

Schröder, David; Cavatorta, Elisa

Conference Paper

Measuring Ambiguity Preferences

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Risk and Ambiguity II, No. E06-V2

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Schröder, David; Cavatorta, Elisa (2014) : Measuring Ambiguity Preferences, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Risk and Ambiguity II, No. E06-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100593>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Measuring Ambiguity Preferences[†]

Elisa Cavatorta

King's College, University of London*

David Schröder

Birkbeck College, University of London**

First version: November 2013

This version: February 28, 2014

Preliminary and incomplete – do not cite or circulate

Abstract

Ambiguity aversion has shown to be economically relevant and has been proposed as an explanation for many phenomena in economics and finance. While the literature has suggested a large variety of elicitation methods to measure ambiguity preferences, their consistency and reliability it is rarely evaluated. This is the first study that systematically analyses the consistency of individual ambiguity preferences elicited using a variety of incentivized tasks, non-incentivized thought experiments and survey questions. We find a high degree of aggregate consistency across elicitation methods, but large discrepancies in degrees of individual consistency in pair-wise tasks comparisons. Finally, the study identifies a set of non-incentivized tasks that predict ambiguity attitudes elicited experimentally which may serve as a viable alternative when running laboratory experiments is unfeasible.

JEL Classification: C81, C83, C91, D81

Keywords: ambiguity, preference measurement, decision making, experimental economics, survey validation

[†]We are grateful to Johannes Abeler, Anke Becker, Syngjoo Choi, Jean-Marc Tallon, Matthias Wiebral and seminar participants at Paris-I for helpful discussion and comments. Any remaining errors are ours.

*King's College, University of London, 22 Kingsway, London WC2B 6NR, United Kingdom. E-mail: elisa.cavatorta@kcl.ac.uk. Phone: +44 20 7848 7495.

**Birkbeck College, University of London, Malet Street, Bloomsbury, London WC1E 7HX, United Kingdom. E-mail: d.schroeder@bbk.ac.uk. Phone: +44 207 631 6408.

1 Introduction

At the latest since the seminal Ellsberg (1961) thought experiment, the literature on decision-making under uncertainty distinguishes two different types of uncertainty: risk and ambiguity. While risk denotes a situation with uncertain outcomes with a probability distribution over these outcomes, ambiguity refers to a situation with uncertain outcomes in the absence of such a probability distribution.¹ Furthermore, the typical finding of empirical studies is that people tend to prefer risky situations over ambiguous situations. This preference pattern is usually referred to as ambiguity aversion.

In the past decades, ambiguity preferences have been established as a category of fundamental economic human preferences on its own. Most of all, ambiguity aversion has been shown meaningful in many different contexts, as it can explain a variety of phenomena and puzzles that cannot be attributed by risk aversion alone.²

Although in many empirical studies the majority of subjects is ambiguity averse, confirming the original Ellsberg (1961) thought experiment, experimental evidence shows that individual attitudes toward ambiguity are not uniform. Evidence of ambiguity seeking behaviour has been found for small probabilities of gains and large probabilities of losses (Einhorn and Hogarth, 1985), and in the domain of losses in general, suggesting a reflection effect over ambiguous prospects (Chakravarty and Roy, 2009).

The abundance of real life situations with undefined probabilities highlights the importance of accurate measures of ambiguity preferences. Following Ellsberg’s seminal work, various methods to elicit ambiguity attitudes have been proposed. The experimental economics literature advocates the use of incentivized elicitation tasks. In the earlier economic literature and in psychology, the use of non-incentivized elicitation tasks, such as thought-experiments and survey questions, is more common. The literature also suggests a diversity of designs to measure ambiguity preferences. These include pair-wise choices (or multiple choice lists) between a risky and an ambiguous lottery, choice lists between an ambiguous lottery and a sure amount of money, as well as monetary valuation tasks of risky and ambiguous lotteries.

Despite this variety, the robustness of ambiguity attitudes elicited across methods and tasks is rarely evaluated. Some attempts in this direction have shown a significant heterogeneity of responses across tasks, which question the consistency of ambiguity attitudes. Stahl (2013) shows that slight changes in the elicitation task design (e.g. the amount of the prize for the ambiguous option) changes substantially the distribution of ambiguity aversion in the sample. Using different models, he estimates that a mere 12% of people show evidence of ambiguity aversion. Similarly, Binmore et al. (2012) report mild degrees of ambiguity aversion across a battery of ambiguity models and weak evidence for consistent ambiguity aversion. This literature has focused on the classical 2 and 3 color-Ellsberg urns tasks. Trautmann et al. (2011) bring evidence that choice

¹Following Knight (1921), ambiguity is sometimes also called Knightian uncertainty.

²Ambiguity aversion has shaped both the empirical and theoretical literature. Ambiguity aversion has been successfully used to explain the equity premium puzzle (Gilboa and Schmeidler, 1989; Epstein and Schneider, 2008) the small-firm effect (Chen and Epstein, 2002; Mukerji and Tallon, 2001; Olsen and Troughton, 2000), climate change (Weitzman, 2009; Millner et al., 2013), adolescent behaviors (Sutter et al., 2013), and employment outcomes among the poor (Spears, 2013). Ambiguity attitudes have shown to persist in experimental market settings (Sarin and Weber, 1993) as well outside the lab, among business owners (Chesson and Viscusi, 2003), children (Sutter et al., 2013), and Ethiopian farmers (Akay et al., 2012).

tasks elicit lower ambiguity aversion than valuation tasks. In sum, the type of elicitation tasks used has an important influence in measuring ambiguity attitudes.

However, the consistency of ambiguity judgements is of central importance for an informative measurement of individual preferences. This is the first study that systematically analyses the consistency of individual responses across elicitation tasks and different incentive methods. Our analysis uses information on ambiguity attitudes elicited using typical experimental tasks incentivized with monetary amounts, not-incentivized tasks and survey questions. Unlike previous studies, which are mainly concerned with consistency across the classical Ellsberg urns tasks, we examine a variety of tasks. Evaluating different designs is important because if the different elicitation tasks lead to a shift in the distribution of ambiguity attitudes, the degree of ambiguity aversion elicited from different methods are not informative, but the ranking of individuals according to their ambiguity attitudes is, a point Trautmann (2013) stress in their review. On the contrary, if elicitation tasks affect the ranking of individuals, then one may ask whether ambiguity attitudes are at all elicitable with the currently used designs. Using different designs allows us to compare our results to the literature and, since different tasks may be more appropriate in certain contexts than others, to offer a comparative guidance to researchers selecting suitable tasks.

To shed light on the consistency of measured ambiguity preferences, we first examine the distributions of revealed ambiguity attitudes across tasks. This is a measure of aggregate consistency across methods and tasks. Secondly, we look at the share of within-subject consistent choices with respect to their ambiguity attitudes in different tasks and briefly comment on the consistency of these measures over time. This is a measure of individual consistency (and individual consistency over time). We then investigate how different tasks rank individual ambiguity attitudes by analyzing the correlation among various ambiguity measures. We conclude by exploring which set of non-incentivized tasks best predict ambiguity attitudes elicited in the experiments incentivized with monetary rewards. This validation exercise indicates candidate measures of ambiguity attitudes which are behaviorally valid and can act as a viable alternative when running experiments is unfeasible.

Our preliminary results show a high degree of aggregate consistency across tasks. The aggregate classifications of our subjects closely mimic results from previous literature using similar tasks, providing an indirect indicator of aggregate consistency. However, individual consistency is more volatile: while we find that subjects are in general more consistent than one would expect under random choice, subjects are not fully consistent in their responses across tasks. Interestingly, tasks inspired by the original Ellsberg thought-experiment elicit more consistent responses than designs framed otherwise. This is intuitively appealing as it is expected that tasks with similar underlying designs would elicit comparable responses. Our results confirm this intuition and show that this is the case across different incentive methods. However, some of the most popular thought-experiments in the literature on ambiguity have no relation to the behavior observed in the incentivized experiments. Finally, our analysis identifies two simple thought-experiments and two survey questions that perform best in predicting behavior in the incentivized experiments: their combination can explain almost 25% of the variation in individual ambiguity preferences elicited experimentally.

The paper develops as follows. The next section describes the experimental design and the various elicitation methods to measure ambiguity preferences. In section 3 we present the main results

of this study. After a brief summary statistics, we examine the aggregate consistency of the measured ambiguity preferences across elicitation methods. Then we analyze whether the measured ambiguity preferences are consistent on the individual level. Next, we analyze the stability of measured ambiguity preferences over time. The last subsection examines whether non-incentivized experiments and survey questions can predict an individual's ambiguity preferences obtained in a standard experimental setup. Section 4 offers some concluding remarks and implications.

2 Experimental design

This study uses incentivized tasks, non-incentivized thought experiments and a survey questionnaire to measure ambiguity preferences. The incentivized tasks and non-incentivized thought experiments are taken from well-known studies on ambiguity preferences in the experimental economics literature. The survey questionnaire is based on various seminal studies on ambiguity tolerance in the psychology literature.

2.1 Incentivized tasks

Subjects participated in three incentivized tasks, two ambiguity tasks and one risk task. The two tasks to elicit ambiguity preferences are based on the works by Lauriola and Levin (2001), Chakravarty and Roy (2009) and Butler et al. (2013). Since ambiguity preferences are closely related to risk preferences (Chakravarty and Roy, 2009), we also include an task to elicit risk preferences, similarly taken from Chakravarty and Roy (2009). A detailed description of the tasks can be found in appendix A.

All three tasks use binary choice lists to measure ambiguity or risk preferences. Similar to the multiple price lists used in Holt and Laury (2002), subjects are asked to state their strict preference between two different lotteries. The ambiguity tasks involve choosing between an ambiguous lottery and a risky lottery, while the risk task involves choices between two different risky lotteries. The lotteries are presented in the form of two-color urns, similar to Ellsberg (1961).³

The first task (ambiguity task 1) is taken from Chakravarty and Roy (2009). Subjects are presented a decision table with 10 choices between drawing a ball from either a risky or an ambiguous urn. The composition and payoff structure of the ambiguous urn is identical in all 10 situations. In contrast, the potential payoffs of the risky urn decreases monotonically from one situation to the next, which consequently lowers its expected payoff. Subjects are therefore expected to switch at some point from drawing a ball from the risky urn to drawing a ball from the ambiguous urn. An ambiguity-neutral decision maker is indifferent between both lotteries in situation number 5. Ambiguity seeking individuals would rationally switch before situation 5, while ambiguity averse decision markers would switch after situation 6.

In the second task (ambiguity task 2), subjects are presented another decision table with 11 choices between a risky or an ambiguous urn. The composition and payoff structure of the ambiguous urn is again identical in all 11 situations. In contrast, the expected payoff of the risky urn increases monotonically. Different from task 1, this change is induced by increasing the probability of winning some prize, while leaving the potential earnings constant. Subjects are therefore expected

³All tasks are over gains, i.e., there is no possibility to lose money.

to switch at some point from the ambiguous urn to the risky urn. An ambiguity-neutral decision maker is indifferent between both lotteries in situation number 5. Ambiguity seeking individuals would rationally switch after situation 6, while ambiguity averse decision makers would switch before situation 5.

