to remain unchanged when we focus specifically on them.

In order to use the MMEU model to study the behavior in our experiment, we need to specify both the space of consequences X and the state space Ω. A natural candidate for the former is the set of possible money amounts that may be won by the agent (e.g. $X = \{0, \ldots, 40\}$). To specify the latter we need to include all the possible events that are subject to uncertainty: the number of red balls ($\{0, \ldots, 40\}$), and the color of the ball that is drawn ($\{R, G, B\}$). We consider therefore $\Omega = \{0, \ldots, 40\} \times \{R, G, B\}$. At the same time, however, we focus on the specific case in which all beliefs in Π are compatible with the actual composition of the urn in the experiment. In particular, we posit two conditions. First, any prior must assigns probability 1/3 to a black ball being extracted (i.e. to the event $\{B\} \times \{0, \ldots, 40\}$). Second, conditional on there being x red balls in the urn, then the agent must assign probabilities $x/20$, $(40-x)/20$, and 1/3 respectively to R, G, B.[24] (The priors we thus identify are the only ones compatible with the given information on the composition of the urn – the *objectively rational beliefs* of Gilboa et al. (2010); see Cerreia-Vioglio et al. (2011) for more.) With these restrictions it is easy to see that the marginal belief assigned to the number of red balls $\{0, \ldots, 40\}$ identifies *uniquely* the belief on the full state space Ω. For this reason, in our analysis below we focus only on the marginal beliefs over $\{0, \ldots, 40\}$, referring to them as 'priors,' which means that our analysis *de facto* proceeds as if the state space were simply $\{0, \ldots, 40\}$. This is without loss of generality if we focus on beliefs satisfying the restrictions above, and it also allows us to express our conditions more clearly.[25] Define $\hat{\Omega} := \{0, \ldots, 40\}$.

We start by analyzing the utility of each of our five gambles under the MaxMin Expected Utility model.[26] By u we understand a generic continuous function from $\mathbb{R}_+$ to $\mathbb{R}$. Let us introduce some additional notation: for every $\pi \in \Delta(\hat{\Omega})$, denote by $\mathbb{E}_\pi[x]$ the expectation of π, i.e. $\mathbb{E}_\pi[x] = \sum_{x=0}^{40} \pi(x)x$. $\mathbb{E}_\pi[x^2]$ and $\mathbb{E}_\pi[u(x)]$ are defined analogously. VAR_π denotes the variance of π. Table 9 displays the utility of each option for ambiguity averse (N$\succcurlyeq$P) agents, and for ambiguity averse and risk neutral ones ($u(x) = x$). (Since every utility is multiplied by $\frac{1}{60}$, for simplicity it is dropped from Table 9.) Notice that the utility of ambiguity loving agents can be obtained trivially by replacing the min with a max in the second and third column.

Our next task is to understand if the observed rankings are compatible with the

[24]For example, if the agent is sure that there are no red balls ($\pi(\{R, G, B\} \times \{0\}) = 1$), then she must assign probability zero to R ($\pi(\{R\} \times \{0, \ldots, 40\}) = 0$).

[25]Indeed we will have to be careful to always refer to the original states space whenever required, e.g. if we wish to characterize sets of priors to be convex.

[26]As we discussed, each alternative D,P,M, and C is presented to the subjects in two forms, depending on the winning color. For simplicity, we consider here only the option in which the 'winning color' is red. (For M we consider the case in which if red ball is extracted, the agent wins \$$g$.) The case of green is symmetric.

