OFF THE CHARTS: MASSIVE UNEXPLAINED HETEROGENEITY IN A GLOBAL STUDY OF AMBIGUITY ATTITUDES

Olivier l'Haridon, Ferdinand M. Vieider, Diego Aycinena, Agustinus Bandur, Alexis Belianin, Lubomír Cingl, Amit Kothiyal, and Peter Martinsson*

Abstract—Ambiguity attitudes have been prominently used in economic models, but we still know little about their demographic correlates or their generalizability beyond the West. We analyze the ambiguity attitudes of almost 3,000 students across thirty countries. For gains, we find ambiguity aversion everywhere, while ambiguity aversion is much weaker for losses. Ambiguity attitudes change systematically with probabilities for both gains and losses. Much of the between-country variation can be explained through a few macroeconomic characteristics. In contrast, we find massive unexplained variation at the individual level. We also find much unexplained heterogeneity in individual responses to different decision tasks.

I. Introduction

AMBIGUITY aversion—the preference for known-probability outcome-generating processes over normatively equivalent processes in which exact probabilities are not known—has attracted considerable interest since the publication of Daniel Ellsberg's seminal article in 1961. In particular, ambiguity aversion has been indicated as a potential explanation of a wide range of economic outcomes in theoretical models. To name but a few, Mukerji (1998) used ambiguity aversion to explain the existence of incomplete contracts, Easly and O'Hara (2009) used it to explain the stock market participation bias, and Bryan (2010) hypothesized that ambiguity aversion may be causing the low uptake of insurance in developing countries in the presence of basis risk, while Alary, Gollier, and Treich (2013) studied the demand for insurance and self-protection under ambiguity aversion. Berger, Bleichrodt, and Eeckhoudt (2013) modeled the incidence of ambiguity aversion on treatment decisions where ambiguity could affect either the diagnosis or the treatment itself, with different consequences for medical decisions.

Given the importance that ambiguity attitudes have been ascribed in the determination of economic outcomes, it seems crucial to know whether the phenomenon is universal or whether there exist qualitative or quantitative differences across countries. In this paper, we examine cross-country differences in ambiguity preferences using a rich data set obtained in experiments with close to 3,000 students from thirty countries.[1] We further obtain rich measurements of preferences at the individual level, which allow us to investigate the variation of ambiguity attitudes over probability levels (ambiguity-insensitivity to probabilities, or a-insensitivity) and for gains versus losses (domain-dependence). This is important inasmuch as ambiguity attitudes have been found to differ systematically across these dimensions (see Trautmann & van de Kuilen, 2015, for a detailed survey), and since such systematic patterns serve to discriminate between different theoretical accounts of behavior under ambiguity (Baillon & Bleichrodt, 2015).

The conclusions we draw from the data are mixed. At the country level, ambiguity aversion is the standard pattern, albeit with significant quantitative differences between countries. We also find systematic patterns across prospect (lottery) characteristics, with a-insensitivity prevailing everywhere, and less ambiguity aversion for losses than for gains (although we fail to find much evidence of ambiguity seeking for either losses or small probabilities, as found in many previous studies; see Trautmann & van de Kuilen, 2015). We furthermore manage to explain cross-country differences in ambiguity attitudes recurring to a set of standard variables from the macroeconomic literature.

This picture changes once we move to lower levels of analysis. We find considerable heterogeneity among subjects, but our best attempts at explaining this variance bear little fruit. At 3% of the variance at the individual level explained by our best model, unexplained heterogeneity remains large even measured against the dismal standard of explaining individual variation in risk preferences (von Gaudecker, van Soest, & Wengström, 2011; l'Haridon & Vieider, in press). This variation could be due to either systematic noise in the responses of some individuals or ambiguity attitudes constituting an idiosyncratic trait that is orthogonal to observable characteristics, an issue to which we will return in section IV.

The situation does not improve when we try to explain differences in behavior depending on prospect characteristics. A model accounting for all possible prospect characteristics manages to explain a mere 10% to 19% of the residual

Received for publication January 27, 2017. Revision accepted for publication September 29, 2017. Editor: Shachar Kariv.

* l'Haridon: University of Rennes; Vieider: University of Reading; Aycinena: Universidad del Rosario; Bandur: Bina Nusantara University: Belianin: Higher School of Economics; Cingl: University of Economics; Kothiyal: Max Planck Institute for Human Development; Martinsson: University of Gothenburg and Linköping University.

F.V. acknowledges financial support from the Excellence Initiative at the University of Munich and the WZB Berlin Social Science Center. O.l'H. acknowledges financial support from the Institut Universitaire de France and Rennes Metropole District. We thank Roy Kouwenberg, Stephen Dimmock, Aurélien Baillon, and Han Bleichrodt for helpful comments. All errors remain our own.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/rest_a_00724.

[1] These data were first presented by Vieider, Lefebvre et al. (2015), who examined risk preferences and uncertainty preferences aggregated over individual choices, and their correlation with survey measures of preferences. They did not in any way discuss ambiguity preferences (here defined as the difference between risk and uncertainty preferences; see below). A second paper based on a subset of the same data, l'Haridon and Vieider (in press), presents a structural analysis of decision making under risk and does not use the data on uncertainty. We will refer to the latter to contrast our findings for ambiguity to those for risk in the same data set.

The Review of Economics and Statistics, October 2018, 100(4): 664–677

doi:10.1162/rest_a_00724

variance. We interpret the unexplained variance at this level as noise, given the common definition of the latter as capturing calculation mistakes, random responses, or response patterns not captured by the model fit to the data. The extent of the noise at this level becomes even more evident when comparing it to the performance of the same model for risk preferences, where the explained variance hovers around 75%.

We are not alone in finding poor predictive performance of experimentally measured ambiguity attitudes (for an exception, see Muthukrishnan, Wathiev, & Xu, 2009). Giné, Townsend, and Vickery (2008) tried to relate experimentally measured ambiguity aversion to insurance uptake in a large field experiment in India but found no correlation. Borghans et al. (2009) found cognitive and psychological characteristics to be predictive of risk attitudes, but not of ambiguity attitudes. Binmore, Stewart, and Voorhoeve (2012) could not find any stable demographic correlates of ambiguity attitudes (as well as finding low ambiguity aversion in general; see also Ahn et al., 2014). Sutter et al. (2013) found only weak predictive power of experimentally measured ambiguity attitudes for the behavior of adolescents. Stahl (2014) also detected high levels of noise in an experimental study of ambiguity attitudes. Even Dimmock, Kouwenberg, and Wakker (2015), while finding some correlations in a representative sample of the Dutch population, manage to explain at most around 5% in the variation of their ambiguity indices—a proportion that roughly corresponds to the one we report in our own data.

We cannot fully exclude that some of these findings may be driven by our measurement method. We measure ambiguity attitudes using two certainty equivalents (CEs)—one for risk (known probabilities) and one for uncertainty (unknown probabilities), and construct an ambiguity premium from the difference of the two. CEs have been shown to provide an unbiased measure of ambiguity attitudes (Trautmann, Vieider, & Wakker, 2011), and they are popular in the literature (Abdellaoui et al., 2011; Sutter et al., 2013). They are easy to implement, including in paper-and-pencil experiments, straightforward to incentivize, and have the great advantage of being simple to administer and understand. Furthermore, they allow for a model-free analysis of ambiguity attitudes. Given that the method consists of two separate measurements for risk and uncertainty attitudes, there could, however, be an accumulation of error terms when deriving our ambiguity premia. Some recent papers have used direct measurements of ambiguity aversion by matching probabilities and have reported some success in correlating the measures to real-world behavior (Dimmock, Kouwenberg, & Wakker, 2015; Dimmock, Kouwenberg, Mitchell et al., 2015). We will provide a more detailed examination of measurement errors across different tasks in the discussion.

We do not mean to conclude from our results that ambiguity aversion is not an empirically meaningful concept. The use of abstract urns and the comparison of extreme situations of completely known and completely unknown probabilities, however, seem to induce high levels of inconsistencies in responses, which may well result from the salient and artificially induced absence of information from one of the urns (Frisch & Baron, 1988; Fox & Tversky, 1995). These issues may then be further exacerbated by measurement problems, which are also well known in the risk preference literature. One solution for applied empirical researchers may then be to recur to more natural sources of uncertainty (Abdellaoui et al., 2011; Baillon et al., 2016). In addition, it appears crucial to obtain a better understanding of measurement issues affecting our estimates of ambiguity attitudes and comparative evidence on the relative merit of different measurement tools. We will return to both of these issues in section IV.