In the risk task, also taken from Chakravarty and Roy (2009), subjects are presented a decision table with 10 choices between drawing a ball from two different risky urns. One urn, which involves only a small risk, is identical in all situations. The other urn is more risky, while its expected payoff increases monotonically. Subjects are therefore expected to switch at some point from the low-risk urn to the more risky urn. A risk-neutral decision maker is indifferent between both lotteries in situation number 6. Risk seekers would rationally switch before situation 6, while risk averse decision makers would switch after situation 7.

2.2 Non-incentivized thought experiments

Besides the incentivized experimental preferences measures, we also use several non-incentivized thought experiments to elicit ambiguity preferences. Similar to the experiments, we use some of the most common thought experiments of the literature.

First, we replicate the seminal two-color Ellsberg (1961) urn experiment. Similar to Butler et al. (2013), subjects are offered five different answer possibilities to allow differentiating between different degrees of ambiguity preferences. The second thought experiment places the Ellsberg (1961) experiment in a more realistic setting of horse betting. The design is inspired by the field experiment of Viscusi and Chesson (1999) who elicit ambiguity preferences of business owners. In the third hypothetical experiment, we ask subjects to price hypothetical complementary bets contingent on future temperatures in a familiar city (London) and an unfamiliar city with a similar climate (Vancouver), similar to Fox and Tversky (1995). Since subjects live in London, ambiguity aversion should imply a higher willingness to pay for London temperature bets than for Vancouver temperature bets.⁴ Finally, we use a new dynamic version of the Ellsberg (1961) thought experiment where subjects are presented four (in some cases five) sequential decisions between a risky and an ambiguous lottery. While the ambiguous urn is identical in all situations, the composition of the in the risky urn changes from one situation to the next (and hence the probability to draw a ball of the winning color) depending on the subject's previous choices. This elicitation method was first used in psychology (Cornsweet, 1962), and is sometimes called staircase method. See appendix B for a detailed description of the four thought experiments.

2.3 Survey questionnaire

In addition to experimental measures of ambiguity preferences and thought experiments, this study uses survey questions to measure ambiguity preferences. While the using survey questionnaires to measure risk preferences (Guiso and Paiella, 2008; Dohmen et al., 2011) is common in the literature, there is yet no study in economics that explicitly uses questionnaires to measure ambiguity preferences.

⁴The temperature of a given city on some day in the future is commonly used in ambiguity experiments, see also Abdellaoui et al. (2011).

The survey or self-assessment questions to elicit ambiguity preferences are taken from self-reporting scales in the psychology literature which has investigated individuals' ambiguity tolerance since the 1960s. In these studies, subjects are asked to indicate the extent to which they agree or disagree with a list of statements on a scale from 1 to 7.

The survey questionnaire of this study is based on the Intolerance of Ambiguity Scale by Kirton (1981), which uses elements of the earlier seminal works by Budner (1962), Rydell and Rosen (1966) and Mac Donald Jr. (1970). In addition, we include selected items from the ambiguity tolerance scales by Budner (1962), Norton (1975) and McLain (2009) that fit well with our working definition of ambiguity. Since ambiguity attitudes might be related to optimism and pessimism (Chateauneuf et al., 2007), we include some optimism/pessimism items from the Extended Life Orientation test by Chang et al. (1997). Furthermore, given that ambiguity shares cognitive similarities with uncertainty, we also added several items of the Uncertainty Response Scale by Greco and Roger (2001). Since ambiguity attitudes might also be influenced by self-esteem (Heath and Tversky, 1991), we add the single self-esteem measure by Robins et al. (2001). Together with some own additions, the survey questionnaire contains 46 items in total. For the complete list see appendix C.

2.4 Procedure

Each session consisted of four parts. The first part included the survey questionnaire on ambiguity preferences, as described in section 2.3. The thought experiments (see section 2.2) followed in part 2. The third part included a standard demographic questionnaire. Finally, the last section consisted of the three tasks, see section 2.1. The first task was the risk task, followed by ambiguity tasks 1 and 2.

This particular sequence was chosen to reduce any potential spillover effects between the various methods to elicit ambiguity preferences. More specifically, we expected students to think more carefully about their choices in the incentivized tasks rather than in the thought experiments or the questionnaire. To avoid subjects to simply repeat their stated preferences of the incentivized tasks in the non-incentivized thought experiments and the questionnaire, we placed the actual tasks at the end of each session. In two sessions of a pilot study that was carried out before the actual experiment, we reversed the sequence of the various parts to test for order effects. The results were not significantly different from the main treatment.

Before the start of the two ambiguity tasks, participants had to select the color of the winning ball, i.e., either black or white. This ensures that subjects had no reason to believe that the experimenter had any strategic incentive to manipulate the color of the balls in the ambiguous urn.⁵ In the incentivized tasks, subjects were told that at the end of the session, the computer would randomly select one out of the situations of each task and then randomly draw one ball from the urn chosen. This procedure ensures that subjects state their true preferences.

The computer-based sessions were conducted in May and October 2013 at Birkbeck College, University of London. We used z-tree (Fischbacher, 2007) as experimental software. Each session

⁵If the color of the winning ball is set by the experimenter, this may lead to some mistrust of the subjects as the experimenter may want to economize on the budget. Consequently, subjects have a pessimistic belief about the fraction of winning balls in the ambiguous urn, which induces a preference of the risky urn over the ambiguous urn even in the absence of ambiguity aversion (Charness et al., 2013; Chow and Sarin, 2002).

lasted about 40 minutes. Earnings from the incentivized tasks were calculated in terms of points, and then converted at a rate of 2:1 into GBP. On average, subjects earned GBP 18.45, which includes a fixed show-up fee of GBP 10. The lowest payment was GBP 10, the highest payment GBP 26.

2.5 Participants

121 subjects participated in our study, all of them students of various University of London colleges. The participants were recruited via electronic mail and announcements at the beginning of graduate and undergraduate lectures of various study programmes. This sample contains 54 (45%) male and 67 (55%) female participants, with an average age of about 26 years. Since the answers of 19 participants are inconsistent with expected utility we had to discard their responses in some of the analyses, leaving us with 102 subjects. For a detailed sample description see appendix D.

3 Results

This section presents the main findings. First, section 3.1 summarizes the results of the incentivized tasks and the non-incentivized thought experiments. Section 3.2 analyzes the overall pattern of ambiguity preferences obtained from the different elicitation methods, and compares the results of previous studies. In section 3.3, we examine whether the various elicitation methods lead to similar preference measures on the individual level. Using a small extension of the experimental design, section 3.4 explores whether the measured ambiguity preferences are consistent over time. Finally, section 3.5 addresses the question whether non-incentivized thought experiments and survey questions can be used to proxy for experimentally measured ambiguity preferences.

3.1 Descriptive statistics

Tables 2 and 3 summarize the subjects' choices between drawing a ball from the risky or the ambiguous urns. In both incentivized tasks, subjects preferred in around 56% of the situations drawing a ball from the risky urn over drawing a ball from the ambiguous urn. Yet, the two tasks differ significantly in the standard deviation of the fraction of risky choices out of all situations. This divergence can be explained by the different design of the two tasks. Most of all, the increments in expected earnings of the risky lottery in task 1 are, on average, smaller compared to task 2. Therefore task 1 allows for a finer assessment of ambiguity preferences compared to task 2, which results in a larger dispersion of the subjects' choices.⁶

In both tasks, a typical strategy is a threshold strategy. Since the relative attractiveness of the risky lottery changes monotonically from situation to situation, many participants prefer one urn over the other up to a switching point, from which they prefer the other urn. In task 1, the

⁶In task 1, expected earnings of the risky urn change by 0.5 points from one situation to the next (the last two situations aside), while the expected earnings of the risky urn in task 2 change by 1 point from situation to situation, see appendix A.

natural choice is to first prefer drawing a ball from the risky urn (urn X) with expected earnings of 7 points, and then to switch to the ambiguous urn (urn Y) at some later stage as the expected earnings of the risky urn decrease. In task 2, the opposite holds true: in situation 0, there is no chance of winning anything in the risky urn (urn 1), such that the ambiguous urn (urn 2) is the preferred choice. However, as the number of winning balls in urn 1 increases, the risky urn becomes more attractive.

Similar to the fraction of risky choices, the standard deviation of switching points is much larger for task 1 compared to task 2. Yet, several participants switched multiple times between the risky and the ambiguous urn. In task 1, about 13% of the subjects switched more than once; in task about 4% displayed multiple switches. Although in the original study by Chakravarty and Roy (2009) no subject switched more than once, such pattern is common. Such a behavior is difficult to rationalize with any of the standard models of decision making under ambiguity, and may therefore indicate the subjects' misapprehension of the tasks. The difference in consistent responses between the two tasks suggest that ambiguity task 2 was easier to understand.

Table 4 reports the summary statistics of the four non-incentivized thought experiments. In the replication of the seminal Ellsberg (1961) thought experiment with five answer possibilities (Panel A), the majority of subjects are ambiguity-averse (64%), about 28% are ambiguity-seeking and 8% are ambiguity-neutral. In the thought experiment that places the Ellsberg urn experiment in a horse betting framework (see panel B), about 60% of the participants are ambiguity-averse, with the remaining subjects being ambiguity-seeking. The risk equivalent elicited in this setting, i.e., the probability for which a risky horse bet is equally attractive to subjects as the ambiguous horse bet is with 30.7% just above the mean of the range of the ambiguous horse bet (20% and 40%). In thought experiment 3 (pricing complementary temperature bets), we first add for each subject his or her willingness to pay for both sides of the complementary bets. In particular, we add together the willingness to pay for both London tickets, and the willingness to pay for both Vancouver tickets, similar to Fox and Tversky (1995). The absolute ambiguity premium is the difference between the combined willingness to pay for both London tickets (risky lottery) less the combined willingness to pay for both Vancouver tickets (ambiguous lottery). The relative ambiguity premium is defined as the absolute ambiguity premium divided by the combined willingness to pay for both London tickets. This normalization is designed to adjust for the subjects' risk aversion. Since the combined willingness to pay for complementary bets depends on a subject's risk aversion, the absolute ambiguity premium is less meaningful when comparing subjects with different a different risk aversion.⁷ In our sample, we find an average ambiguity premium of GBP 6.97 and a relative ambiguity premium of 13.5%. These results mean that subjects were willing to pay more to bet on familiar London temperature than on unfamiliar Vancouver temperature, consistent with the notion of ambiguity aversion. Finally, in the dynamic Ellsberg urn, ambiguity preferences can be inferred indirectly by the risk equivalent for each subject. This risk equivalent is defined as the probability which makes a participant

⁷A very risk-averse subject who has a combined willingness to pay for the two London bets of GBP 2 and a combined willingness to pay for the two Vancouver bets of GBP 1 has the same absolute ambiguity premium (GBP 1) than a less risk averse subject who is willing to pay 51 for the London bets and GBP 50 for the Vancouver bets. However, the relative ambiguity premia would differ, being 50% for the risk-averse subject and 0.2% for the less risk averse subject. The relative ambiguity premium is more meaningful, since the absolute ambiguity premium reflects both the subject's ambiguity aversion and risk aversion.

indifferent between the risky and the ambiguous urn.⁸ The average risk equivalent is around 44%, fairly concentrated in the range from 37.5% to 57.5%.