Table 9: Utility of each act using the MaxMin Expected Utility Model

Act	*Amb. Averse*	*Amb Av and Risk Neutral*
N	$20\ u(20)$	20^2
P	$\min_{\pi\in\Pi} \mathbb{E}_\pi[x]\ u(20)$	$\min_{\pi\in\Pi} \mathbb{E}_\pi[x]\ 20$
C	$\min_{\pi\in\Pi} 20\ \mathbb{E}_\pi[u(x)]$	$\min_{\pi\in\Pi} 20\ \mathbb{E}_\pi[x]$
D	$\min_{\pi\in\Pi} \mathbb{E}_\pi[xu(x)]$	$\min_{\pi\in\Pi} \mathbb{E}_\pi[x^2] = \min_{\pi\in\Pi} \text{VAR}_\pi + \mathbb{E}_\pi[x]^2$
M	$\min_{\pi\in\Pi} \mathbb{E}_\pi[xu(40-x)]$	$\min_{\pi\in\Pi} 40\mathbb{E}_\pi[x] - \text{VAR}_\pi - \mathbb{E}_\pi[x]^2$

MMEU model, and if they are, what are the implied restrictions on the parameters of the model. For simplicity, we address these questions by first assuming risk neutrality. We then extend our conclusions to the general case.[27]

We start by considering whether a risk neutral MMEU subject could rank D$\succcurlyeq$N.

Observation 1. *A risk neutral MMEU subject with set of priors* Π *ranks D*$\succcurlyeq$*N if, and only if,*

$$\text{VAR}_\pi \geq 20^2 - (\mathbb{E}_\pi[x])^2 \quad \forall \pi \in \Pi. \tag{2}$$

Moreover, the condition above implies that there is no $\pi \in \Pi$ *such that* $\mathbb{E}_\pi[x] < 20$ *and* $\pi(\{21, \ldots, 40\}) = 0$. *That is, every prior in* Π *with an expected value below 20 must assign a* strictly positive *probability to there being* strictly *more than 20 red balls.*

(If we consider subjects who are not risk neutral, condition (2) becomes $\mathbb{E}_\pi[xu(x)] \geq 20^2$.) Observation 1 shows that the ranking D$\succcurlyeq$N$\succcurlyeq$P *is compatible* with the MMEU model, but only if the variance of *all* priors in the set of priors Π is above a certain threshold. In particular, Condition (2) posits that if an agent has a prior in his set of priors with an expected value below the 'symmetric' value of 20 – i.e. if she has a 'pessimistic' prior in terms of the number of red balls – then this prior *should also have a non-trivial variance.* This variance must be higher than the difference between the square of 20 and the square of the expected value.

One possible interpretation of Condition (2) is the following. While subjects could act as if they were pessimistic about the number of red or geen balls – they have a prior with a low expected value – they cannot act as if they were *sure* about this pessimistic valuation: they should incorporate in this prior the awareness that they

[27]One the one hand, risk neutrality allows us to obtain much simpler expressions while conveying the main idea of the restrictions that we find. On the other hand, as we have seen in our data the role of the agent's risk attitude seems to be secondary, since when we focus on (weakly) risk averse subjects the rankings change minimally.

are being pessimistic. In fact, as mentioned in the second part of Observation 1, Condition (2) also implies that every pessimistic prior in Π must also assign a strictly positive probability to there being strictly more than 20 red balls: that is, even for the pessimistic priors, while the expected value can be below 20 ($\mathbb{E}_\pi[x] < 20$), the agent must admit the possibility that there are, in fact, *strictly more* than 20 balls ($\pi(\{21, \ldots, 40\}) > 0$). We denote this behavior *Skeptical Pessimism*.

As we discussed in the introduction, Condition (2) is not compatible with the casual "classroom" example of MMEU. For instance, any degenerate prior according to which there are only $r < 20$ red balls ($\pi(r) = 1$) is not compatible with Condition (2) – this prior has an expected value below 20, but a zero variance.

Condition (2) gives us the restrictions on Π that are necessary and sufficient for a risk neutral MMEU agent to rank D$\succcurlyeq$N. We derive similar conditions that guarantee that the agent ranks D$\succcurlyeq$ P and D$\succcurlyeq$M.