II. Experiment and Analysis

A. *The Experiment*

We conducted the same experiment in thirty countries. The countries were selected with an eye to diversification along several dimensions that we deemed potentially important for our study. These included inter alia geographic representation and extension, level of income per capita, and importance in economic and population terms. A total of 2,939 subjects participated in the experimental sessions. Students were used since they were deemed more comparable across countries than other population groups. Subjects were recruited at major public universities in the different countries, with a few exceptions where no collaborators at public universities could be found (Brazil, Guatemala, Malaysia, Saudi Arabia, and Tunisia). Care was taken to obtain a subject sample that was balanced in terms of sex and study major, although this was not always successful (e.g., in Saudi Arabia, only men could be recruited because our male contact was not allowed to interact with female students). In universities with a standing subject pool, we recruited only subjects who had participated in at most two experiments before, so that they would be similar to subjects in developing countries for whom experiments were new. The online appendix provides a summary of the main subject characteristics country by country.

All experiments were run between September 2011 and October 2012. Experiments across countries were kept as comparable as possible. The experiment was run in the teaching language of the university, since many countries included in the study are multilingual, so that the official teaching language is the only one all students shared. Instructions were translated from English and back-translated into English by a different person (Brislin, 1970). Differences were then eliminated by discussion. The payoffs were carefully converted using World Bank purchasing power parity (PPP) data and then double-checked using PPP conversion rates calculated from net wages of student assistants at the university where the experiments took place. Vieider (2012) tested explicitly whether small variations in payoffs in the order of ±20% would make a difference in terms of measured risk or uncertainty attitudes and found none. Also, the experiment was run in two cities in China, Beijing and Shanghai, and on two campuses in Addis Ababa, Ethiopia, to determine whether differences found could be ascribed to differences in the subject pool, which would be troubling for an international comparison. No such differences were found once observable subject characteristics had been controlled for (see Vieider, Chmura, et al. (2015). Subjects were reassured of their anonymity, an element that is potentially important for measures of ambiguity, as the potential for scrutiny may increase ambiguity aversion (Baltussen, van den Assem, & van Dolder, 2016).

We elicited CEs for 44 binary prospects that differed by outcomes, probabilities, decision domain (gains versus losses), and source of uncertainty (known probabilities versus vague probabilities). In this paper, we concentrate specifically on the sixteen pairs of prospects that were administered using both objective and vague probabilities. In particular, we had six prospects providing a prize of 20 euros PPP with probability $\frac{i}{8}$, $i = \{1, 2, 3, 5, 6, 7\}$ or else nothing. For the extreme probabilities of $\frac{1}{8}$ and $\frac{7}{8}$, we further had prospects offering a prize of 20 euros or else 5 euros PPP. For our main analysis, we focus on the prospects with lower outcomes of 0, since this simplifies our analysis considerably at little or no cost in terms of insights. We use the four additional prospect pairs for consistency checks and out-of-sample predictions.

Subjects were asked to make a choice between the prospect and different sure amounts (those paid with certainty) of money contained between y and x. For gains, the sure amounts increased from y to x. For losses, the sure amount decreased from $-y$ to $-x$. For gains, subjects generally choose the prospect for small sure amounts and switch to preferring the sure amount as the latter gets larger (and vice versa for losses). The CE of a subject was encoded as the average of the last sure amount for which the prospect was chosen and the first sure amount chosen (vice versa for losses).

The tasks were distributed across different categories and domains (see the online appendix for the full experimental instructions). By "decision domain," we mean that choices were framed as either gains or losses. Losses were always administered in a second part of the experiment and took place from an endowment. This endowment was given conditional on the second part being selected for real play and was equivalent to the highest loss of 20 euros no matter what the selected choice. Etchart-Vincent and l'Haridon (2011) tested whether decisions from an endowment are different from decisions involving real losses and found no differences. In each of the two domains, we had tasks with known probabilities, which we call *risky*, and decisions involving unknown probabilities, which we call *uncertain*. The tasks were always kept in the same order, starting with risky gains and then uncertain gains, and in a second part, risky losses were followed by uncertain losses. This was done so as to facilitate the logistics and avoid mix-ups, as well as to reduce the cognitive burden on subjects. A large-scale pilot with 330 subjects showed that such a fixed ordering was less demanding for subjects while not significantly affecting the measures used in this paper.

In the experiment, the urns were not called risky or uncertain, but rather "transparent" and "opaque." Concerning the risky urn, subjects simply learned that the urn contained exactly eight balls, numbered from 1 to 8. About the uncertain urn, they were told, "You cannot see what numbers the balls contained in the urn have. This means that you do not know the exact numbers that are present in that urn. All balls bear a number between 1 and 8 inclusive (they have either 1, 2, 3, 4, 5, 6, 7, or 8 written on them), but it is possible that some numbers are absent from this urn while others occur repeatedly. Thus, you do not know the exact composition of the urn." This implementation of uncertainty permits centering the uncertainty around a known probability distribution. In this sense, a prospect offering a given prize when a ball with the number 1 or 2 is extracted offers a vague probability interval that is centered on a probability of two-eights. The vagueness derives from the fact that in reality, the probability may be lower or higher than two-eights.

Due to logistic reasons, we could not allow subjects to choose the colors on which to bet. This may result in distrust of the experimenter to contaminate the results (Charness, Karni, & Levin, 2013), although others studies failed to find such effects (Oechssler & Roomets, 2015), and the evidence accumulated to date on the relevance of this issue is inconclusive (Trautmann & van de Kuilen, 2015). We took two steps to counteract this issue. First, we explicitly encouraged subjects to check the contents of the ambiguous urn after the experiment. This was meant to reassure them that while some balls could be absent from the urn as described in the instructions, we did not systematically manipulate the urn to their disadvantage. Second, we conducted a pilot on whether letting subjects choose the winning number of the balls would make a difference in our particular design. We did not find any evidence that this was the case (the online appendix provides details about this pilot). We will see that the data do not appear to support explanations based on suspicion against the experimenter (see also Dimmock, Kouwenberg, Mitchell, et al., 2015, for further evidence on this).[2]

B. Analysis and Econometrics

For our analysis, we construct ambiguity premia in the following way. We have eight CEs for uncertain gains and eight CEs for uncertain losses. We contrast them with the equivalent risky CEs for which probabilities are known. The ambiguity premium is then defined as

[2] One may also argue that in some sense, dislike of asymmetric knowledge by the experimenter is a part of ambiguity aversion. For instance, Chow and Sarin (2002) showed that subjects disliked the information asymmetry inherent in the ambiguity problem even when distrust of the experimenter was completely excluded.

$$\pi_{si} = \frac{ce^r_{si} - ce^u_{si}}{x_{si} - y_{si}}, \tag{1}$$

where the subscript s indicates the sign (gains or losses), i indicates the probability level or prospect, and the superscripts r and u indicate risk and uncertainty, respectively.[3] The division by the outcome range of the prospect serves to normalize the value of the certainty equivalent. A CE under uncertainty that is smaller than the corresponding CE under risk thus indicates ambiguity aversion for both gains and losses. The normalization allows us to interpret the difference directly on a probability scale. For instance, an ambiguity premium of 0.1 signifies that a decision maker is willing to give up an additional 10 percentage points of the probability of winning the extra prize $x - y$ in order to avoid ambiguity, measured against her risk aversion.