3.2 Aggregate consistency

This section analyses the overall distribution of ambiguity attitudes across the different elicitation methods, and compares our results to corresponding results of previous studies, where available. For each of the ambiguity tasks and thought experiments considered in this paper it is possible to classify subjects into different categories depending on their broad ambiguity preferences. Subjects can be sorted either in two categories (ambiguity-averse and ambiguity-loving decision makers), or three categories (ambiguity-seeking, ambiguity-neutral, and ambiguity-averse decision makers). In the tasks and some of the thought experiments, this classification is not obvious but requires choosing a definition of ambiguity neutrality. There are several possibilities and a certain degree of arbitrariness. We apply the most narrow definition of ambiguity neutrality for each elicitation method.⁹ Since Ellsberg thought experiment and the complementary temperature bets explicitly allow for ambiguity neutrality, it is not possible to apply a binary classification for these two elicitation methods.

Table 5 shows the fraction of subjects that are classified in these broad ambiguity preference groups. When using a binary classification of ambiguity preferences between ambiguity seeking and ambiguity averse behavior, the various tasks generate a remarkably similar distribution of preferences. Consistently between 60 and 70% of the subjects can be classified as ambiguity averse, with the remaining subjects being ambiguity seeking. A three-group classification, allowing for ambiguity neutrality, shows a different and less consistent picture. The valuation task, i.e., the willingness to pay for complementary temperature lottery tickets (Fox and Tversky, 1995) elicit higher ambiguity aversion than choice tasks, a result consistent with Trautmann et al. (2011). The two incentivized choice tasks by Lauriola and Levin (2001) and Chakravarty and Roy (2009) suggest broadly similar classifications, but the Lauriola and Levin (2001) design leads to a higher share of ambiguity neutral subjects (63% versus 50%). This difference can be explained by the larger increments of the expected earnings in the Lauriola and Levin (2001) task relative to Chakravarty and Roy (2009), inducing a smaller dispersion of switching points, as mentioned earlier. Interestingly, the Dynamic Ellsberg urn thought experiment, which uses a design very similar to Lauriola and Levin (2001), but presents the choices in different order, elicits a substantially lower share of ambiguity neutrality (39%). It may be possible that the Lauriola and Levin (2001) task is subject to a framing effect of a typical number induced by the 5 white balls/5 black balls choice appearing in the middle of the list. In addition, the distributions are more similar if the definition of ambiguity neutrality is set to include subjects with risk equivalent from 40% and 60%, similar to Lauriola and Levin (2001). Finally, the low fraction of ambiguity-neutral subjects in the simple Ellsberg thought experiment might again partly be explained by a

⁸The risk equivalent is calculated as the mid-point of the interval that is consistent with his or her choices in the thought experiment. For example, a subject that prefers a risky urn with a probability of winning of 50% over the ambiguous urn, but prefers the ambiguous urn over a risky urn with a probability of winning of 45%, has an estimated risk equivalent of 47.5%.

⁹For example, in the complementary temperature bets, an ambiguity premium of GBP 0 implies ambiguity neutrality. However, one may argue that a small interval around GBP 0 may be plausibly interpreted as ambiguity neutrality as well.

framing effect, as the ambiguity-neutral option represents only one out of five answer possibilities. Based on the classifications, we test whether the subjects are on average ambiguity neutral using a Wilcoxon (1945) signed-rank test. Consistently across the various elicitation methods, the null hypothesis of the subjects being on average ambiguity neutral is rejected at high confidence levels. Our results are consistent with previous studies that use similar tasks and thought experiments to measure ambiguity preferences. The two incentivized tasks elicit a distribution remarkably similar to those presented in the original studies by Lauriola and Levin (2001) and Chakravarty and Roy (2009). The horse betting thought experiment, which closely follows the design by Viscusi and Chesson (1999), presents an almost identical binary classification; in the three-category classification, our sample exhibits a slightly lower share of ambiguity neutrality than Viscusi and Chesson (1999). Yet, their three-category classification is based on a small sub-sample of their data, consisting of only 38 students.¹⁰ The non-incentivized version of the Ellsberg urn experiment with five answer possibilities shows a somewhat smaller fraction of ambiguity neutral subjects compared to Butler et al. (2013). The results for the subjects' willingness to pay for complementary temperature lottery tickets are qualitatively similar to Fox and Tversky (1995), although their average relative ambiguity premium is with 39% around twice as large as in our experiment that exhibits an average relative ambiguity premium of 20%.¹¹ Yet, they do not provide any distributional information on their sample. The dynamic Ellsberg urn experiment was to our knowledge first implemented in this paper.

To conclude, this section shows that the various tasks and thought experiments exhibit a remarkably high level of aggregate consistency with each other as well as to previous findings in the literature. This result is especially true when distinguishing only between ambiguity seeking and ambiguity averse behavior. Yet, a finer three-way classification shows less consistency across tasks. Nevertheless, the findings imply that the evidence of ambiguity aversion is robust to different elicitation methods and experimental designs used when looking at the aggregate sample.

3.3 Individual consistency

This section examines whether the measured ambiguity preferences are consistent across the various elicitation methods at the individual level. As a first assessment, we compare whether subjects exhibit the same broad ambiguity attitude across elicitation methods using the same classification as in the previous section. In a second step, we analyze the correlation statistics of the ambiguity preferences as measured by each of the elicitation methods.

We consider the choices of a subject pair-wise consistent if they reveal the same ambiguity attitude in two given elicitation methods. Table 6 presents the individual consistency of ambiguity preferences for each pair of elicitation methods. Panel A presents the individual consistency when using the binary classification that distinguishes between ambiguity seeking and ambiguity averse decision makers; panel B present the average consistency based on a classification into three groups, including ambiguity neutrality. Since the Ellsberg thought experiment and the complementary temperature bets explicitly allow for ambiguity neutrality, these tasks have been excluded from

¹⁰It has to be noted that the Viscusi and Chesson (1999) design is framed as a loss, while ours is framed as a gain.

¹¹The average individual relative ambiguity premium is different from the aggregate relative ambiguity premium.

panel A. The first column of each panel presents for each pair of tasks the fraction of consistent preferences, i.e., the fraction of subjects that exhibit the same ambiguity attitude for two given ambiguity measures. The second column presents the difference between the observed fraction of consistent answers and the fraction of consistent answers that would be generated if the measured ambiguity preferences were completely random. In the binary classification, this random benchmark equals 50%; when using three categories the benchmark equals 33.3%. Only if the fraction of consistent answers is statistically higher than the random benchmark, two elicitation methods show some individual consistency.¹² Furthermore, we perform a Fisher (1922) test of association between each pair of elicitation methods. If two measures are individually consistent, a Fisher (1922) test of no association should be rejected.

The table shows that the various elicitation methods used to measure the individuals' ambiguity attitude, broadly defined, tend to be consistent with each other. In all but two pair-wise comparisons, the difference between the observed consistency and the random benchmark is positive. Yet, not for all pairs the difference is statistically significant. In the binary classification, the highest consistency in preferences can be observed between the ambiguity attitudes measured by task 2 and the dynamic Ellsberg thought experiment (69.8%). The individual consistency between the two incentivized tasks is almost equally high (68.6%), followed by the combination of task 1 (Chakravarty and Roy, 2009) and the dynamic Ellsberg thought experiment (64.8%). When ambiguity neutrality is included as a category on its own, see panel B, the static and dynamic Ellsberg thought experiments show the highest individual consistency (60.3%), followed by the three pairs of elicitation methods that are most consistent in the binary classification.

The Fisher (1922) test offers another perspective on the consistency of the various elicitation methods. While the results are similar for the binary classification, there are substantial differences in panel B that differentiates between three classes of ambiguity attitudes since the Fisher (1922) test distinguishes also between different types of non-consistent outcomes. As a result, even pairs of elicitation methods that score high on individual consistency in terms of the fraction of consistent classifications, a Fisher test cannot reject the hypothesis that the two elicitation methods generate completely independent ambiguity classifications. Most important, the classification into ambiguity seeking, ambiguity neutral, and ambiguity averse decision makers obtained from task 1 is statistically independent from the results of ambiguity task 2.

Taken together, the analysis shows that most of the different approaches to measure the overall ambiguity attitude yield rather similar results. Perhaps not surprisingly, the highest consistency can be observed between the rather similar tasks derived from the original Ellsberg thought experiment, involving choices between risky and ambiguous urn lotteries. In contrast, ambiguity preferences obtained from the complementary temperature bets and the horse bet setting exhibit, on average, rather low levels of individual consistency.

Another possibility to measure the consistency of the ambiguity preferences obtained from the various elicitation methods is to calculate their correlation with each other. The correlation matrix of table 7 is directly calculated from the switching points (incentivized tasks) and answers

¹²The random benchmark refers to an entirely random classification into the different ambiguity preference groups, not to completely random choices in the tasks and thought experiments. An entirely random choice pattern in the tasks would not even generate single switching points in most cases.

(non-incentivized thought experiments).¹³ Apart from a few cases, the correlation between the various ambiguity measures is positive. Yet, only very few correlations are positively significant. Interestingly, although the overall preference pattern of both incentivized tasks is very similar (see table 5), the correlation table reveals that individual ambiguity preferences elicited from task 1 and task 2 are very different. The correlation of switching points is very small, reaching just about 12%, without being statistically significant. Furthermore, the correlation matrix shows that the ambiguity preferences obtained from the thought experiments are only little correlated with each other, and similarly exhibit only a low correlation with the switching points from the tasks. Only the static and dynamic Ellsberg thought experiments show more than one significant correlation with some other ambiguity measure.¹⁴

As such, the correlation matrix confirms the previous results. Across the various ambiguity measures, it is again mostly the Ellsberg type of urn experiments that are significantly related with each other. The results of this section show that commonly used elicitation methods of ambiguity preferences yield only similar results when looking at broad categories of ambiguity attitudes, such as ambiguity aversion and ambiguity love. In contrast, they fail to accurately capture finer degrees of ambiguity preferences. Very clearly, different mechanisms to measure ambiguity preferences result in partly very different results, i.e., ambiguity preferences depend very much on the experimental design. As a consequence, the implications drawn from any ambiguity experiment, incentivized or not, are also likely to depend on the experimental design. Furthermore, while some thought experiments are lead to similar ambiguity preferences compared to standard incentivized experiments, this section shows that not all thought experiment are equally good to measure ambiguity preferences in the sense of economic experiments. Especially the willingness to pay for complementary temperature bets following Fox and Tversky (1995) does not exhibit any association with experimental measures of ambiguity preferences.

3.4 Time consistency

The previous section shows that, at the individual level, measures of ambiguity preferences depend very much on the method and experimental design used to elicit preferences. While some ambiguity measures are highly related to each other, some other measures exhibit even a negative association. Especially the low correlation between the switching points obtained from the two incentivized experiments questions the reliability of the various methods to elicit ambiguity preferences.

There are two explanations for these findings. First, the wording, context and answer possibilities of the various techniques differs. Even if there exists something like a fundamental or universal ambiguity preference, such framing effects can cause differences in the measured ambiguity attitudes.¹⁵ Second, and more important, all elicitation methods considered are subject to measurement errors. A perfect correlation between the various measures cannot possibly be expected, since some of the tasks are rather complex resulting in measurement errors.

¹³The correlation table is calculated such that the measured ambiguity preferences are comparable. A high number indicates ambiguity aversion for all elicitation methods.