Observation 2. *A risk neutral MMEU subject with set of priors* Π *ranks* $D \succcurlyeq P$ *if and only if*

$$\min_{\pi\in\Pi} \mathrm{VAR}_\pi + \mathbb{E}_\pi[x]^2 \geq \min_{\pi\in\Pi} 20\mathbb{E}_\pi[x] \tag{3}$$

If the most pessimistic priors for the evaluation of D and P were the same and equal to π*, then (3) becomes*

$$\mathrm{VAR}_\pi \geq 20\mathbb{E}_\pi[x] - \mathbb{E}_\pi[x]^2 \tag{4}$$

where $\mathbb{E}_\pi[x] \leq 20$ *(because the agent ranks* $N \succcurlyeq P$*). Notice that Condition (4) is strictly weaker than Condition (2).*

Observation 3. *A risk neutral MMEU subject with set of priors* Π *ranks* $D \succcurlyeq M$ *if and only if*

$$\min_{\pi\in\Pi} 2 * (20\mathbb{E}_\pi[x]) - \mathrm{VAR}_\pi - \mathbb{E}_\pi[x]^2 \geq \min_{\pi\in\Pi}\mathrm{VAR}_\pi + \mathbb{E}_\pi[x]^2 \tag{5}$$

If the most pessimistic priors for the evaluation of D and M were the same and equal to π'*, then, the (5) becomes*

$$\mathrm{VAR}_{\pi'} \geq 20\mathbb{E}_{\pi'}[x] - \mathbb{E}_{\pi'}[x]^2 \tag{6}$$

where $\mathbb{E}_{\pi'}[x] \leq 20$ *(because the agent ranks* $N \succcurlyeq P$*). Notice that Condition (6) is strictly weaker than Condition (2).*

(If we consider subjects who are not risk neutral, (3) becomes $\min_{\pi\in\Pi} \mathbb{E}_\pi[xu(x)] \geq \min_{\pi\in\Pi} 20\mathbb{E}_\pi[x]$, while (4) becomes $\mathbb{E}_\pi[xu(x)] \geq 20^2$. Condition (5) becomes $\min_{\pi\in\Pi} \mathbb{E}_\pi[xu(x)] \geq \min_{\pi\in\Pi} 2 * (20\mathbb{E}_\pi[u(x)]) - \mathbb{E}_\pi[xu(x)]$.))

Observation 2 and 3 show that to obtain D$\succcurlyeq$P or D$\succcurlyeq$M we need to impose on Π conditions that are reminiscent of Condition (2), but are strictly *weaker*. In fact, while they apply only to *some* of the priors in Π (whereas (2) applies to all), they also posit restrictions on the minimal variance of the members of Π in line with the notion of skeptical pessimism. Both conditions, for example, are violated by the 'classroom example' of MMEU mentioned before.

If we look at our data in light of the results above, we notice that if subjects were modeled using the MMEU representation, then almost all of the ambiguity averse subjects would have to satisfy at least one of the conditions above. In fact, 97% rank D$\succcurlyeq$M and would satisfy Condition (5); 74% rank D$\succcurlyeq$P and would satisfy Condition (3); 40% would rank D$\succcurlyeq$N and satisfy Condition (2). Moreover, while Condition (2) applies to ambiguity averse subjects, it is easy to see that an ambiguity loving one (P$\succcurlyeq$N) who is also risk neutral must have D$\succcurlyeq$N.[28] This is true in our data: every subject who ranks P$\succcurlyeq$N also ranks D$\succcurlyeq$N.

Finally, the discussion above emphasizes how the MMEU model need not be a model of 'extreme pessimism,' as one might be tempted to say from its functional form. As we have seen, while in this model agents use priors with a 'low' expected value, at the same time these priors could also have a high variance and allow for the possibility of good outcomes – inducing the agent to chose an ambiguous option, D, against a non-ambiguous one, N.