Our index of ambiguity attitudes is completely model free. To show this, it is instructive to examine the relation of our ambiguity premium to other measures used in the literature, and particularly the ones derived from matching probabilities by Dimmock, Kouwenberg, Mitchell, et al. (2015). Following the latter in applying the source method of Abdellaoui et al. (2011) to our stimuli (see Dimmock, Kouwenberg, Mitchell et al., 2015, section 3.2 for a detailed discussion of the theoretical assumptions), and dropping the subscripts si to the prospect characteristics to avoid notational clutter, we can define our certainty equivalent under risk as

$$ce^r = u^{-1}\left[w^r_s(p)u(x) + (1 - w^r_s(p))u(y)\right], \tag{2}$$

where u is a utility function transforming outcomes into utiles and w is a source or probability weighting function transforming probabilities into decision weights, and the power -1 indicates the inverse of the utility function. The superscript r to w serves to remind us that this function is applied to objectively given probabilities or risk, while the subscript s emphasizes its sign dependence. Now take an uncertain event E to represent likelihoods in our uncertain urn. Assigning a probability $P(E)$ to the uncertain event, we can now derive the CE under uncertainty as follows:

$$ce^u = U^{-1}\left[w^u_s(P(E))U(x) + (1 - w^u_s(P(E)))U(y)\right], \tag{3}$$

where U denotes utility under uncertainty and the superscript u to w indicates that this is the source function for uncertainty. Assuming $y = 0$ for simplicity (the reasoning is similar for $y \neq 0$) and normalizing utility such that $u(0) = U(0) = 0$ and $u(x) = U(x) = 1$, we can write the ambiguity premium as

$$\pi = \frac{u^{-1}\left[w^r_s(p)\right] - U^{-1}\left[w^u_s(P(E))\right]}{x - y}. \tag{4}$$

[3] In particular, we can obtain a normalized CE for risk, $\pi^r_{si} = \frac{ce^r_{si} - y_{si}}{x_{si} - y_{si}}$, and a normalized CE for uncertainty, $\pi^u_{si} = \frac{ce^u_{si} - y_{si}}{x_{si} - y_{si}}$. Equation (1) then results from $\pi^r_{si} - \pi^u_{si}$.

Equation (4) shows that our ambiguity premia capture not only differences in decision weights but also any potential differences in utility between risk and uncertainty. By assuming equality in utility between risk and uncertainty, matching probabilities capture differences only in decision weights. Thus, if utility is different between risk and uncertainty but decision weights are not, the matching probabilities will not detect ambiguity attitudes, but our ambiguity premia will.

We analyze the data using a hierarchical or multilevel model with three levels of analysis (see Snijders & Bosker, 2012, for an introduction to multilevel modeling). Our fundamental units of analysis are ambiguity premia for single-prospect comparisons, of the type indicated in equation (1). The next higher level is constituted by subjects, so that the standard errors are clustered at that level and we can capture the between-subject variance. The highest level is constituted by countries. We express this as

$$\pi_{inc} = \beta_0 + X_{inc}\beta_{nc} + \nu_c + \mu_{nc} + \varepsilon_{inc}, \tag{5}$$

where i is now a running index including all prospect characteristics, having dropped s to avoid notation overload. The part $X_{inc}\beta_{nc}$ constitutes the so-called fixed part of the model, with X containing observable characteristics at the prospect (i), individual (n), and country (c) level, and β constituting a vector of regression coefficients. The remaining part of the equation is referred to as the random part, with ν_c capturing the random intercept at the country level and μ_{nc} the random intercept at the subject level. These two terms thus capture systematic variation in the residuals at these two levels. The remaining term ε_{inc} captures the residual variance of the model. By using an empty set of observables X and dropping the fixed part of the model, we can estimate a baseline model that purely quantifies the variance at the different levels of analysis. The extent to which the variance at various levels is reduced by adding independent variables at the different levels to X then serves to assess our success in explaining the overall variation in the data. We will run all regressions separately for gains and losses. While it is in principle possible to run one large regression on all ambiguity premia, this requires interaction terms to be included at all times, given the opposite signs for gains and losses. Running two separate regressions for gains and losses thus avoids having to run triple interactions when considering crossed effects.

The random intercepts captured in ν and μ allow the intercept to be different by country and by subject within countries and thus capture general differences in ambiguity aversion, while the regressands contained in X are forced to show the same effect for the whole sample. This has the advantage of quantifying the variance at each level, while adjusting the standard errors for the level of analysis. While the above model constitutes a first step in the direction of accounting for heterogeneity on various levels, the assumption of equal effects of observable characteristics across different levels is still unsatisfactory. In particular, we may want to allow for the possibility that individual-level regression results may differ between countries—for instance, to what extent the gender effect in ambiguity attitudes is different across countries. We thus amend our model by allowing for so-called random slopes:

$$\pi_{inc} = \beta_0 + X_{inc}\beta_{nc} + Y_{in}\nu_c + Z_i\mu_{nc} + \varepsilon_{inc}. \qquad (6)$$

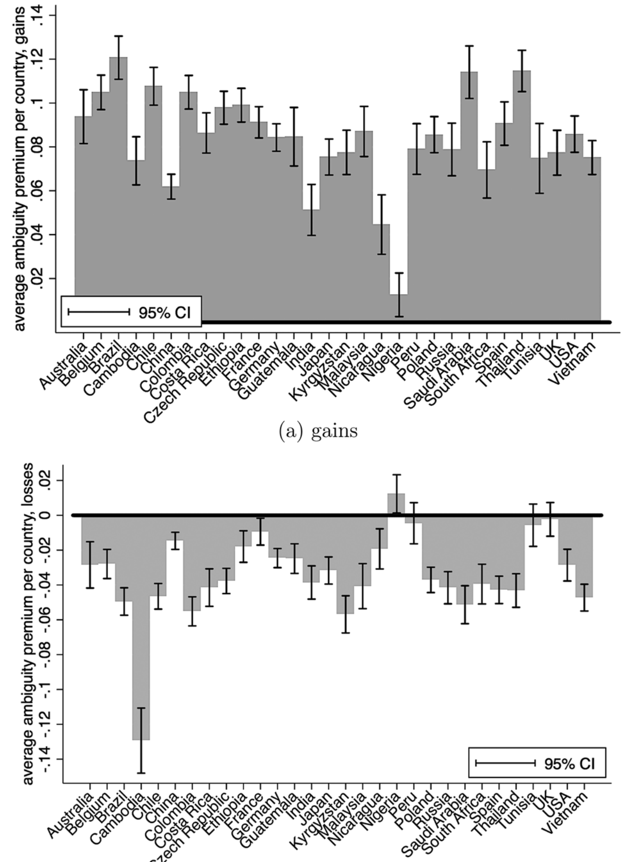
Thus, we now also allow the effects of prospect and individual characteristics to differ across subjects and countries. For instance, Y_{in} may contain both prospect characteristics, such as the probability of winning in a given prospect, and individual characteristics, such as, for instance, a female dummy. The prospect characteristic would capture differences in the effect of prospect characteristics across countries, in our example differential effects of the probability of winning across countries. For the latter, we may furthermore observe differences across subjects (e.g., different subjects may react differently to the probability of winning). This is captured by inserting prospect characteristics as random slopes at the individual level into Z_i.

The results then correspond to those of an OLS regression, with standard errors adjusted for the specific level of analysis. The probability inserted as an explanatory characteristic into matrix Y will allow us to quantify a-insensitivity, the phenomenon by which ambiguity attitudes change less than proportionally with the probability of winning. Technically, a-insensitivity in a given country is indeed simply measured by the regression coefficient of probability at this level, with a positive value indicating a-insensitivity, a zero coefficient indicating perfect sensitivity, and a negative coefficient indicating excess sensitivity. The random intercept measures the level of ambiguity aversion. Inserting the probability into Z will achieve the same at the individual level; it will capture heterogeneity in reactions to probabilities across subjects. Finally, inserting the probability into the fixed part X and comparing this to a baseline model empty of covariates will allow us to determine to what extent we can reduce the residual variance ε by controlling for this prospect characteristic.

III. Results

We present the results in several steps. We start with a nonparametric analysis of the ambiguity premia. We then proceed to documenting differences between countries using random intercepts—which serve to quantify country-level differences in aggregate ambiguity aversion—and random slopes for probability—which quantify between-country differences in a-insensitivity. We then quantify the variance in the data and explore the extent to which we can explain the overall level of variance observed at each of the three levels: between countries, between individuals, and between ambiguity premia.

Figure 1.—Average Ambiguity Premia by Country for Gains and Losses with 95% Confidence Intervals



(a) gains

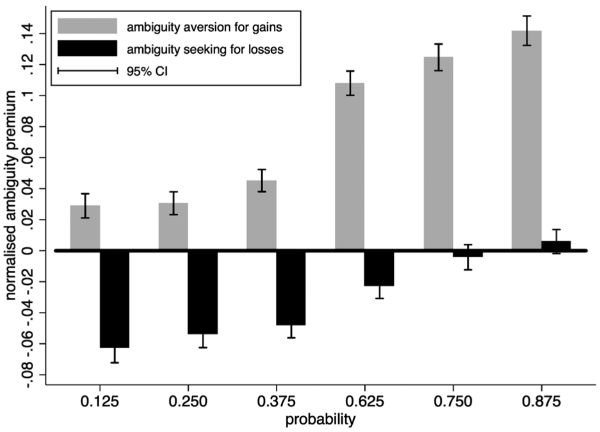
(b) losses

A. Descriptive Results

We start by presenting some descriptives of the nonparametric data. Figure 1 shows the average ambiguity premia per country, together with their 95% confidence intervals. The premia were obtained in the following way. We first calculated the average ambiguity premium per subject, separately for gains and losses. We then took the average of this mean premium per country, again separately by decision domain.