¹⁴In addition, there is also a substantial correlation between the two measures of ambiguity premia obtained from the temperature bets and the two ambiguity preferences from in the horse betting framework.

¹⁵Beauchamp et al. (2012) find similar results when measuring risk aversion.

In this section we analyze the reliability of the various elicitation methods by examining their consistency over time. The conjecture is that elicitation methods that exhibit less measurement errors should lead to more consistent results when subjects are confronted with the same mechanism again. To analyze the time consistency, and hence accuracy, of the elicitation methods, we invited a sub-sample of subjects to participate in the experiments twice. This allows us to obtain two measures of ambiguity preferences for each of the ambiguity mechanisms. 26 subjects participated in the experiment twice, around 5 month after their first participation.¹⁶

Table 8 compares the ambiguity preferences obtained from the same subjects at two points in time. The first two columns report the correlation and p-values of each of the measured ambiguity preferences, the last two columns the explained variance (R^2 and adjusted R^2) when regressing the second observation of ambiguity preferences on the first observation.

The table shows that the time consistency of the measured ambiguity preferences is not very pronounced. Several measures even exhibit a negative correlation over time. Especially the complementary temperature bets as proposed by Fox and Tversky (1995) to measure ambiguity does not produce consistent preference measures, and might therefore not generate very reliable ambiguity measures. Only the ambiguity preferences obtained from the hypothetical horse betting question show some significant time consistency at the 10% level. As it is the simplest of the four thought experiments, this suggests that task complexity and measurement error are positively related. In direct comparison of the two experimental measures of ambiguity, it seems that the ambiguity experiment 2 generates slightly smaller measurement errors as the preferences obtained are more consistent over time. Yet, the sample size of this extension to the main study is rather small.

3.5 Prediction of experimental ambiguity measures

Although there are many techniques to measure human preferences, including incentivized tasks, non-incentivized thought experiments as well as survey questionnaires, the experimental economics literature advocates the use of incentivized laboratory experiments. While this method is recognized as best practice in experimental research, this procedure is sometimes very difficult to implement (requiring a laboratory) and expensive (payments to participants). Especially in field experiments, incentivized experiments face serious challenges due to the high administrative costs and the relatively complex design of experimental games.

To facilitate the measurement of fundamental human preferences, an emerging literature is concerned with finding simpler and more practical methods to measure preferences. These studies identify a set of simple survey questions and non-incentivized thought experiments that allow for a preference measurement equally reliable as elaborate laboratory experiments.¹⁷ In this section, we similarly examine whether it is possible to use non-incentivized thought experiments and simple survey questions to predict or proxy for standard experimental measures of ambiguity preferences. Since it is a priori not possible to tell which out the ambiguity preferences obtained from the

¹⁶14 of these subjects participated in a pilot study that took place several months before the main study. The other 12 subjects participated in both experiments in May and October 2013. Given some modifications of the experimental design, only a subset of ambiguity measures was available in both the pilot experiment and the main study. However and most important, both incentivized experiments remained unchanged.

¹⁷Examples are by Fehr et al. (2003), Dohmen et al. (2011), Hardeweg et al. (2011) and Vischer et al. (2011).

two incentivized tasks truly reflects individual ambiguity preferences, we first create a synthetic ambiguity preference measure using both tasks. This joint measure then serves as our experimental benchmark of ambiguity preferences. To this end, we calculate the average switching points AMB_s of the two experiments. This procedure allows also to reduce the measurement error of each of the experimental measures. Since the relative attractiveness of the risky urn changes in opposite directions, we need to adjust the original switching points. Furthermore, we need to correct for the number of situations in the two experiments:

$$AMB_s = \frac{1}{2} \left(\frac{\text{switching point experiment 1}}{10} + \frac{11 - \text{switching point experiment 2}}{11} \right)$$

The higher the switching point, the more ambiguity averse are the subjects. Since AMB_s requires a unique switching point in both tasks, AMB_s is only available for 102 subjects.

As a first assessment of the quality of the synthetic ambiguity measure, we evaluate the time consistency of the measured ambiguity preferences, similar to the previous section. In untabulated calculations, we find that the synthetic ambiguity measure exhibits a correlation over time of 35.6% (p-value: 0.106), which is higher than each of the two experimental measures, see table 8. Similarly, a regression of the second observation of AMB_s on the first observation yields a R^2 of 11.95%, which is also higher than each of the tasks. Hence, as conjectured, combining the two experimental ambiguity measures reduces the measurement error.

We next analyze whether it is possible to use four non-incentivized thought experiments to predict individual ambiguity preferences obtained from the tasks. Put differently, we seek to identify the thought experiments that explain the highest share of variance in the synthetic experimental ambiguity measure. In simple OLS regressions, this corresponds to identifying the combination of predictor variables that maximizes the adjusted R^2 . We adopt the following procedure. First, we identify the set of predictors that is positively correlated with AMB_s , the combined measure of experimental ambiguity preferences. Second, for a given number of predictor variables, we estimate the linear regression model using every possible combination of thought experiments as predictor variables. Finally, we select the best combination of predictors in terms of the adjusted R^2 .¹⁸

The results are shown in table 9. The best predictor for the experimental measure of ambiguity preferences is the original Ellsberg (1961) thought experiment, explaining almost 11% of the variance. When adding the thought experiment that asks subjects for their risk equivalent to an ambiguous horse bet, the adjusted R^2 increases from 9.7% to 13.6%. Although the strict preference between a risky and an ambiguous horse bet is not significant, it still slightly increases the adjusted R^2 . The remaining two thought experiments further increase the explanatory power of the regression, but they fail to improve the adjusted R^2 . Hence, the maximum adjusted R^2 is obtained by using just three thought experiments.

In an extension of this analysis, we also include survey questions as potential predictors for individual ambiguity attitudes. Since these items (see appendix C) are taken from the psychology literature, it is a priori not clear whether they measure ambiguity attitudes in a similar way

¹⁸The results are unchanged if we use the Akaike information criterion, AIC (Akaike, 1974) or Bayesian information criterion, BIC (Schwarz, 1978) as regression model selection criteria. All predictors of the selected regression models are jointly significant using an F-test. Furthermore, the results are identical when using an ordered probit regression model instead of OLS, using the pseudo R^2 as model selection criterion.

as the experimental economics literature. In fact, out of the 46 survey items, only 6 items show some weakly significant correlation with the synthetic benchmark of experimental ambiguity preferences, see table 10. Since one item (item number 11) has the wrong correlation with the experimental preference measure, we discard this survey questions.¹⁹ Two survey items, items 7 and 33, however show a strong correlation with the experimental preference measure.

We add these 6 survey items to the pool of explanatory variables to predict the experimental ambiguity preferences. That is, we repeat the analysis of the previous section, using now a set of 11 explanatory variables. Again, we then identify the combination of predictor variables that maximizes the adjusted R^2 . The results are shown in table 11. The survey items increase the explained variance of the regressions significantly. Although the original Ellsberg (1961) thought experiment and the thought experiment that asks subjects for their risk equivalent to an ambiguous horse bet remain the two best predictors, survey items 45, 33, 38 and 7 further increase the explained variance by more than 8%. The optimal set of predictor variables includes four thought experiments and two survey questions.²⁰

From a practical perspective, how can we assess the quality of prediction of experimental ambiguity preferences using non-incentivized thought experiments and survey questions? At first sight, an explained variance of 23% seems very low, as this implies that the predictors together fall obviously short from exactly predicting experimental ambiguity measures. Yet, even when repeating both incentivized experiments with the same participants some time later, the obtained synthetic ambiguity preferences only correlate 33% with each other. Hence, the predictive regressions should not be benchmarked against an explained variance of one, but rather against a correlation of 33% or a R^2 of 10.8%, the explained variance when regression the two synthetic ambiguity measures on each other, as this level of would be equally good as the experiment itself, once taking measurement errors into account. In this light, the selected thought experiments and survey questions are slightly better in measuring ambiguity preferences than the experiment itself.²¹

4 Concluding remarks

Following Ellsberg’s seminal thought experiment, the literature has proposed many different task designs to measure ambiguity preferences. These tasks are based on incentivized laboratory experiments, non-incentivized thought experiments and survey questionnaires. This study compares individual estimates of ambiguity preferences obtained from a large variety of elicitation tasks proposed in the literature.

We find a high degree of aggregate consistency across tasks closely in line with the results from previous literature using similar elicitation methods. However, individual consistency is more

¹⁹Subjects were asked to respond to the statements by indicating the extent to which you agree or disagree with them on a scale from 1 (I strongly agree) to 7 (I strongly disagree). Hence, a high value in the question “Perfect balance is the essence of all good composition.” rather implies ambiguity-seeking preferences than ambiguity-averse preferences.

²⁰Adding other survey questions or thought experiments increases the explanatory power of the regression, but not the adjusted R^2 .

²¹It should be noted that the second experiment took place five months after the first experiment, see section 3.4. Furthermore, the sample size of the retest is much smaller. Hence, a direct comparison of the correlation measures and R^2 is not possible.

volatile: while we find that subjects are in general more consistent than one would expect under random choice, subjects are not fully consistent in their responses across tasks. In particular, some of the most popular thought experiments in the literature on ambiguity have no relation to the behavior observed in incentivized experiments. While the concept of ambiguity tolerance in psychology is partly distinct from the notion of ambiguity in the economics literature, a selection of survey questions widely-used in psychology correlates with ambiguity attitudes elicited using incentivized games. We identify two simple thought experiments and four survey questions which explain about 25% of the variation in individual ambiguity preferences. Taken together, these items might offer a reliable alternative to measure ambiguity preferences when running experiments is unfeasible.

The results of this study may have some important implications. On the one hand, the study highlights the importance of carefully selecting the experimental design when estimating ambiguity preferences. Since the concept of ambiguity is more difficult to grasp compared to other preference dimensions, such as risk, the elicitation of ambiguity attitudes is more prone to errors and therefore more challenging for a researcher.

On the other hand, the results indicate that it is possible to estimate ambiguity preferences without relying on laboratory experiments, at least to some extent. As such, this insight makes it easier to measure ambiguity preferences in large-scale field surveys. There might be also some applications for practitioners. While simple risk assessment questions are standard when advising banking clients about their optimal investment portfolio, the clients' ambiguity attitudes have been left unexplored because of the lack of simple assessment methods. With the help of simpler ambiguity measures, private banks can improve the personality assessment of their clients to tailor better asset allocation strategies.

Appendix A: Incentivized experiments

This appendix describes the three incentivized experiments in detail. Before each experiment, subjects were presented examples to familiarize them with the design of the experiment. In addition, subjects were asked several control questions to ensure that they understood the payoff structure of each experiment.

Ambiguity experiment 1, Chakravarty and Roy (2009): This experiment is taken from decision sheet D of Chakravarty and Roy (2009).

This task involves a decision table with 10 situations. Each situation offers you a choice between drawing a ball from two different urns, urn X or urn Y. Both urns contain 10 balls, either white or black.