3.2 Second-Order Expected Utility Models

The MaxMin Expected Utility model is only one amongst the many models of ambiguity aversion that have been suggested in the literature.[29] Some of these models are strict generalization of MMEU, and for each of them it is not hard to derive conditions that guarantee D$\succcurlyeq$N$\succcurlyeq$P in the spirit of those derived for the MMEU model.[30] Other well-known models represent ambiguity-averse behavior via Second-Order Expected Utility (SOEU). In these models, subjects evaluate each act using a functional of the form

$$U(f) = \int_{\Delta(\Omega)} \mu(\pi)\phi\big(\int_{\Omega}\int_{X} u(f(w)(x))\mathrm{d}x \ \mathrm{d}\omega\big)\mathrm{d}\pi \tag{7}$$

where μ is a prior over priors and ϕ is a concave function. (Ambiguity loving subjects are modeled using a convex ϕ.) The idea is that subjects have a prior over priors μ, and they aggregate the expected utility computed with each prior via a concave function ϕ, where the concavity of ϕ leads to ambiguity aversion just like the concavity of a

[28] To see why, notice that we have D$\succcurlyeq$P iff $\max_{\pi\in\Pi} VAR_\pi + \mathbb{E}_\pi[x]^2 \geq \max_{\pi\in\Pi}\mathbb{E}_\pi[x]\ 20$. Consider now $\bar{\pi} = \text{argmax}\mathbb{E}_\pi[x]$; since the agent is ambiguity loving it must be that $\mathbb{E}_{\bar{\pi}}[x] \geq 20$. But then $\max_{\pi\in\Pi}\mathbb{E}_\pi[x]\ 20 = \mathbb{E}_{\hat{\pi}}[x]\ 20 \leq \mathbb{E}_{\hat{\pi}}[x]^2 \leq \mathbb{E}_{\hat{\pi}}[x]^2 + VAR_{\hat{\pi}} \leq \max_{\pi\in\Pi} VAR_\pi + \mathbb{E}_\pi[x]^2$. Therefore, D$\succcurlyeq$P. Since P$\succcurlyeq$N, we have D$\succcurlyeq$P$\succcurlyeq$N.

[29] Halevy (2007) surveys many of these models, and conducts an ingenious experiment in which he compares behavior in the standard Ellsberg treatment with treatments with objective risk in order to differentiate which of these models provides a more accurate description of subjects' behavior.

[30] Among many, see for example the Variational Preferences of Maccheroni et al. (2006), the Uncertainty Averse Preferences of Cerreia et al. (2010), the Biseparable Preferences of Ghirardato et al. (2004), or the MBC preferences of Ghirardato and Siniscalchi (2010).

utility function leads to risk aversion in von-Neumann Morgenstern Expected Utility. Models of this kind appear in Klibanoff et al. (2005), Ergin and Gul (2009), and Seo (2009).

One of the main differences between the MMEU and SOEU models is that in the former the agent either considers a prior π, or disregards it entirely; by contrast, in latter the agent may assign different weights to different priors. This is a relevant difference if we aim to find the equivalent to Condition (2) for models of the latter class. In fact, for the MMEU model we have argued that priors that are pessimist and have a low variance should not be considered at all by an agent who ranks D$\succcurlyeq$N. This allows us to suggest restrictions on the set of priors Π. By contrast, an agent with SOEU preferences could rank D$\succcurlyeq$N and *also* consider priors that are both pessimistic and have a low variance, in the sense that they belong to the support of μ, by assigning to them a 'smaller' weight. In particular, in the SOEU case we observe D$\succcurlyeq$N as long as μ satisfies some specific restrictions on the induced variance on the number of red balls. It would hold in two cases. First, if μ gave a relatively high weight to the priors with 'high' variance. In this case we would have D$\succcurlyeq$N even if the agent is 'very ambiguity averse' (ϕ is 'very concave'). Alternatively, it would also hold if μ admits in its support priors with a low variance, but these priors are very different from one another, in such a way that the expectation of μ, which is a prior over Ω, has a relatively high variance over the number of red balls. In this latter case, we would have D $\succcurlyeq$ N as long as the agent is not 'too ambiguity averse' (ϕ is mildly concave).[31]

Almost identical arguments could be used to find equivalent restrictions to the model of prior uncertainty of Cerreia-Vioglio et al. (2011), and similar arguments lead to similar restrictions for other well-known models like Segal (1987), Schmeidler (1989), Halevy and Feltkamp (2005), and Siniscalchi (2009).