Figure 1a shows the premia thus obtained for gains. We find significant ambiguity aversion on average in all countries. At the same time, there is significant heterogeneity in average ambiguity attitudes across countries. The premia can be interpreted on a probability scale, meaning that, for instance, subjects in Belgium are on average willing to sacrifice 10 percentage points of the probability of winning to play a risky rather than an ambiguous prospect. Figure 1b shows the average ambiguity premia per country for losses, where a positive premium now indicates ambiguity seeking. Most country premia are negative, showing that the prevalent pattern is still one of ambiguity aversion. This ambiguity aversion is considerably weaker than for gains (with the exception of Cambodia, where ambiguity aversion

FIGURE 2.—AMBIGUITY AVERSION BY PROBABILITY



for losses is extremely strong). For a few countries, such as Peru, Tunisia, and the United Kindom, we cannot exclude ambiguity neutrality, while Nigeria again stands out from the general trend, this time by showing ambiguity seeking. Lower levels of ambiguity aversion for losses than for gains are consistent with previous evidence. Cohen, Jaffray, and Said (1987) found ambiguity aversion for gains but ambiguity neutrality for losses. Similar results were obtained, among others, by Friedl, Lima de Miranda, and Schmidt (2014). However, we do not find much evidence for ambiguity seeking for losses. This contrasts with the previous literature, where most of the studies report either ambiguity neutrality or ambiguity seeking for losses (see Trautmann & van de Kuilen, 2015).

We have so far looked only at average attitudes per subject. This masks the fact that ambiguity attitudes have been found to differ systematically across the probability spectrum (Tversky & Fox, 1995; Abdellaoui et al., 2011). Figure 2 depicts ambiguity attitudes by probability level, aggregating over the whole sample across subjects and countries. For gains, we find relatively low ambiguity aversion for small probabilities, with ambiguity aversion increasing as probabilities get larger. A jump occurs between $p = 3/8$ and $p = 5/8$, which suggests that people are particularly insensitive to changes in probability to either side of $p = 0.5$. For losses, we find the exact opposite pattern: ambiguity aversion starts out large for small probabilities and becomes smaller as probabilities increase, resulting in a trend toward ambiguity seeking for the largest probability (although this fails to reach significance in the aggregate data). There also is a similar jump around the middle of the probability spectrum as observed for gains.

Most previous studies have found ambiguity seeking for small probability gains (Dimmock, Kouwenberg, & Wakker, 2015), but some have found attitudes closer to ambiguity neutrality for the same type of prospects (Abdellaoui et al., 2011). The reason for this may well be that previous studies that found ambiguity seeking used comparative environments in which uncertainty was directly compared to risk. In our design, there is no direct comparison, and we elicited certainty equivalents for risk and uncertainty separately in separate blocks of the experiment. It is well known that ambiguity attitudes are strongest in settings that allow the direct comparison of objectively given and vague probabilities (Fox & Tversky, 1995; Chow & Sarin, 2001). In this sense, our measures can be seen as a lower bound on ambiguity attitudes. Bouchouicha, Martinsson, Medhin, and Vieider (2017) furthermore report evidence that a-insensitivity increases in monetary stakes, so that ambiguity seeking for small probabilities may emerge only at higher stake levels. While they found no evidence for ambiguity seeking at small probabilities for the stake levels used here, such ambiguity seeking started to emerge once stakes were doubled.

B. *Ambiguity Aversion and A-Insensitivity by Country*

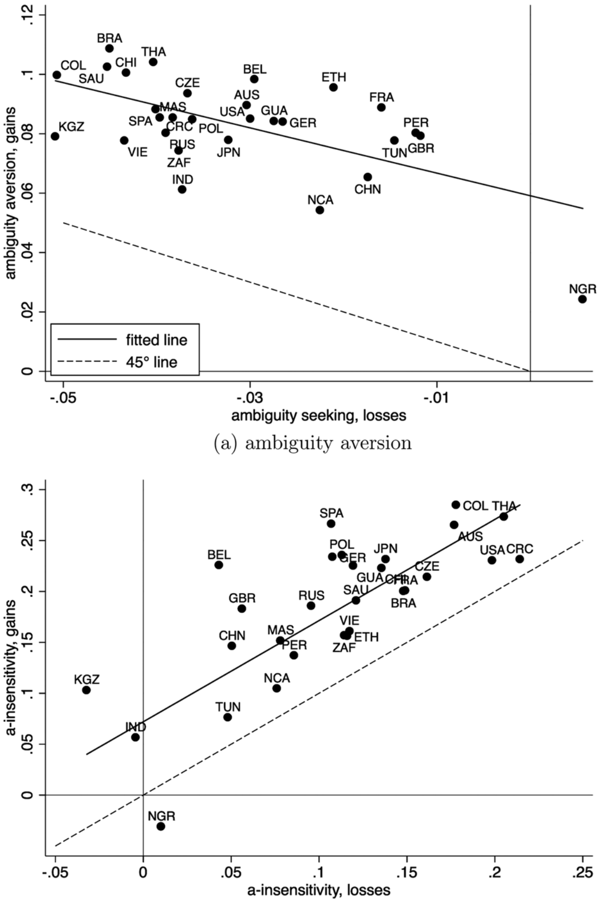
We start our parametric analysis by quantifying the aggregate ambiguity aversion and a-insensitivity at the country level for gains and losses. We quantify the two measures by running multilevel regressions separately for gains and losses and using probability as the sole explanatory variable, whereby we normalize the probability to be centered around 0.5 (i.e., we take $p - 0.5$). We further obtain random slopes for probabilities at the country level, allowing us to investigate differences in a-insensitivity across countries.

Figure 3 shows scatter plots of the two measures for gains and losses, with figure 3a showing a scatter plot of ambiguity attitudes for gains and losses, and figure 3b showing a scatter plot of a-insensitivity for gains and losses. Ambiguity attitudes for gains and losses show a clear negative correlation at the country level ($\rho = -0.44, p = 0.017$). Given the inverse interpretation for gains and losses, this means that ambiguity aversion for gains is negatively correlated with ambiguity seeking for losses. A-insensitivity, indicated in figure 3b, shows a clear, positive correlation across domains ($\rho = 0.74, p < 0.001$). Overall, the patterns we observe at the country level thus indicate a remarkable stability of ambiguity attitudes between gains and losses.

C. *Explaining Residual Variance (within Subjects)*

We now examine how much of the variation in ambiguity premia we can explain at the prospect or premium level—in other words, to what extent we can explain the residual variance in our model. Table 1 shows regressions of the ambiguity premia on a number of prospect characteristics. The first column presents a benchmark model empty of covariates that serves purely to quantify the variance at the various levels. At the residual level, we observe a value of 0.043 for gains, corresponding to about 84% of the overall variance. The picture is similar for losses, where again about

FIGURE 3.—SCATTER PLOTS OF AMBIGUITY AVERSION AND A-INSENSITIVITY FOR GAINS AND LOSSES



(a) ambiguity aversion

(b) a-insensitivity

Cambodia is excluded from both graphs since it constitutes an outlier.

84% of the variance is to be found at the residual level.[4] The goal of the subsequent regressions will now be to see how much of this variance we can explain through characteristics of the choice problem.

Model 1 adds the most basic characteristic of the decision problem: the probability of winning the prize. This probability shows a positive effect for both gains and losses, indicating a-insensitivity, whereby ambiguity aversion (seeking) increases in probabilities for gains (losses). The effect is weaker for losses than it is for gains, further confirming the effects already discussed at the country level. The proportion of variance explained by this variable is, however, rather modest, amounting to 4.7% of the overall residual variance for gains and only 2.3% for losses. Model 2 further adds a dummy indicating whether the probability is larger than 0.5. This is motivated by the jump observed in figure 2 and performs better than, for instance, inserting the square of the probability. Inserting this dummy constitutes a significant improvement in terms of goodness of fit relative to model 1 for gains ($\chi^2(1) = 24.42, p < 0.001$; likelihood ratio test) but not for losses ($\chi^2(1) = 0.38, p = 0.578$). The improvement in terms of variance explained remains quite modest, however, with a mere 0.2% of the variance explained relative to model I for gains (and no improvement for losses).