- *Urn X: You know the composition of urn X. It contains 5 white balls and 5 black balls.*
- *Urn Y: In contrast you don't know the composition of urn Y. All you know is that it may contain either 10 white balls or 10 black balls.*

Note that the composition of the urns does not change from one situation to the next. The points you can earn depend on the color of the ball drawn. Only one color yields some points. You can choose whether the color that yields points is white or black. For example, suppose you chose black. In this case, you earn some points if the ball drawn is black, but you don't earn any points if the ball drawn is white. Please choose the color of the ball that provides you points:

- *white*
- *black*

Please look at the decision table below.²² In each of the 10 situations, we would like you to indicate from which urn (urn X or urn Y) you prefer drawing a ball. If white ball is drawn, you earn the points indicated in the table. If a black ball is drawn, you earn no points. The points you earn when drawing a white ball from urn X change from one situation to the next. However, the points you earn when drawing a white ball from urn Y is always 10.

At the end of the session, the computer will randomly select one out of the 10 situations. Then, depending on whether you have chosen urn X or urn Y in that situation, the computer will randomly draw one ball from that urn. Depending on the color of the ball, you earn the points indicated in the table. Notice that even though you will make 10 decisions, only one of these will determine the points you earn, but you will not know in advance which situation will be selected (they are equally likely to be selected).

In each situation, from which urn do you prefer to draw a ball, urn X or urn Y?

²²The actual decision table presented to the subjects depends on the color chosen. In this appendix, we assume that the selected color is white. If the selected color is black, the word “white” has to be replaced with “black”, and vice versa.

	URN X: Composition: 5 white balls, 5 black balls	URN Y: Composition: 10 white balls OR 10 black balls	Your choices
1	If a white ball is drawn you earn 14 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
2	If a white ball is drawn you earn 13 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
3	If a white ball is drawn you earn 12 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
4	If a white ball is drawn you earn 11 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
5	If a white ball is drawn you earn 10 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
6	If a white ball is drawn you earn 9 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
7	If a white ball is drawn you earn 8 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
8	If a white ball is drawn you earn 7 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
9	If a white ball is drawn you earn 4 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y
10	If a white ball is drawn you earn 2 points	If a white ball is drawn you earn 10 points	Urn X <input type="radio"/> <input type="radio"/> Urn Y

Ambiguity experiment 2: This experiment is an extension of the Ellsberg (1961) thought experiment, similar to Lauriola and Levin (2001) and Butler et al. (2013).

In this task, we present you a decision table with 11 situations. Each situation offers you a choice between drawing a ball from two different urns, urn 1 or urn 2. Both urns contain 10 balls, either white or black.

- *Urn 1: The composition of urn 1 changes from one situation to the next. While the number of balls in one color (e.g., white) increases incrementally from 0 to 10, the number of balls of the other color (e.g., black) decreases accordingly.*
- *Urn 2: The composition of urn 2 is identical in each situation. However, you don't know how many balls are white and how many balls are black. Any combination is possible. There might be from 0 to 10 white balls, with the remaining balls being black.*

One ball will be drawn from the urn you choose. The points you can earn depend on the color of the ball drawn. Only one color yields some points. You can choose whether the color that yields points is white or black. Please choose the color of the ball that provides you points:

- *white*
- *black*

Please look at the decision table below.²³ In each of the 11 situations, we would like you to indicate from which urn (urn 1 or urn 2) you prefer drawing a ball. As explained before, both urns contain 10 balls, either white or black.

- *Urn 1: The composition of urn 1 changes from one situation to the next. The number of white balls increases incrementally from 0 white balls in situation 0 to 10 white balls in situation 10, while the number of black balls decreases accordingly.*
- *Urn 2: The composition of urn 2 is identical in all situations. However, the exact composition of urn 2 is unknown. Any combination of white and black balls is possible: there might be 10 white balls, or 10 black balls, or any other possible combination of white and black balls.*

²³The actual decision table presented to the subjects depends on the color chosen. In this appendix, we assume that the selected color is white. If the selected color is black, the word “white” has to be replaced with “black”, and vice versa.

If a white ball is drawn, you earn 10 points. If a black ball is drawn, you earn no points.

At the end of the session, the computer will randomly select one out of the 11 situations. Then, depending on whether you have chosen urn 1 or urn 2 in that situation, the computer will randomly draw one ball from that urn. Depending on the color of the ball, you earn the points indicated in the table. Notice that even though you will make 11 decisions, only one of these will determine the points you earn, but you will not know in advance which situation will be selected (they are equally likely to be selected).

In each situation, from which urn do you prefer to draw a ball, urn 1 or urn 2?

	URN 1: If a white ball is drawn you earn 10 points	URN 2: If a white ball is drawn you earn 10 points	Your choices
0	0 white balls, 10 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
1	1 white ball, 9 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
2	2 white balls, 8 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
3	3 white balls, 7 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
4	4 white balls, 6 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
5	5 white balls, 5 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
6	6 white balls, 4 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
7	7 white balls, 3 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
8	8 white balls, 2 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
9	9 white balls, 1 black ball	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2
10	10 white balls, 0 black balls	unknown composition	Urn 1 <input type="radio"/> <input type="radio"/> Urn 2

Risk experiment (Chakravarty and Roy, 2009): This experiment is taken from decision sheet B of Chakravarty and Roy (2009).

In this task you need to fill in the decision table shown below. The decision table consists of 10 different situations, listed 1 to 10. Each situation offers you a choice between drawing a ball from two different urns, urn A or urn B. Both urns contain 10 balls, either white or black.

- *The composition of urn A is identical in all 10 situations. There are 5 white balls and 5 black balls.*
- *The composition of urn B changes from one situation to the next. The number of white balls increases incrementally from 0 white balls in situation 1 to 9 white balls in situation 10, while the number of black balls decreases accordingly.*

At the end of the session, the computer will randomly select one out of the 10 situations. Then, depending on whether you have chosen urn A or urn B in that situation, the computer will randomly draw one ball from that urn. Depending on the color of the ball, you earn the points indicated in the table. Notice that even though you will make 10 decisions, only one of these will determine the points you earn, but you will not know in advance which situation will be selected (they are equally likely to be selected).

In each situation, from which urn do you prefer to draw a ball, urn A or urn B?

	URN A: If a white ball is drawn you earn 6 points If a black ball is drawn you earn 4 points	URN B: If a white ball is drawn you earn 10 points If a white ball is drawn you earn 0 points	Your choices
1	5 white balls, 5 black balls	0 white balls, 10 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
2	5 white balls, 5 black balls	1 white ball, 9 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
3	5 white balls, 5 black balls	2 white balls, 8 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
4	5 white balls, 5 black balls	3 white balls, 7 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
5	5 white balls, 5 black balls	4 white balls, 6 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
6	5 white balls, 5 black balls	5 white balls, 5 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
7	5 white balls, 5 black balls	6 white balls, 4 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
8	5 white balls, 5 black balls	7 white balls, 3 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
9	5 white balls, 5 black balls	8 white balls, 2 black balls	Urn A <input type="radio"/> <input type="radio"/> Urn B
10	5 white balls, 5 black balls	9 white balls, 1 black ball	Urn A <input type="radio"/> <input type="radio"/> Urn B

Appendix B: Non-incentivized thought experiments

Ellsberg urn experiment

Please imagine the following situation: You can choose between drawing a ball from two different urns, urn A and urn B. Urn A contains 100 balls, some are white and some are black. However, you don't know how many balls are white and how many balls are black. Any combination is possible. There might be from 0 to 100 white balls, with the remaining balls being black. Urn B contains 100 balls as well, but you know that it contains exactly 50 white balls and 50 black balls. Now choose a color, either white or black. Suppose you win £100 if you draw a ball of the color you have selected. If the ball is of the other color, you win nothing. From which urn would you prefer drawing a ball?

- 1. I have a strong preference to draw a ball from urn A.*
- 2. I have a slight preference to draw a ball from urn A.*
- 3. I am indifferent between drawing a ball from urn A or from urn B.*
- 4. I have a slight preference to draw a ball from urn B.*
- 5. I have a strong preference to draw a ball from urn B.*

Horse betting

Please imagine the following situation: Suppose two friends invite you to accompany them to a horse race. Both are experts in horse races. Your friends agree that the probability of horse A winning the race is 30%. In contrast, they disagree on the chances of horse B. One friend suggests that the probability of horse B winning is 20%, the other friend suggests the probability of horse B winning is 40%. If you had to bet on one of the horses, which horse would you rather bet on?

- 1. Horse A*
- 2. Horse B*

Suppose there is a third horse C. Which probability of horse C winning would make a bet on horse C equally attractive to you as betting on horse B? In other words, we ask you to select the probability of horse C winning such that you are indifferent between betting on horse B or on horse C. Remember that your friends do not agree on the chances of horse B. One friend suggests that the probability of horse B winning is 20%, the other friend suggests the probability of horse B winning is 40%. Please select the probability of horse C winning such that you are indifferent between betting on horse B or on horse C. Insert a probability between 20% and 40% in the box below.

Pricing complementary temperature bets

Imagine you are offered some lottery tickets whose payoff depends on the maximum temperature in a specific city, on a specific day in the future. There are two types of such tickets. The first type of ticket pays you £100 if the maximum temperature is above a certain level. If the maximum temperature is below that level, the ticket pays you nothing. The second type of ticket pays you £100 if the maximum temperature is below a certain level. If the maximum temperature is above that level, the ticket pays you nothing.

Example of a lottery ticket of type 1: A ticket that pays you £100 if the maximum temperature in Paris is at least 20 degrees Celsius on the 01/07/13, and nothing if the maximum temperature is below 20 degrees Celsius.

On the next two screens you are offered four temperature lottery tickets. We would like to know the highest price you would be willing to pay for each of the tickets.

- 1. Imagine you have been offered a lottery ticket that pays you £100 if the maximum temperature in London is at least 15 degrees Celsius exactly one week from today, and nothing if the temperature is below 15 degrees Celsius. What is the highest price you would be willing to pay for such a ticket? (in £)*
- 2. Imagine you have been offered a lottery ticket that pays you £100 if the maximum temperature in London is below 15 degrees Celsius exactly one week from today, and nothing if the temperature is above 15 degrees Celsius. What is the highest price you would be willing to pay for such a ticket? (in £)*
- 3. Imagine you have been offered a lottery ticket that pays you £100 if the maximum temperature in Vancouver is at least 15 degrees Celsius exactly one week from today, and nothing if the temperature is below 15 degrees Celsius. What is the highest price you would be willing to pay for such a ticket? (in £)*
- 4. Imagine you have been offered a lottery ticket that pays you £100 if the maximum temperature in Vancouver is below 15 degrees Celsius exactly one week from today, and nothing if the temperature is above 15 degrees Celsius. What is the highest price you would be willing to pay for such a ticket? (in £)*

Dynamic Ellsberg urn experiment

Please imagine the following situation: You can choose between drawing a ball from two different urns, urn A and urn B. Urn A contains 100 balls, some are white and some are black. However, you don't know how many balls are white and how many balls are black. Any combination is possible. There might be from 0 to 100 white balls, with the remaining balls being black. Put differently, you do not know the probability of drawing a white or a black ball. Urn B contains 100 balls as well, but you know the exact number of white and black balls in this urn. In other words, you know the exact probability of drawing a white or a black ball. Now choose a color, either white or black. Suppose you win £100 if you draw a ball of the color you have selected. If the ball is of the other color, you win nothing. Please choose the color of the ball that provides you £100:

- white*
- black*

We present you now several situations.²⁴ The composition of urn A is identical in each situation. The composition of urn B is different in each situation.