4 Conclusion

This paper extends the classical Ellsberg experiment to investigate decision-makers' reaction to gambles in which the prize amount, the probability of winning, or both, may depend on the composition of an ambiguous urn. Most of our subjects exhibit the same attitude towards ambiguity when either the prize or the probability (but not both) are unknown. However, a large fraction of our subjects display a different attitude when both the winning probability and the amount of the prize are positively correlated and depend on the unobserved composition of the urn: If we focus on the subjects who prefer the gamble with no ambiguity to the gamble in which the probability of winning depends on the urn composition, then 40% of them rank both

[31]For example, consider some μ with these two degenerate priors in its the support: 1) there are 40 red balls and the probability of R is $\frac{2}{3}$; 2) there are no red balls and the probability of R is zero. Then, for mildly concave ϕ such agent would rank N$\succcurlyeq$P and D$\succcurlyeq$N.

of these gambles *below* the one in which both the prize amount and the probability of winning depend on the urn's composition.We show that this behavior could be rationalized in a MMEU model by requiring that every prior in the set of priors of the agents has a non-trivial variance. In particular, the variance should be above a threshold which is higher the more pessimistic the prior is – a condition that we called skeptical pessimism.

While our result tests for the presence of skeptical pessimism with a simple modification of the typical Ellsberg experiment, it would be interesting for future work to test if similar conditions on the variance of the priors hold in other settings as well.

Appendix A: Instructions

Subjects received the following instructions.[32]

INSTRUCTIONS

This is an experiment in decision making. Various research institutions have provided the money for this experiment.

The experimenter will stand in front of you with a non-see-through bag that contains 60 poker chips.

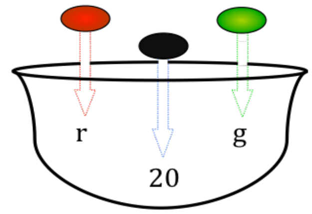
- **20** of these chips are BLACK,
- **r** chips are RED and
- **g** chips are GREEN,
- where **r+g=40**.

You are not told how many RED and GREEN chips there are. All you know is that the number of RED chips plus the number of GREEN chips is 40. Thus, the number of RED chips and the number of GREEN chips may be any whole number between 0 and 40 such that the sum of these two numbers is 40.

Note that you are also not told *how* the number of RED and GREEN chips (i.e. r and g) is determined.

[32] What follows are the instructions that subjects received during sessions 3 and 4. They differ from those they received during sessions 1 and 2, since in those sections subjects were also asked to specify their willingness to pay for each option, with a BDM mechanism. (The instructions therefore contained an explanation of the mechanism.) As we mention in the main text, however, we disregard these answers in our analysis, and conduct robustness tests that shows that the presence of these questions had no (significant) impact on the answers. For these reason, we report here the instructions that subjects received for sections 3 and 4 only. (The instructions for the sessions 1 and 2 are available upon request.)

60 Chips in total are in the bag
20 Black chips are in the bag
r Red chips are in the bag
g Green chips are in the bag
r+g = 40 chips



At the end of the experiment, each of you may inspect the content of the bag

In this experiment you are asked to answer a questionnaire. In the questionnaire you are first asked to tell us your gender, your SAT score, and your year of birth. This is followed by list of questions in which you are asked to choose among *lotteries*. The following are examples of what we call a *lottery*.

Example 1: If the chip drawn is BLACK, you win \$12, but if the chip drawn is either RED or GREEN, you win nothing.
Example 2: If the chip drawn is BLACK or RED, you win nothing, but if the chip drawn is GREEN you win an amount of money equal to the number of GREEN chips.