Model 3 further adds random slopes to our model. That is, we let the effect of probabilities and of the $p > 0.5$ dummy be different from subject to subject (i.e., we insert the random slopes at the next higher level while keeping the random intercept model at the country level). For gains, we find significant variance to be captured by both the probability and the probability cutoff dummy, resulting in a significant improvement of fit relative to model 2 ($\chi^2(2) = 782.95, p < 0.001$). We can explain 14% of the variance remaining unexplained in model 2, which brings us to 19% of the overall residual variance being explained by prospect characteristics. For losses, the picture is similar, albeit somewhat more muted. While there is a clear improvement in model fit ($\chi^2(2) = 207.93, p < 0.001$), no significant variance is explained by the random slope for probability. There is, however, significant variance being explained by the individual-level random slope of the cutoff dummy. The variance explained improves by 8.9% relative to model 2, which brings us to 10.5% relative to the overall residual variance.

The regressions just discussed show that choices are systematically related to prospect characteristics. The relatively large improvement between models 2 and 3, where we allowed reactions to probabilities to differ between subjects, shows further that there is considerable individual heterogeneity in reactions to prospect characteristics. Nonetheless, our success in explaining overall variance is modest, with 81% of the residual variance at the premium level remaining unexplained in our best-performing model 3 for gains and fully 89.5% for losses. We can put this figure into context by comparing it to the same figure for risk, obtained by running the exact same models directly on the normalized certainty equivalent for risk, ce^r_{is}. The equivalent of model 1 for risk, using probability alone to explain choices and assuming uniform effects across subjects, can explain 58% of the residual variance for gains. For losses, the figure is 61%. Further adding probability squared (which for risk works better than the cutoff dummy) and allowing for random effects to emulate model 3, we manage to explain 73% of the total residual variance for gains and 75% for losses (i.e., only 25% to 27% of the residual variance remains unexplained).

The large level of residual variance that remains unexplained by prospect characteristics for ambiguity is particularly surprising in the light of the observation that no other prospect characteristics exist that one could conceivably control for (given that other nonlinear elements we tried did not show any effect). We thus conclude from this that the largest part of the variance—around 80% for gains and 90% for

[4] The proportion of variance occurring at the residual level is influenced by our decision to analyze gains and losses separately. Given the large heterogeneity between domains, the variance at the residual level increases to 95% when we enter both domains jointly into the regressions. The subsequent analysis, however, does not change in any substantive way by treating the two decision domains separately.

TABLE 1.—RESIDUAL LEVEL REGRESSIONS

	Gains (Ambiguity Aversion)				Losses (Ambiguity Seeking)			
	Benchmark	Model 1	Model 2	Model 3	Benchmark	Model 1	Model 2	Model 3
Probability		0.168***	0.100***	0.100***		0.095***	0.087***	0.087***
		(0.020)	(0.020)	(0.020)		(0.017)	(0.017)	(0.017)
$p > 0.5$ dummy			0.040***	0.040***			0.005	0.005
			(0.008)	(0.008)			(0.008)	(0.008)
Random slopes				✓				✓
Constant	0.083***	−0.001	0.013*	0.011*	−0.034***	−0.081***	−0.079***	−0.079***
	(0.004)	(0.007)	(0.007)	(0.006)	(0.005)	(0.009)	(0.009)	(0.009)
Country VAR	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Subject VAR	0.008***	0.008***	0.008***	0.009***	0.008***	0.008***	0.008***	0.000
	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
Residual VAR	0.043***	0.041***	0.041***	0.035***	0.044***	0.043***	0.043***	0.039***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	17,629	17,629	17,629	17,629	17,605	17,605	17,605	17,605
Countries	30	30	30	30	30	30	30	30
Log likelihood	1,535.487	1,968.557	1,980.767	2,372.240	1,426.181	1,557.752	1,557.908	1,661.871

Standard errors in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

losses—would appear to consist of pure noise (defined as inconsistency of behavior with our modeling assumptions). We next explore to what extent we can uncover more stable patterns once we take the analysis to higher levels of analysis and aggregate across the single choices.

D. *Explaining Individual Variance*

We next look at the variance at the individual level. Table 2 shows a number of regressions on individual characteristics. All regressions also control for the prime characteristics at the premium level: the probability of winning and the probability cutoff dummy at 0.5. The proportion of the overall variance occurring at the individual level is highly significant and constitutes 17.6% for the overall variance for gains and 15.8% for losses.[5]

Model 1 regresses the ambiguity premium on biological characteristics, including a female dummy and the age of the respondent (entered as a z-score), and on study characteristics, including the (z-score of) grade point average (GPA) and a series of dummies capturing the study major. We find women to be more ambiguity averse than men for gains. This corresponds to the findings of Powell and Ansic (1997), while Borghans et al. (2009) and Dimmock, Kouwenberg, Mitchell, et al. (2015) found an effect going in the opposite direction, and Dimmock, Kouwenberg, and Wakker (2015) found no gender effect. There is no gender effect for losses. We also register some effects of study characteristics, but none of them are consistent between gains and losses. For gains, we find arts students to be less ambiguity averse and medicine students more ambiguity averse (marginally significant). For losses, we find a correlation indicating that students with a higher GPA tend to be more ambiguity averse—contrary to what one might expect if ambiguity aversion is taken to be a decision bias. The regression performs poorly in terms of the variance explained. For gains, model 1 explains 1% of the variance in the benchmark model, an improvement that is only marginally significant ($\chi^2(10) = 17.33, p = 0.067$). For losses, the improvement is significant ($\chi^2(10) = 21.11, p = 0.020$), but at 1.3%, the proportion of variance explained remains underwhelming.

Model 2 adds responses to a questionnaire eliciting cultural attitudes developed by Hofstede (1980). The cultural attitudes are then grouped into four dimension or attitudes—individualism, uncertainty avoidance, masculinity, and power distance—among which especially uncertainty avoidance has been indicated to capture something akin to ambiguity aversion (see also Rieger, Wang, & Hens, 2014, for correlations with risk aversion). For gains, we observe a marginally significant positive correlation between individualism and ambiguity aversion, while we observe no significant correlations with losses. The model does not represent a significant improvement over model 1 for gains ($\chi^2(4) = 4.58, p = 0.333$) or for losses ($\chi^2(4) = 1.00, p = 0.910$). Using aggregate country-level measures of the Hofstede measures jointly with individual deviations instead of the individual measures used here does not improve these results.

Model 3 adds interaction effects to model 2. That is, all individual-level variables included in model 2 are crossed with the probability of winning. This allows us not only to investigate average ambiguity attitudes at the individual level as done previously, but also whether a-insensitivity varies with subject characteristics. The probability is always recentered to make $p = 0.5$ coincide with 0 so as to facilitate the interpretation of the pure effects. This model constitutes a significant improvement over model 2 for both

[5] These figures obtain after controlling for the prospect caracteristics of the probability of winning and the probability cutoff at $p = 0.5$. The proportion of variance occurring at the individual level without the inclusion of these controls is somewhat lower, coming to 16% for gains and 15.4% for losses. We take the model including the prospect characteristics as our benchmark model, since we later need those characteristics when inserting crossed effects, and this could result in explained variance decreasing (i.e., the proportion of remaining variance increasing) relative to the benchmark model if it were not to include those characteristics because of the decrease in residual variance.