²⁴The actual situations presented to the subjects depend on the color chosen. In this appendix, we assume that the selected color is white. If the selected color is black, the word “white” has to be replaced with “black”, and vice versa.

1. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 45 white balls and 55 black balls, i.e., there is a 45% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
 - (a) Urn A \rightarrow go to situation 2
 - (b) Urn B \rightarrow go to situation 3
2. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 65 white balls and 35 black balls, i.e., there is a 65% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
 - (a) Urn A \rightarrow go to situation 4
 - (b) Urn B \rightarrow go to situation 5
3. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 25 white balls and 75 black balls, i.e., there is a 25% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
 - (a) Urn A \rightarrow go to situation 6
 - (b) Urn B \rightarrow go to situation 7
4. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 75 white balls and 25 black balls, i.e., there is a 75% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
 - (a) Urn A \rightarrow go to situation 8
 - (b) Urn B \rightarrow go to situation 9
5. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 55 white balls and 45 black balls, i.e., there is a 55% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
 - (a) Urn A \rightarrow go to situation 10
 - (b) Urn B \rightarrow go to situation 11
6. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 35 white balls and 65 black balls, i.e., there is a 35% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
 - (a) Urn A \rightarrow go to situation 12
 - (b) Urn B \rightarrow go to situation 13
7. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 15 white balls and 85 black balls, i.e., there is a 15% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.

- (a) Urn A \rightarrow go to situation 14
 - (b) Urn B \rightarrow go to situation 15
8. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 80 white balls and 20 black balls, i.e., there is a 80% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
- (a) Urn A \rightarrow go to situation 16
 - (b) Urn B
9. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 70 white balls and 30 black balls, i.e., there is a 70% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
- (a) Urn A
 - (b) Urn B
10. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 60 white balls and 40 black balls, i.e., there is a 60% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
- (a) Urn A
 - (b) Urn B
11. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 50 white balls and 50 black balls, i.e., there is a 50% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
- (a) Urn A
 - (b) Urn B
12. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 40 white balls and 60 black balls, i.e., there is a 40% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
- (a) Urn A
 - (b) Urn B
13. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 30 white balls and 70 black balls, i.e., there is a 30% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.
- (a) Urn A
 - (b) Urn B
14. What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 20 white balls and 80 black balls, i.e., there is a 20% chance of

drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.

(a) Urn A

(b) Urn B

15. *What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 10 white balls and 90 black balls, i.e., there is a 10% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.*

(a) Urn A

(b) Urn B → go to situation 17

16. *What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 90 white balls and 10 black balls, i.e., there is a 90% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.*

(a) Urn A

(b) Urn B

17. *What would you prefer? Drawing a ball from urn A with unknown composition, i.e., you do not know the probability of drawing a white ball or drawing a ball from urn B in which there are 5 white balls and 95 black balls, i.e., there is a 5% chance of drawing a white ball? Remember, you win £100 if you draw a white ball, and nothing otherwise.*

(a) Urn A

(b) Urn B

Appendix C: Survey questionnaire

The survey or self-assessment questions to elicit ambiguity preferences are mostly taken from self-reporting scales in the psychology literature.

In this part, we present you a list of statements. Please indicate the extent to which you agree or disagree with them. Please do not spend too much time on each statement. There are no right or wrong answers and therefore your first response is important. Nevertheless, try to be as honest as you can be. Answer according to your own feelings, rather than how you think most people would answer. Don't worry about being consistent in your responses. Be sure to answer every statement.

Please respond to the following statements by indicating the extent to which you agree or disagree with them on a scale from 1 (I strongly agree) to 7 (I strongly disagree).

Intolerance of Ambiguity Scale by Kirton (1981). Items based on Mac Donald Jr. (1970) and Rydell and Rosen (1966):

- 1 *There's a right way and a wrong way to do almost everything.*
- 2 *Practically every problem has a solution.*
- 3 *I have always felt that there is a clear difference between right and wrong.*
- 4 *Nothing gets accomplished in this world unless you stick to some basic rules.*
- 5 *If I were a doctor, I would prefer the uncertainties of a psychiatrist to the clear and definite work of someone like a surgeon or a x-ray specialist.*

- 6 *Vague and impressionistic pictures really have little appeal for me.*
- 7 *Before an examination, I feel much less anxious if I know how many questions there will be.*
- 8 *The best part of a jigsaw puzzle is putting in that last piece.*
- 9 *I don't like to work on a problem unless there is a possibility of coming out with a clear-cut and unambiguous answer.*
- 10 *I like to fool around with new ideas, even if they turn out later to be a total waste of time.*
- 11 *Perfect balance is the essence of all good composition.*

Items based on Budner (1962):

- 12 *An expert who doesn't come up with a definite answer probably doesn't know too much.*
- 13 *There is really no such thing as a problem that can't be solved.*
- 14 *A good job is one where what is to be done and how it is to be done are always clear.*
- 15 *In the long run it is possible to get more done by tackling small, simple problems rather than large and complicated ones.*
- 16 *What we are used to is always preferable to what is unfamiliar.*
- 17 *A person who leads an even, regular life in which few surprises or unexpected happenings arise, really has a lot to be grateful for.*
- 18 *I like parties where I know most of the people more than the ones where all or most of the people are complete strangers.*

Item from the Tolerance of Ambiguity Scale by Budner (1962):

- 19 *I would like to live for a while in a foreign country that is new to me.*

Items from the Ambiguity Tolerance Scale by Norton (1975):

- 20 *In a decision-making problem in which there is not enough information to process the problem, I feel very uncomfortable.*
- 21 *I am tolerant of ambiguous situations.*
- 22 *Vague and impressionistic pictures appeal to me more than realistic pictures.*
- 23 *I like movies or stories with definite endings.*
- 24 *The best part about reading a poem is then being able to read a commentary explaining the poem's meanings.*

Items from the Extended Life Orientation test by Chang et al. (1997):

- 25 *In uncertain times, I usually expect the best.*
- 26 *When I undertake something new, I expect to succeed.*
- 27 *If something can go wrong for me, it will.*
- 28 *I rarely count on good things happening to me.*

Items from the Uncertainty Response Scale by Greco and Roger (2001):

- 29 *When making a decision, I am deterred by the fear of making a mistake.*
- 30 *When a situation is uncertain, I generally expect the worst to happen.*

- 31 *I find the prospect of change exciting and stimulating.*
- 32 *I enjoy unexpected events.*
- 33 *The idea of taking a trip to a new country fascinates me.*
- 34 *Before making any changes, I need to think things over thoroughly.*
- 35 *I feel relieved when an ambiguous situation suddenly becomes clear.*
- 36 *When uncertain, I act very cautiously until I have more information about the situation.*
- 37 *I prefer to stick to tried and tested ways of doing things.*

Items from the Ambiguity Tolerance Scale II by McLain (2009):

- 38 *I try to avoid situations that are ambiguous.*
- 39 *I avoid situations that are too complicated for me to easily understand.*
- 40 *I generally prefer novelty over familiarity.*
- 41 *I find it hard to make a choice when the outcome is uncertain.*

Single-item measure of self-esteem by Robins et al. (2001)²⁵:

- 42 *I have high self-esteem.*

Other items (own additions):

- 43 *I like novelties.*
- 44 *I voluntarily accept new challenges.*
- 45 *When a situation is uncertain, I never take action until I know all the risks involved.*
- 46 *Do you consider yourself as a pessimist or an optimist?*

²⁵Similar to Robins et al. (2001), this item uses a 5 point answer scale.

Appendix D: Sample description

The table of this appendix summarizes our sample of participants.

Table 1: Summary of sample characteristics

Panel A: Gender		
	Observations	Percentage
male	54	44.6%
female	67	55.4%
Panel B: Marital status		
	Observations	Percentage
single	112	92.6%
married	5	4.1%
divorced	3	2.5%
widowed	1	0.8%
Panel C: Age		
	Observations	Percentage
up to 20	11	9.1%
21 - 25	59	48.8%
26 - 30	29	24.0%
over 30	22	18.2%
Panel D: Nationality		
	Observations	Percentage
United Kingdom	45	37.2%
Italy	12	9.9%
Germany	9	7.4%
China	7	5.8%
Poland	6	5.0%
other countries	42	34.7%
Panel E: Main field of studies		
	Observations	Percentage
Politics and International Relations	14	11.6%
Economics	11	9.1%
Business Studies	10	8.3%
Modern Languages and Cultures	10	8.3%
Psychology	9	7.4%
Development Studies	9	7.4%
Humanities	9	7.4%
other subjects	49	40.5%

References

- Abdellaoui, Mohammed, Aurélien Baillon, Laetitia Placido and Peter P. Wakker (2011), ‘The rich domain of uncertainty: Source functions and their experimental implementation’, *American Economic Review* **101**(2), 695–723.
- Akaike, Hirotugu (1974), ‘A new look at the statistical model identification’, *IEEE Transactions on Automatic Control* **19**(6), 716–723.
- Akay, Alpaslan, Peter Martinsson, Haileselassie Medhin and StefanT. Trautmann (2012), ‘Attitudes toward uncertainty among the poor: an experiment in rural ethiopia’, *Theory and Decision* **73**(3), 453–464.
- Beauchamp, Jonathan P., Daniel J. Benjamin, Christopher F. Chabris and David I. Laibson (2012), How malleable are risk preferences and loss aversion? working paper.
- Binmore, Ken, Lisa Stewart and Alex Voorhoeve (2012), ‘How much ambiguity aversion? finding indifference between ellsberg’s risky and ambiguous bets’, *Journal of Risk and Uncertainty* **45**(3), 215–238.
- Budner, S. (1962), ‘Intolerance for ambiguity as a personal variable’, *Journal of Personality* **30**, 29–50.
- Butler, Jeffrey V., Luigi Guiso and Tullio Jappelli (2013), The role of intuition and reasoning in driving aversion to risk and ambiguity. working paper.
- Chakravarty, Sujoy and Jaideep Roy (2009), ‘Recursive expected utility and the separation of attitudes towards risk and ambiguity: an experimental study’, *Theory and Decision* **66**(3), 199–228.
- Chang, Edward C., Albert Maydeu-Olivares and Thomas D’Zurilla (1997), ‘Optimism and pessimism as partially independent constructs: relationship to positive and negative affectivity and psychological well-being’, *Personality and Individual Differences* **23**(3), 433–440.
- Charness, Gary, Edi Karni and Dan Levin (2013), ‘Ambiguity attitudes and social interactions’, *Journal of Risk and Uncertainty* **46**, 1–25.
- Chateauneuf, Alain, Jürgen Eichberger and Simon Grant (2007), ‘Choice under uncertainty with the best and worst in mind: Neo-additive capacities’, *Journal of Economic Theory* **137**, 538–567.
- Chen, Zengjing and Larry Epstein (2002), ‘Ambiguity, risk, and asset returns in continuous time’, *Econometrica* **70**(4), 1403–1443.
- Chesson, HarrellW. and W.Kip Viscusi (2003), ‘Commonalities in time and ambiguity aversion for long-term risks*’, *Theory and Decision* **54**(1), 57–71.
- Chow, Clare Chua and Rakesh K Sarin (2002), ‘Known, unknown, and unknowable uncertainties’, *Theory and Decision* **52**, 127–138.
- Cornsweet, Tom N. (1962), ‘The staircase-method in psychophysics’, *The American Journal of Psychology* **75**(3), 485–491.

- Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp and Gert G. Wagner (2011), 'Individual risk attitudes: Measurement, determinants and behavioral consequences', *Journal of the European Economic Association* **9**(3), 522–550.
- Einhorn, Hillel J. and Robin M. Hogarth (1985), 'Ambiguity and uncertainty in probabilistic inference', *Psychological Review* **92**(4), 433–461.
- Ellsberg, D. (1961), 'Risk, ambiguity and the savage axioms', *Quarterly Journal of Economics* **75**(4), 643–669.
- Epstein, Larry G. and Martin Schneider (2008), 'Ambiguity, information quality and asset pricing', *Journal of Finance* **63**(1).
- Fehr, E., U. Fischbacher, B. Rosenbladt, J. Schupp and G. Wagner (2003), A nationwide laboratory: Examining trust and trustworthiness by integrating behavioural experiments into representative surveys. working paper.
- Fischbacher, Urs (2007), 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics* **10**(2), 171–178.
- Fisher, R. A. (1922), 'On the interpretation of χ^2 from contingency tables, and the calculation of p', *Journal of the Royal Statistical Society* **85**(1), 87–94.
- Fox, Craig R. and Amos Tversky (1995), 'Ambiguity aversion and comparative ignorance', *Quarterly Journal of Economics* **110**(3), 585–603.
- Gilboa, Itzhak and David Schmeidler (1989), 'Maxmin expected utility with non-unique prior', *Journal of Mathematical Economics* **18**, 141–153.
- Greco, Veronica and Derek Roger (2001), 'Coping with uncertainty: the construction and validation of a new measure', *Personality and individual Differences* **31**, 519–534.
- Guiso, Luigi and Monica Paiella (2008), 'Risk aversion, wealth and background risk', *Journal of the European Economic Association* **6**(6), 1109–1150.
- Hardeweg, B., L. Menkhoff and H. Waibel (2011), Experimentally-validated survey evidence on individual risk attitudes in rural thailand. working paper.
- Heath, Chip and Amos Tversky (1991), 'Preference and belief: Ambiguity and competence in choice under uncertainty', *Journal of Risk and Uncertainty* **4**(1), 5–28.
- Holt, Charles A. and Susan K. Laury (2002), 'Risk aversion and incentive effects', *American Economic Review* **92**(5), 1644–1655.
- Kirton, M. J. (1981), 'A reanalysis of two scales of tolerance of ambiguity.', *Journal of Personality Assessment* **45**(4), 407–414.
- Knight, Frank H. (1921), *Risk, Uncertainty, and Profit*, Houghton Mifflin Company, Boston, MA.
- Lauriola, Marco and Irwin P. Levin (2001), 'Relating individual differences in attitude toward ambiguity to risky choices', *Journal of Behavioral Decision Making* **14**(2), 107–122.
- Mac Donald Jr., A. P. (1970), 'Revised scale for ambiguity tolerance: reliability and validity', *Psychological Reports* **26**, 791–798.

- McLain, David L. (2009), ‘Evidence of the propoerties of an ambiguity tolerance measure: the multiple stimulus types ambiguity tolerance scale-ii (mstat-ii)’, *Psychological Reports* **105**(3), 975–988.
- Millner, Antony, Simon Dietz and Geoffrey Heal (2013), ‘Scientific ambiguity and climate policy’, *Environmental and Resource Economics* **55**(1), 21–46.
- Mukerji, Sujoy and Jean-Marc Tallon (2001), ‘Ambiguity aversion and incompleteness of financial markets’, *The Review of Economic Studies* **68**(4), 883–904.
- Norton, Robert W. (1975), ‘Measurement of ambiguity tolerance’, *Journal of Personality Assessment* **39**(6), 607–619.
- Olsen, Robert A. and George H. Troughton (2000), ‘Are risk premium anomalies caused by ambiguity?’, *Financial Analysts Journal* **56**(2), 24–31.
- Robins, Richard W., Holly M. Hendin and Kali H. Trzesniewski (2001), ‘Measuring global self-esteem: Construct validation of a single-item measure and the rosenberg self-esteem scale’, *Personality and Social Bulletin* **27**(2), 151–161.
- Rydell, S.T. and E. Rosen (1966), ‘Measurement and some correlates of need cognition’, *Psychological Reports* **19**, 139–165.
- Sarin, R. K. and M. Weber (1993), ‘Effects of ambiguity in market experiments’, *Management Science* **39**(5), 602–615.
- Savage, Leonard J. (1954), *The Foundations of Statistics*, Wiley, New York.
- Schwarz, Gideon E. (1978), ‘Estimating the dimension of a model’, *Annals of Statistics* **6**(2), 461–464.
- Spears, Dean (2013), ‘Poverty and probability: aspiration and aversion to compound lotteries in el salvador and india’, *Experimental Economics* **16**(3), 263–284.
- Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler and Stefan T. Trautmann (2013), ‘Impatience and uncertainty: Experimental decisions predict adolescents’ field behavior’, *The American Economic Review* **103**(1), 510–531.
- Trautmann (2013), *Blackwell Handbook of Judgment and Decision Making*, Blackwell, chapter Ambiguity Attitudes. forthcoming.
- Trautmann, Stefan T., Ferdinand M. Vieider and Peter P. Wakker (2011), ‘Preference reversals for ambiguity aversion’, *Management Science* **57**(7), 1320–1333.
- Vischer, T., T. Dohmen, A. Falk, D. Huffman, J. Schupp, U. Sunde and G. Wagner (2011), Measuring time preferences in representative samples by combining experimental elicitation and survey measures. working paper.
- Viscusi, W. Kip and Harrell Chesson (1999), ‘Hopes and fears: the conflicting effects of risk ambiguity’, *Theory and Decision* **47**, 153–178.
- Weitzman, Martin L. (2009), ‘On modeling and interpreting the economics of catastrophic climate change’, *Review of Economics and Statistics* **91**(1), 1–19.
- Wilcoxon, F. (1945), ‘Individual comparisons by ranking methods’, *Biometrics* **1**, 80–83.

Tables

Table 2: Descriptive statistics of incentivized ambiguity task 1

Panel A: Choices and switching points					
Risky choices			Switching point		
Risky choices	Observations	Fraction	Switching point	Observations	Fraction
0/10	1	0.8%	0	1	1.0%
1/10	0	0.0%	1	0	0.0%
2/10	3	2.5%	2	1	1.0%
3/10	2	1.7%	3	2	1.9%
4/10	32	26.5%	4	28	26.7%
5/10	26	21.5%	5	25	23.8%
6/10	13	10.7%	6	8	7.6%
7/10	22	18.2%	7	19	18.1%
8/10	17	14.1%	8	16	15.2%
9/10	3	2.5%	9	3	2.9%
10/10	2	1.7%	10	2	1.9%
Total	121	100.0%	Total	105	100.0%

Panel B: Summary statistics					
	Observations	Mean	Standard deviation	Lowest	Highest
Risky choices	121	56.6%	0.18%	0%	10%
Switching point	105	5.72	1.81	0	10

The table summarizes the results of the incentivized ambiguity task 1. In panel A, the three columns on the left report the number of situations subjects preferred drawing a ball from the risky urn over drawing a ball from the ambiguous urn. The three columns on the right report the switching points of the subjects. More precisely, it indicates the last urn before a subject switched from the risky to the ambiguity urn. A switching point of 0 means that the subject always preferred the ambiguous urn; a switching point of 10 means that the subject always preferred the risky urn. For a detailed description of the task, see appendix A. Panel B presents the summary statistics of the fraction of risky choices and the switching points.

Table 3: Descriptive statistics of incentivized ambiguity task 2

Panel A: Choices and switching points					
Risky choices			Switching point		
Risky choices	Observations	Fraction	Switching point	Observations	Fraction
0/11	0	0.0%	0	0	0.0%
1/11	0	0.0%	1	0	0.0%
2/11	0	0.0%	2	2	1.7%
3/11	0	0.0%	3	8	6.9%
4/11	7	5.8%	4	27	23.3%
5/11	26	21.5%	5	48	41.4%
6/11	50	41.3%	6	25	21.6%
7/11	28	23.1%	7	6	5.2%
8/11	8	6.6%	8	0	0.0%
9 /11	2	1.7%	9	0	0.0%
10/11	0	0.0%	10	0	0.0%
11/11	0	0.0%	11	0	0.0%
Total	121	100.0%	Total	105	100.0%

Panel B: Summary statistics					
	Observations	Mean	Standard deviation	Lowest	Highest
Risky choices	121	55.3%	0.10%	36.4%	81.8%
Switching point	116	4.90	1.04	2	7

The table summarizes the results of incentivized ambiguity task 2. In panel A, the three columns on the left report the number of situations subjects preferred drawing a ball from the risky urn over drawing a ball from the ambiguous urn. The three columns on the right report the switching points of the subjects. More precisely, it indicates the last urn before a subject switched from the ambiguous to the risky urn. A switching point of 0 means that the subject always preferred the risky urn; a switching point of 11 means that the subject always preferred the ambiguous urn. For a detailed description of the task, see appendix A. Panel B presents the summary statistics of the fraction of risky choices and the switching points.

Table 4: Descriptive statistics of the non-incentivized thought experiments

Panel A: Ellsberg (1961) urn		
	Observations	Fraction
Strong preference for ambiguous urn	12	9.9%
Slight preference for ambiguous urn	22	18.2%
Indifferent between both urns	10	8.3%
Slight preference for risky urn	40	33.1%
Strong preference for risky urn	37	30.6%
Total	121	100.0%
Panel B: Horse betting		
	Observations	Fraction
Preference for ambiguous horse bet	48	38.7%
Preference for risky horse bet	73	60.3%
Total	121	100.0%
	Mean	Standard deviation
Risk equivalent for ambiguous horse bet	30.7%	5.4%
Panel C: Complementary temperature bets		
	Mean	Standard deviation
Combined willingness to pay (London, in GBP)	34.49	31.31
Combined willingness to pay (Vancouver, in GBP)	27.52	27.76
Absolute ambiguity premium (in GBP)	6.97	15.89
Relative ambiguity premium	13.5%	53.4%
Panel D: Dynamic Ellsberg urn		
Risk equivalent	Observations	Fraction
2.5%	1	0.8%
12.5%	4	3.3%
22.5%	5	4.1%
27.5%	5	4.1%
32.5%	11	9.1%
37.5%	25	20.7%
42.5%	6	5.0%
47.5%	20	16.5%
52.5%	27	22.3%
57.5%	10	8.3%
62.5%	1	0.8%
67.5%	2	1.7%
72.5%	2	1.7%
77.5%	1	0.8%
95%	1	0.8%
Total	121	100.0%
	Mean	Standard deviation
Risk equivalent	43.9%	13.8%

The table summarizes the results of the four non-incentivized thought experiments. For a detailed description of the thought experiments, see appendix B.