Note that the prize amount in a lottery may depend on the number of RED and GREEN chips as in Example 2 above.

In the questionnaire, you will be given a list of lotteries and will be asked to choose ONE lottery from the list, which you most prefer. In each question the number of lotteries will vary between 2 and 4. The following is an example of a question where you are asked to choose ONE lottery out of THREE:

Which lottery do you prefer?

If BLACK then \$12, if RED then \$0 and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$g	
If BLACK then \$0, if RED then \$r and if GREEN then \$0	

Please answer **all the questions**. If you wish to change your answer, please mark your final answer clearly.

Please take your time in answering all the questions. You have plenty of time to think about each question. When you have finished answering all of the questions, **please review all of your answers by reading the whole questionnaire again**. Note that there is no advantage to finishing quickly as the experiment will end only when everyone has finished answering.

Once all participants have finished answering and reviewing their questionnaire, we will proceed as follows:

1. The experimenter will randomly draw a chip from the cloth-bag (in front of everybody).

2. Using dice (in front of you), the experimenter will draw one question number from the questionnaire.

3. You will then be paid according to the amount specified by the lottery you have chosen for that question, plus a show-up fee of \$7. (This might involve counting the number of red and green chips in the bag.) For example, suppose the lottery you chose for the question stated that "If the chip drawn is BLACK, you win \$12, but if the chip drawn is either RED or GREEN, you win nothing." Then if the experimenter draws a black chip you win \$12; if the experimenter draws a red or a green chip, you win nothing.

Payments

- You will receive a show-up fee of **\$7**, which is your to keep *regardless* of the decisions you make in the experiment.
- You will also be paid an amount that will depend on the question that is randomly selected, the answer you gave to that question and the chip that is drawn.

Summary To summarize:

1. You will be presented with a cloth-bag containing 60 chips, 20 of which are BLACK, the others are GREEN and RED. The total number of RED and GREEN chips is 40 but you are not told how many of these are RED or GREEN. In addition, you are not told how the number of RED or GREEN chips was determined.
2. You will be presented with a list of questions in which you will be asked to choose between lotteries.
3. After you have answered all the questions, please review them to make sure you are satisfied with the answers you gave.
4. Finally, the experimenter will draw a chip from the cloth-bag, and then:
   (a) For each of you, she will roll a dice to randomly select one of the questions of the questionnaire;
   (b) You will be paid the amount specified by the lottery you selected, plus a show-up fee.

Appendix B: Questionnaire

Subjects answered the following questionnaire.[33]

QUESTIONNAIRE

[33] This is the Questionnaire that subjects answered to in sessions 3 and 4. As we mentioned before, the Questionnaire in sessions 1, 2 was different (see footnote 32), and it is available upon request.

Please Indicate

Gender	
Year of Birth	
SAT Score	

Please answer each of the following questions:

Question 1

Which lottery do you prefer?

If BLACK then \$20, if RED then \$0 and if GREEN then \$0.	
If BLACK then \$0, if RED then \$20 and if GREEN then \$0.	
If BLACK then \$0, if RED then \$0 and if GREEN then \$20.	

Question 2

Which lottery do you prefer?

If BLACK then \$20, if RED then \$0 and if GREEN then \$0	
If BLACK then \$r, if RED then \$0 and if GREEN then \$0	
If BLACK then \$g, if RED then \$0 and if GREEN then \$0	

Question 3

Which lottery do you prefer?

If BLACK then \$20, if RED then \$0 and if GREEN then \$0	
If BLACK then \$0, if RED then \$r and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$g	

Question 4

Which lottery do you prefer?

If BLACK then \$0, if RED then \$20 and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$20	
If BLACK then \$r, if RED then \$0 and if GREEN then \$0	
If BLACK then \$g, if RED then \$0 and if GREEN then \$0	

Question 5

Which lottery do you prefer?