TABLE 2.—INDIVIDUAL-LEVEL REGRESSIONS

	Gains (Ambiguity Aversion)				Losses (Ambiguity Seeking)			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Female	0.010**	0.008*	0.020**	0.021***	−0.005	−0.004	0.015	0.013
	(0.004)	(0.005)	(0.008)	(0.008)	(0.005)	(0.004)	(0.011)	(0.010)
Age	0.003	0.003	0.005	0.006	0.005	0.005	0.007	0.008
	(0.004)	(0.004)	(0.008)	(0.007)	(0.004)	(0.004)	(0.006)	(0.005)
Gpa	0.002	0.002	0.001	0.001	−0.008***	−0.008***	−0.009**	−0.009**
	(0.002)	(0.002)	(0.004)	(0.004)	(0.002)	(0.002)	(0.004)	(0.004)
Mathematics	0.004	0.004	−0.014	−0.012	0.001	0.001	−0.016	−0.018
	(0.010)	(0.011)	(0.013)	(0.014)	(0.008)	(0.008)	(0.015)	(0.014)
Natural sciences	−0.010	−0.010	0.014	0.013	0.015	0.015	0.053*	0.048*
	(0.008)	(0.008)	(0.016)	(0.016)	(0.011)	(0.011)	(0.027)	(0.025)
Medicine	0.019*	0.019*	−0.041**	−0.041**	0.002	0.003	−0.016	−0.017
	(0.012)	(0.011)	(0.020)	(0.020)	(0.013)	(0.013)	(0.042)	(0.042)
Social sciences	−0.010	−0.011	0.008	0.007	0.010	0.011	0.027	0.025
	(0.009)	(0.010)	(0.020)	(0.020)	(0.010)	(0.010)	(0.020)	(0.020)
Humanities	−0.006	−0.007	0.013	0.013	−0.011	−0.011	0.007	0.003
	(0.012)	(0.012)	(0.027)	(0.027)	(0.018)	(0.019)	(0.022)	(0.021)
Arts	−0.028**	−0.027**	0.033*	0.041*	0.002	0.002	0.056**	0.044**
	(0.013)	(0.013)	(0.019)	(0.021)	(0.019)	(0.019)	(0.022)	(0.022)
Study other	−0.014*	−0.014*	−0.016*	−0.016*	0.001	0.002	0.003	0.001
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)
Power distance		−0.001	−0.001	−0.001		0.001	0.003*	0.003*
		(0.001)	(0.002)	(0.002)		(0.001)	(0.001)	(0.001)
Individualism		0.001*	0.000	0.000		0.000	0.001	0.001
		(0.001)	(0.001)	(0.001)		(0.001)	(0.002)	(0.002)
Uncertainty avoid.		−0.000	0.002*	0.002*		0.000	0.002	0.003
		(0.001)	(0.001)	(0.001)		(0.001)	(0.002)	(0.002)
Masculinity		−0.000	0.005***	0.005***		0.000	0.001	0.001
		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Female × probability			−0.024	−0.023			−0.038**	−0.038**
			(0.017)	(0.017)			(0.017)	(0.017)
Natural × probability			−0.047*	−0.047*			−0.076	−0.077
			(0.026)	(0.026)			(0.048)	(0.048)
Medicine × probability			0.121***	0.121***			0.038	0.038
			(0.045)	(0.045)			(0.063)	(0.063)
Arts × probability			−0.121***	−0.121***			−0.108***	−0.108***
			(0.041)	(0.041)			(0.031)	(0.031)
Other × probability			0.255***	0.270***			−0.176***	−0.181***
			(0.015)	(0.014)			(0.010)	(0.009)
Power × probability			−0.000	−0.000			−0.004*	−0.004*
			(0.004)	(0.004)			(0.002)	(0.002)
Masculinity × probability			−0.010***	−0.010***			−0.002	−0.002
			(0.003)	(0.003)			(0.002)	(0.002)
Prospect characteristics	✓	✓	✓	✓	✓	✓	✓	✓
Random slopes				✓				✓
Constant	0.012	0.061***	0.061***	0.061***	−0.079***	−0.079***	−0.036***	−0.032***
	(0.008)	(0.007)	(0.007)	(0.008)	(0.010)	(0.009)	(0.006)	(0.006)
Country VAR	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Individual VAR	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Residual VAR	0.041***	0.041***	0.040***	0.040***	0.043***	0.043***	0.043***	0.043***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	17,629	17,629	17,629	17,629	17,605	17,605	17,605	17,605
Subjects	2,939	2,939	2,939	2,939	2,939	2,939	2,939	2,939
Countries	30	30	30	30	30	30	30	30
Log likelihood	1,989.434	1,991.725	2,051.410	2,057.241	1,568.462	1,568.961	1,602.948	1,612.540

Standard errors, in parentheses. *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$. All interactions with probability are included win models 2 and 3, but only the significant ones are shown.

gains ($\chi^2(14) = 119.37, p < 0.001$) and losses ($\chi^2(14) = 67.97, p < 0.001$). Some interesting effects emerge (we show only significant interactions in the table to keep it tractable). For gains, we now find medical students to be less ambiguity averse than economics students, a correlation between a high score on masculinity and ambiguity aversion and a marginally significant positive correlation between uncertainty avoidance and ambiguity aversion (individualism is no longer significant). These effects are balanced by effects on a-insensitivity, with medical students being considerably more a-insensitive compared to economics students, and masculinity being associated with slightly reduced a-insensitivity. Once again, however, these effects are not reproduced for losses, where some other effects emerge, such as women being less a-insensitive than men and arts students being both more ambiguity averse than economics students and less a-insensitive.

Finally, model 4 allows for random slopes at the country level for some basic characteristics (biological and study characteristics). This further improves the fit relative to model 3 for both gains ($\chi^2(4) = 11.66, p = 0.020$) and losses ($\chi^2(4) = 19.18, p < 0.001$). More interesting, however, is to determine to what extent we have succeeded in explaining the individual-level variance. For gains, model 4 explains about 2.7% of the individual-level variance. For losses, the figure is 3.6%. These figures look low even compared to the low standards for risk preferences, where large unexplained heterogeneity at the individual level is the norm (von Gaudecker et al., 2011). Although they reach similar conclusions, l'Haridon and Vieider (in press) found several explanatory variables to be significant at the individual level, with remarkable consistency in these effects between gains and losses. It is also informative to once again consider the interclass correlation coefficient (ICC) at the individual level.[6] For ambiguity, we found it to be in the range of 0.15 to 0.18, indicating that any two randomly chosen choices from one randomly chosen individual would have a correlation of 0.15 to 0.18. For risk, the equivalent range is 0.38 to 0.40—still far from perfect, but a clear improvement over the figure for ambiguity.

We conclude this section in a similar vein as for the premium-level section above. While we can account for some heterogeneity in the data at the individual level recurring to observable characteristics of the decision makers, the proportion of variance thus explained is low in both absolute terms and in comparison to results for risk preferences.

E. *Explaining Country-Level Variance*

We once again start by quantifying the variance occurring at this level, based on the benchmark model, which in this case comprises both prospect characteristics as above and some basic individual characteristics (sex, age, and GPA). We find the ICC to be very low at 0.8% for both gains and losses.[7] This is, however, a relative figure, rendered artificially low by the massive amount of residual variance discussed above. Indeed, we have seen in the descriptive country comparison at the beginning of section III that countries do differ significantly in terms of their ambiguity attitudes (as is also apparent from the high statistical significance of the variance at the country level).[8] It thus seems well worth our while to try to account for this variance by the use of country-level indicators. Prime candidates for such indicators are economic, geographical, and institutional variables as used in the macroeconomic literature on comparative development and growth. Once again, we mainly focus on the extent to which we can explain the between-country variance.

Table 3 shows the regression results. The benchmark model again serves to quantify the variance occurring at this level, which we show multiplied by 1,000 in the table to be able to discern changes across regressions. Model 1 introduces income and institutional variables. The income variables include GDP per capita (in PPP and taking logs; World Bank data for 2010), as well as the Gini coefficient to capture income inequality, and a dummy capturing whether the university where the experiment was run was private (which is likely to indicate students who are well off relative to the rest of the country). The institutional data comprise democracy, a variable capturing the institutional and political stability over the previous decades, and legal origins dummies (La Porta et al., 1997). The regression explains a marginally significant proportion of the variance for gains ($\chi^2(7) = 13.66, p = 0.058$)[9], but not for losses ($\chi^2(7) = 9.36, p = 0.228$). For gains, we find a significant effect of GDP per capita, going in the direction of richer countries being more ambiguity averse. We also find some effects of legal origins, with subjects in countries with U.K. legal origins and socialist legal origins showing less ambiguity aversion than subjects in French legal origins countries. For losses, we find subjects from countries with more income inequality to be more ambiguity averse; we also find marginally significant effects of democracy (more institutionally stable countries are more ambiguity averse). We also find more ambiguity aversion in private universities, and socialist legal origins countries are more ambiguity averse for losses.

Model 2 further adds variables capturing genetic diversity and a geographical control. The former contain a measure of

[6] The ICC is defined as the proportion of overall variance captured at the individual level. Let σ_n be the variance at the individual level in a model empty of covariates, and let σ_c be the country-level variance and σ_ρ the residual varaince. We then have $ICC_n = \frac{\sigma_n}{\sigma_n+\sigma_c+\sigma_\rho}$.