Table 5: Aggregate consistency across elicitation methods

Elicitation method	n	Groups	Ambiguity seeking	Ambiguity neutral	Ambiguity averse	Aggregate ambiguity neutrality rejected	Definition of ambiguity neutrality
Ambiguity task 1	105	2	32 (30.5%)	–	73 (69.5%)	***	–
<i>Chakravarty and Roy (2009)</i>	81	3	4 (3.8%)	53 (50.5%)	48 (45.7%)	***	Switching points 4 and 5
		2	30%	–	70%		–
		3	8%	52%	40%		Switching points 4 and 5
Ambiguity task 2	116	2	31 (26.7%)	–	85 (73.3%)	***	–
<i>Lauriola and Levin (2001)</i>	62	3	6 (5.2%)	73 (62.9%)	37 (31.9%)	***	Switching points 5 and 6
		2	32.3%	–	67.7%		–
		3	11.3%	59.7%	29.0%		Crossover point $\in [0.41; 0.6]$
Thought experiment 1 (Ellsberg urn)	121	3	34 (28.1%)	10 (8.3%)	77 (63.6%)	***	Indifference between urns
<i>Butler et al. (2013)</i>	1,686	3	23.8%	24.7%	51.5%		Indifference between urns
Thought experiment 2 (Horse bet)	121	2	48 (39.7%)	–	73 (60.3%)	**	–
<i>Viscusi and Chesson (1999)</i>	269	3	27 (22.3%)	47 (38.8%)	47 (38.8%)	**	Risk equivalent = 30%
		2	41.6%	–	58.4%		–
		3	23.7%	52.6%	23.7%		Risk equivalent = 30%
Thought experiment 3 (Complementary temperature bets)	121	3	23 (19.0%)	27 (22.3%)	71 (58.7%)	***	Ambiguity premium = 0
<i>Fox and Thersky (1995)</i>			<i>(no data available)</i>				
Thought experiment 4 (Dynamic Ellsberg urn)	121	2	44 (36.4%)	–	77 (63.6%)	***	–
		3	17 (14.1%)	47 (38.8%)	57 (47.1%)	***	Risk equivalent $\in [0.475; 0.525]$

The table summarizes the fraction of subjects that can be classified into different broadly defined ambiguity preferences across the various elicitation methods. The first rows use a binary classification that distinguishes between ambiguity seeking and ambiguity averse decision makers. The second row presents the results based on a classification into three groups, including ambiguity neutrality. Since the Ellsberg thought experiment and the complementary temperature bets explicitly allow for ambiguity neutrality, these elicitation methods have been excluded from the first row. Based on the classifications, the last but one column reports the level of statistical significance at which the null hypothesis of the subjects being on average ambiguity neutral can be rejected using a Wilcoxon (1945) signed-rank test. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. The last column presents the exact definition of ambiguity neutrality applied to each of the elicitation methods. Below each set of results, the table presents the corresponding results of previous studies, where available.

Table 6: Individual consistency across elicitation methods

Panel A: Classification in ambiguity seeking/averse decision makers			
Ambiguity measures	Consistent preferences	Difference to random preferences	Test of association
Task 1 / Task 2	68.6%	18.6%***	**
Task 1 / horse bet	60.0%	10.0%*	
Task 1 / dynamic Ellsberg urn	64.8%	14.8%***	**
Task 2 / horse bet	58.6%	8.6%*	
Task 2 / dynamic Ellsberg urn	69.8%	19.8%***	***
Horse bet / dynamic Ellsberg urn	57.0%	7.0%	
Panel B: Classification in ambiguity seeking/neutral/averse decision makers			
Ambiguity measures	Consistent preferences	Difference to random preferences	Test of association
Task 1 / task 2	53.9%	20.6%***	
Task 1 / static Ellsberg urn	45.7%	12.4%***	**
Task 1 / horse bet	40.0%	6.7%	**
Task 1 / ambiguity premium	30.5%	−2.9%	
Task 1 / dynamic Ellsberg urn	54.3%	21.0%***	**
Task 2 / static Ellsberg urn	35.3%	2.0%	
Task 2 / horse bet	39.7%	6.3%	
Task 2 / ambiguity premium	35.3%	2.0%	
Task 2 / dynamic Ellsberg urn	52.6%	19.3%***	***
Static Ellsberg urn / horse bet	33.1%	−0.3%	*
Static Ellsberg urn / ambiguity premium	47.9%	14.6%***	
Static Ellsberg urn / dynamic Ellsberg urn	60.3%	27.0%***	***
Horse bet / ambiguity premium	33.9%	0.6%	
Horse bet / dynamic Ellsberg urn	35.5%	2.2%	
Ambiguity premium / dynamic Ellsberg urn	37.2%	3.9%	

The table summarizes the individual consistency of ambiguity preferences across the various elicitation methods. Panel A presents the average consistency based on a binary classification that distinguishes between ambiguity seeking and ambiguity averse decision makers; panel B presents the average consistency based on a classification into three groups, including ambiguity neutrality. Since the Ellsberg thought experiment and the complementary temperature bets explicitly allow for ambiguity neutrality, these tasks have been excluded from panel A. The first column presents for each pair of elicitation methods the fraction of individually consistent preferences, i.e., the fraction of subjects that exhibit the same ambiguity attitude for two given elicitation methods. The second column presents the difference of consistency compared to random preferences which equal 50% in the binary classification and 33.3% when using three categories. The last column shows the statistical significance of a Fisher (1922) test of association between each pair of ambiguity measures. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 7: Correlation of thought experiments

	Experiment 1 (switching point)	Experiment 2 (switching point)	Static Ellsberg urn	Horse betting (preference)	Horse betting (risk equivalent)	Absolute ambiguity premium	Relative ambiguity premium	Dynamic Ellsberg urn
Experiment 1 (switching point)	1.000							
	105							
Experiment 2 (switching point)	0.124 (0.214)	1.000						
	102	116						
Static Ellsberg urn	0.251*** (0.010)	0.180* (0.054)	1.000					
	105	116	121					
Horse betting (preference)	0.149 (0.131)	-0.006 (0.950)	0.025 (0.788)	1.000				
	105	116	121	121				
Horse betting (risk equivalent)	0.242** (0.013)	0.045 0.635	0.002 (0.985)	0.329*** (0.000)	1.000			
	105	116	121	121	121			
Absolute ambiguity premium	-0.062 (0.527)	-0.151 (0.105)	0.048 (0.600)	0.021 (0.822)	0.010 (0.916)	1.000		
	105	116	121	121	121	121		
Relative ambiguity premium	0.000 (0.998)	0.021 (0.822)	-0.024 (0.792)	-0.012 (0.893)	0.069 (0.450)	0.564*** (0.000)	1.000	
	105	116	121	121	121	121	121	
Dynamic Ellsberg urn	0.075 (0.447)	0.195** (0.036)	0.515*** (0.000)	0.233** (0.010)	0.079 (0.390)	0.023 (0.806)	-0.138 (0.131)	1.000
	105	116	121	121	121	121	121	121

The table presents the correlation of ambiguity preferences based on the various incentivized and non-incentivized thought experiments. p-values are given in parenthesis below the correlation coefficient. The number below indicates the number of observations. The correlation table is calculated such that the ambiguity preferences are comparable. A high number indicates ambiguity aversion for all measures. *, **, and *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Table 9: Predicting experimental ambiguity preferences with thought experiments

Explanatory variables	(1)	(2)	(3)	(4)	(5)
Ellsberg urn	0.026*** (3.45)	0.026*** (3.59)	0.026*** (3.56)	0.026*** (3.55)	0.026*** (2.96)
Horse betting (risk equivalent)		0.004** (2.33)	0.003* (1.77)	0.003* (1.74)	0.003* (1.72)
Horse betting (preference)			0.023 (1.04)	0.023 (1.04)	0.023 (0.99)
Relative ambiguity premium				0.004 (0.22)	0.004 (0.22)
Dynamic Ellsberg urn					0.000 (0.06)
R^2	0.1062	0.1526	0.1618	0.1622	0.1622
adj. R^2	0.0973	0.1355	0.1361	0.1276	0.1186

The table presents the regressions of the experimental ambiguity preferences (AMB_s) on the preferences obtained from the thought experiments. For a given number of predictors, only the specification with the highest adjusted R^2 is given. T-statistics are given in the parenthesis below. *, **, and *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Table 8: Time consistency of ambiguity measures

Ambiguity measure	Obs.	Correlation	p-value	R^2	adj. R^2
Experiment 1 (fraction)	26	0.2621	0.1958	0.069	0.030
Experiment 1 (switching point)	23	0.2633	0.2248	0.069	0.025
Experiment 2 (fraction)	26	0.3285	0.1014	0.108	0.071
Experiment 2 (switching point)	26	0.3285	0.1014	0.108	0.071
Ellsberg urn	26	0.0429	0.8352	0.002	-0.040
Horse betting (preference)	26	0.3587*	0.0719	0.129	0.092
Horse betting (risk equivalent)	26	-0.0585	0.7764	0.003	-0.038
Absolute ambiguity premium	26	-0.1007	0.6246	0.010	-0.031
Relative ambiguity premium	26	-0.0324	0.8750	0.001	-0.041

The table presents the time consistency of the ambiguity preferences. The first two columns report the correlation and p-value of the ambiguity measures, the last two columns the explained variance (R^2 and adjusted R^2) when regressing the second observation of the ambiguity preferences on the first observation. *, **, and *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Table 10: Correlation between survey items and experimental ambiguity preferences

Item	Description	Correlation	p-value	right sign
7	Before an examination, I feel much less anxious if I know how many questions there will be.	−0.2121**	0.0323	yes
11	Perfect balance is the essence of all good composition.	0.1380	0.1667	no
19	I would like to live for a while in a foreign country that is new to me.	0.1363	0.1718	yes
32	I enjoy unexpected events.	0.1436	0.1500	yes
33	The idea of taking a trip to a new country fascinates me.	0.1660*	0.0954	yes
38	I try to avoid situations that are ambiguous.	−0.1358	0.1735	yes
45	When a situation is uncertain, I never take action until I know all the risks involved.	−0.1523	0.1264	yes

The table presents all correlations between survey questionnaire items and experimental ambiguity preferences (AMB_s) for which the correlation is significant at the 20% level.

Table 11: Predicting experimental ambiguity preferences

Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)
Ellsberg urn	0.026*** (3.45)	0.026*** (3.59)	0.028*** (3.81)	0.028*** (3.83)	0.028*** (3.86)	0.026*** (3.54)
Horse betting (risk equivalent)		0.004** (2.33)	0.005*** (2.65)	0.005*** (2.67)	0.005*** (2.77)	0.005*** (2.65)
Survey item (45)			−0.012** (−2.25)	−0.010* (−1.90)	−0.010* (−1.65)	−0.008 (−1.36)
Survey item (33)				0.013 (1.53)	0.011 (1.28)	0.011 (1.29)
Survey item (38)					−0.008 (−1.16)	−0.008 (−1.18)
Survey item (7)						−0.006 (−1.07)
R^2	0.1062	0.1526	0.1941	0.2130	0.2240	0.2333
adj. R^2	0.0973	0.1355	0.1694	0.1806	0.1835	0.1849

The table presents the regressions of the experimental ambiguity preferences (AMB_s) on the preferences obtained from the thought experiments and the survey questionnaire. For a given number of predictors, only the specification with the highest adjusted R^2 is given. T-statistics are given in the parenthesis below. *, **, and *** denote statistical significance at the 10%, 5% and 1% level, respectively.