If BLACK then \$0, if RED then \$20 and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$20	
If BLACK then \$0, if RED then \$r and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$g	

Question 6

Which lottery do you prefer?

If BLACK then \$r, if RED then \$0 and if GREEN then \$0	
If BLACK then \$g, if RED then \$0 and if GREEN then \$0	
If BLACK then \$0, if RED then \$r and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$g	

Question 7

Which lottery do you prefer?

If BLACK then \$40, if RED then \$0 and if GREEN then \$0	
If BLACK then \$0, if RED then \$20 and if GREEN then \$20	

Question 8

Which lottery do you prefer?

If BLACK then \$13, if RED then \$13 and if GREEN then \$13	
If BLACK then \$40, if RED then \$0 and if GREEN then \$0	

Question 9

Which lottery do you prefer?

If BLACK then \$20, if RED then \$20 and if GREEN then \$20	
If BLACK then \$r, if RED then \$r and if GREEN then \$r	
If BLACK then \$g, if RED then \$g and if GREEN then \$g	

Question 10

Which lottery do you prefer?

If BLACK then \$20, if RED then \$20 and if GREEN then \$20	
If BLACK then \$r, if RED then \$20 and if GREEN then \$20	
If BLACK then \$g, if RED then \$20 and if GREEN then \$20	

Question 11

Which lottery do you prefer?

If BLACK then \$20, if RED then \$0 and if GREEN then \$0	
If BLACK then \$0, if RED then \$g and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$r	

Question 12

Which lottery do you prefer?

If BLACK then \$0, if RED then \$r and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$g	
If BLACK then \$0, if RED then \$g and if GREEN then \$0	
If BLACK then \$0, if RED then \$0 and if GREEN then \$r	

References

Camerer, C. (1995): "Individual decision making," in *Handbook of Experimental Economics*, ed. by A. R. J. Kagel, Princeton University Press, Princeton, vol. 2.

Casadesus-Masanell, R., P. Klibanoff, and E. Ozdenoren (2000): "Maxmin expected utility through statewise combinations," *Economics Letters*, 66, 49–54.

Cerreia, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2010): "Uncertainty Averse Preferences," .

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011): "Ambiguity and Robust Statistics," Mimeo, Universita' Bocconi.

Ellsberg, D. (1961): "Risk, ambiguity, and the Savage axioms," *The Quarterly Journal of Economics*, 75, 643–669.

Ergin, H. and F. Gul (2009): "A theory of subjective compound lotteries," *Journal of Economic Theory*, 144, 899–929.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004): "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory*, 118, 133–173.

Ghirardato, P. and M. Siniscalchi (2010): "A more robust definition of multiple priors ," 1–57.

Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with a non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Gilboa, I., D. Schmeidler, F. Maccheroni, and M. Marinacci (2010): "Objective and subjective rationality in a multi-prior model," *Econometrica*, 78, 755–770.

Halevy, Y. (2007): "Ellsberg Revisited: An Experimental Study," *Econometrica*, 75, 503–536.

Halevy, Y. and V. Feltkamp (2005): "A Bayesian Approach to Uncertainty Aversion," *Review of Economic Studies*, 72, 449–466.

Karni, E. and Z. Safra (1987): "" Preference reversal" and the observability of preferences by experimental methods," *Econometrica: Journal of the Econometric Society*, 55, 675–685.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A smooth model of decision making under ambiguity," *Econometrica*, 73, 1849–1892.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498.

Schmeidler, D. (1989): "Subjective probability and expected utility without additivity," *Econometrica*, 57, 571–587.

Segal, U. (1987): "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 28, 175–202.

Seo, K. (2009): "Ambiguity and second-order belief," *Econometrica*, 77, 1575–1605.

Siniscalchi, M. (2009): "Vector Expected Utility and Attitudes Toward Variation," *Econometrica*, 77, 801–855.