[7] These figures are also low compared to previous results obtained for risk preferences. Using risk data from this same data set, l'Haridon and Vieider (in press) reported considerably more heterogeneity between individuals than between countries, with the country-level variance quantified at between 7% and 33% depending on the model parameter. Using a hypothetical question obtained from representative samples from 76 countries and 80,000 respondents, Falk et al. (2015) found the country-level variance to be 9%. Using a different survey question in a sample of over 100,000 repondents from 78 countries, Bouchouicha and Vieider (2017) found this figure to be 10%.

[8] It seems desirable to obtain a measure of between-country variation that is independent of the level of variance at the residual level. A very simple measure consists in pairwise between-country comparisons of average ambiguity premia aggregated per subject for gains and losses. Of all possible comparisons, 88% are significant for gains and 89% for losses (at the 5% level, without adjustment for multiple testing). The corresponding result is somewhat lower for risk, coming in at 72% for gains and 75% for losses. This shows that there is indeed substantial between-country variation of ambiguity attitudes.

[9] Given the highly significant correlations with, especially, GDP, it may appear odd that the improvement in variance explained is only marginally significant. This is in part due to the penalty imposed for using up several degrees of freedom. Indeed, if we add only GDP per capita in the regression, the proportion of variance explained is significant at conventional levels ($\chi^2(1) = 5.29, p = 0.022$).

TABLE 3.—COUNTRY-LEVEL REGRESSIONS

	Gains (Ambiguity Aversion)				Losses (Ambiguity Seeking)			
	Benchmark	Model 1	Model 2	Model 3	Benchmark	Model 1	Model 2	Model 3
Log GDP per capita		0.010**	0.024***	0.023***		0.005	−0.003	−0.003
		(0.005)	(0.009)	(0.009)		(0.006)	(0.007)	(0.007)
Probability × GDP				0.057***				0.042**
				(0.014)				(0.019)
Gini coefficient		0.003	0.003	0.003		−0.008**	−0.006**	−0.006**
		(0.004)	(0.004)	(0.004)		(0.003)	(0.003)	(0.003)
Private university		0.019	0.011	0.011		−0.019*	−0.014	−0.014
		(0.013)	(0.013)	(0.013)		(0.011)	(0.011)	(0.011)
Democracy		0.001	−0.001	−0.001		−0.003*	−0.002	−0.002
		(0.001)	(0.001)	(0.001)		(0.002)	(0.002)	(0.002)
British legal origins		−0.021*	−0.002	−0.002		0.006	0.002	0.002
		(0.011)	(0.009)	(0.009)		(0.009)	(0.008)	(0.008)
Socialist legal origins		−0.006	0.007	0.007		−0.030**	−0.033***	−0.033***
		(0.010)	(0.009)	(0.009)		(0.014)	(0.013)	(0.013)
German legal origins		−0.022***	−0.009	−0.009		−0.007	−0.009	−0.009
		(0.008)	(0.007)	(0.007)		(0.010)	(0.010)	(0.010)
OPEC dummy			−0.057**	−0.057**			0.025	0.025
			(0.022)	(0.022)			(0.020)	(0.020)
Predicted genetic diversity			−0.342***	−0.342***			0.030	0.030
			(0.111)	(0.111)			(0.127)	(0.127)
Predicted genetic diversity squared			0.350***	0.350***			−0.034	−0.034
			(0.112)	(0.112)			(0.125)	(0.125)
Degrees latitude			−0.001*	−0.001*			0.001*	0.001*
			(0.000)	(0.000)			(0.000)	(0.000)
Probability | country random slope			✓				✓	
Prospect controls	✓	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Constant	0.059***	−0.031	−0.127*	−0.127*	−0.034***	−0.062	−0.007	−0.007
	(0.005)	(0.048)	(0.065)	(0.065)	(0.005)	(0.046)	(0.049)	(0.049)
Country VAR (×1000)	0.401***	0.0202***	0.061***	0.061***	0.394***	0.252***	0.217***	0.217***
	(0.201)	(0.102)	(0.048)	(0.048)	(0.241)	(0.148)	(0.139)	(0.139)
Subject VAR	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Residual VAR	0.041***	0.041***	0.041***	0.040***	0.043***	0.043***	0.043***	0.043***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	17,629	17,629	17,629	17,629	17,605	17,605	17,605	17,605
Subjects	2,939	2,939	2,939	2,939	2,939	2,939	2,939	2,939
Countries	30	30	30	30	30	30	30	30
Log likelihood	1,983.099	1,989.931	1,995.732	2,099.960	1,565.737	1,570.417	1,571.647	1,648.301

Standard errors in parentheses. *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$. Country-level variance is shown multiplied by 1,000 to be able to assess changes across regressions.

predicted genetic diversity developed by Ashraf and Galor (2013) and its square. Similar measures were used by Spolaore and Wacziarg (2009) to explain income differences deriving from trading behavior. Becker et al. (2015) found correlations of such genetic measurements with risk preferences, explaining the correlations by a "serial founder effect." The geographical variable consists of absolute distance from the equator (in degrees of latitude). For gains, these variables further reinforce the effect of GDP. Subjects in OPEC countries are found to be less ambiguity averse than what their GDP would suggest. Genetic diversity shows an inverse U-shape, whereby moderately diverse countries (typically in Europe) are the least ambiguity averse (after controlling for the other variables, including GDP), and the most diverse countries in Africa and least diverse ones in Latin America and East Asia are the most ambiguity averse. This is interesting because this parallels the relationship between genetic diversity and GDP per capita pointed out by Ashraf and Galor (2013). We are, however, unable to replicate these effects for losses. Once again, the model significantly improves the fit over model 1 for gains ($\chi^2(3) = 11.60, p = 0.021$) but not for losses ($\chi^2(3) = 2.46, p = 0.652$).

Finally, in model 3, we add the random slopes of the prospect characteristics to account for differences across countries. In addition, we try to explain difference in a-insensitivity across countries by adding an interaction with GDP per capita to the fixed part of the model. This again significantly improves model fit—this time for both gains ($\chi^2(2) = 208.46, p < 0.001$) and for losses ($\chi^2(2) = 153.31, p < 0.001$). The effects previously found remain generally stable. The interaction term (or crossed effect) between GDP per capita and probability furthermore shows that a-insensitivity is larger in our subject pool for richer countries, an effect that holds for both gains and losses.

We conclude this section by assessing the success of our modeling exercise. Overall, our heaviest model 3 explains fully 85% of the between-country variance for gains and a somewhat more modest 45% for losses. The results are thus qualitatively different from those seen at lower levels.

Indeed, the proportion of variance explained is higher even than for risk, albeit similar in the sense that aggregated preferences prove easier to explain than preferences at the individual level. One explanation for this may be the lower total variance to be explained. Another may lie in the sheer number of macroeconomic characteristics to pick from. Ultimately, however, we believe that aggregating ambiguity preferences at this level achieves something very important that could not be achieved at lower levels: the emergence of stable trends from a sea of noise.

IV. Discussion and Conclusion

We found our measures of ambiguity aversion to reproduce typical aggregate trends observed in previous experiments. Aggregating across different prospects and across individuals, we found ambiguity aversion to be the prevalent pattern for gains in all countries. For losses, we still found a prevalence of ambiguity aversion in the aggregate. This ambiguity aversion was, however, much weaker than for gains, and for some countries, ambiguity neutrality could not be rejected (with one country being significantly ambiguity seeking). We furthermore replicated patterns whereby ambiguity attitudes change less than proportionally with the probability, known as a-insensitivity, in nearly all countries, and for both gains and losses. However, we did not find significant levels of ambiguity seeking for either losses or small probabilities, as reported in most of the previous literature (see Trautmann & van de Kuilen, 2015, for a review). This may be due to our measurement tasks, which are not directly comparative in nature, so that they can be seen as a lower bound on ambiguity attitudes.

When it came to explaining the patterns found at the macroeconomic level, our regressions proved quite successful. GDP per capita was significantly related to ambiguity aversion, with poorer countries being significantly less ambiguity averse. This parallels recent findings according to which poorer countries tend to be less risk averse than rich ones (l'Haridon & Vieider, in press; Vieider et al., 2018). We further found an inverse U-shaped effect of predicted genetic diversity. This effect seems interesting in conjunction with a similarly shaped relationship between the same genetic diversity measure and GDP uncovered by Ashraf and Galor (2013). Indeed, there may be wider implications considering the link of genetic measures to trade (Spolaore & Wacziarg, 2009) and the complementary relationship between unfamiliarity and trade (Huang, 2007). However, it seems premature to overemphasize this point based on a single data set.

Once we move from the group levels to the individual and premium level, our results change. Large amounts of heterogeneity are the preponderant finding at the individual level. We succeed in explaining very little of the individual-level variance. Our attempts show limited success not only in absolute terms but even compared to similar exercises carried out for risk preferences, which may themselves be considered to have had mixed success. This conclusion is not unique to our data, since the few previous studies investigating demographic correlates of ambiguity preferences do not seem to have uncovered any regularities either (see the discussion of Trautmann & van de Kuilen, 2015, on this point). Any effects of demographics we find at the individual level are inconsistent between gains and losses and tend to be economically weak.

The picture becomes even bleaker when moving from the subject level to the premium level. Even after including all conceivable prospect characteristics and allowing those characteristics to have individual-specific effects, we manage to explain a meager 19% of the variance at that level for gains, and not much more than 10% for losses. For risk, the result using similar models (which have not been optimized specifically to account for risk preferences) is close to 75%. In other words, the amount of variance we manage to explain at the premium level is particularly low when compared to similar exercises for risk preferences. There is only one conclusion to be drawn from this: the data are plagued by massive amounts of noise. These findings are consistent with those of Stahl (2014), who also found high levels of noise in measures of ambiguity aversion.

Our results can shed some light on the many unsuccessful attempts at detecting stable correlates of ambiguity preferences at the individual level. Borghans et al. (2009), Binmore, Stewart, and Voorhoeve (2012), and Sutter et al. (2013) all express some frustration in detecting such correlates. Dimmock, Kouwenberg, and Wakker (2015) and Dimmock, Kouwenberg, Mitchell, et al. (2015) report better success at predicting economic outcomes. While their main focus is on correlates of ambiguity attitudes, their results are more in agreement with ours than one might infer from the tone of the papers. For instance, in the Dutch sample, they find that a comprehensive set of observable characteristics of respondents could account for between 2.2% and 5.4% of the variance in the different parameters. In the U.S. population, individual characteristics can explain 2.6% of the overall variance. Furthermore, for losses, the authors could not find any correlates of ambiguity attitudes except for risk aversion. Overall, our results thus appear to be in line with the bulk of the previous literature when it comes to the detection of correlates of ambiguity attitudes.

This leaves the question of whether these low correlations are due to noise in the data or to heterogeneity that is orthogonal to observable characteristics. While we cannot provide a definite answer to this question, we strongly suspect that noise plays an important role. To see this, consider the correlation of ambiguity premia for (1/8:20;0) and (1/8:20; 5) and for (7/8:20; 0) and (7/8:20; 5). These two tasks are very similar and can thus be used to examine the test-retest reliability of our measures. The correlation we find between the two is, however, very low, at $\rho = 0.24$ and $\rho = 0.53$, respectively (Spearman rank correlations; results for losses are similar and can be found in the online appendix). Out-of-sample predictions based on parameters estimated from our model perform even more poorly (see the online appendix for details). According to classical measurement theory, this means that any correlation with observable characteristics would be attenuated significantly due to measurement issues of ambiguity attitudes. This in turn means that noise in measurement is almost certainly a substantial part of the story when it comes to explaining the weak correlations with observable characteristics we and others have observed.

Given the good results at the aggregate level and the much more stable trends uncovered for risk in this same data set (l'Haridon & Vieider, in press), we think that the high levels of noise are unlikely to result from particular measurement problems affecting only our experiment. Noise may, however, be exacerbated by the separate measurement of valuations for risk and for uncertainty, from which ambiguity attitudes are then derived. The CEs for risk and for uncertainty may be affected by separate error terms, which could increase the incidence of noise. An examination of violations of first-order stochastic dominance under risk and uncertainty, however, does not reveal any unusual noisiness in our measurements (see the online appendix for details). An interesting question is nevertheless to what extent our particular measurement method may have influenced the noise levels we find and whether these noise levels are absent in tasks involving direct comparisons between risky and uncertainty tasks, such as matching probabilities.

In particular, Dimmock, Kouwenberg, Mitchell, et al. (2015) report better success in explaining variations in matching probabilities across the probability spectrum. The difference between the two tasks may, however, be smaller than one might think. Binmore et al. (2012) draw similar conclusions to ours, even though they use matching probabilities to measure ambiguity attitudes. Dimmock, Kouwenberg, Mitchell, et al. (2015), using two choices as consistency checks, report that fully 42% of such responses result in inconsistencies. We can compare these findings to our own data by examining the consistency of choices between the premia used in the analysis and the premia for prospects with nonzero lower outcomes. Using all four such pairs, we find violation rates between 24% and 28% (see the online appendix for details). These rates are indeed remarkably consistent with those reported by Dimmock, Kouwenberg, Mitchell, et al. (2015). While only a direct methodological comparison can give conclusive evidence about the relative merits of different measurement tasks, our tentative conclusion from this evidence is thus that the difference seems to be at best one of degree. For instance, Dimmock, Kouwenberg, Mitchell, et al. (2015) report that 43% of the within-subject variance can be explained by probability variations. While this is certainly better than our 20%, it still falls far short of the 75% we observe under risk.[10] Ultimately, only direct comparative evidence will allow for a clear conclusion on the relative merits of different methods used to measure ambiguity attitudes.

Our conclusions about measuring ambiguity attitudes are rather negative. We do not want to say with this that ambiguity does not matter. The point is, rather, that processes resembling ambiguity are quite artificial and may thus have limited real-world applications. There furthermore appear to be severe measurement problems. Ambiguity may well matter where it occurs naturally in the real world; for example, Kunreuther et al. (1995) provide evidence that the presence of ambiguity about the precise probabilities underlying a process affects the pricing decisions of insurance underwriters. The way forward may then be to investigate naturally occurring uncertainty rather than artificial ambiguity that is rare in the real world. This will mean focusing on natural sources of uncertainty, as some studies have already done (Abdellaoui et al., 2011; Baillon et al., 2018). Baillon et al. (2016) proposed a method for the nonparametric measurement of ambiguity attitudes and showed that it exhibits high levels of measurement reliability. This may also mean moving away from a comparison point of risk, which can hardly ever be found in reality, and toward varying degrees of ambiguity underlying outcome-generating processes. Time will tell whether such approaches will indeed perform better in terms of the external validity of experimentally measured preferences.

[10] The difference between our figure and theirs may further be overestimated by the different methods we use to account for explained variance. For instance, Dimmock et al. use an "overall R^2" to account for the explained variance. This captures changes in within- and between-subject variance, while our measure only looks at one type of variance in isolation.

REFERENCES

Abdellaoui, Mohammed, Aurélien Baillon, Lætitia Placido, and Peter P. Wakker, "The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation," *American Economic Review* 101 (2011), 695–723.

Ahn, David, Syngjoo Choi, Douglas Gale, and Shachar Kariv, "Estimating Ambiguity Aversion in a Portfolio Choice Experiment," *Quantitative Economics* 5 (2014), 195–223.

Alary, David, Christian Gollier, and Nicolas Treich, "The Effect of Ambiguity Aversion on Insurance and Self-Protection," *The Economic Journal* 123 (2013), 1188–1202.

Allen, Mary J., and Wendy M. Yen, *Introduction to Measurement Theory* (Long Grove, IL: Waveland Press, 2001).

Ashraf, Quamrul, and Oded Galor, "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review* 103 (2013), 1–46.

Baillon, Aurélien, and Han Bleichrodt, "Testing Ambiguity Models through the Measurement of Probabilities for Gains and Losses," *American Economic Journal: Microeconomics* 7 (2015), 77–100.

Baillon, Aurélien, Han Bleichrodt, Umut Keskin, Olivier l'Haridon, and Chen Li, "Learning under Ambiguity: An Experiment Using Initial Public Offerings on a Stock Market," *Management Science* 64 (2018), 2181–2198.

Baillon, Aurélien, Zhenxing Huang, Asli Selim, and Peter P. Wakker, "Measuring Ambiguity Attitudes for All (Natural) Events," Erasmus University Rotlerdam working paper (2016).

Baltussen, Guido, Martijn J. van den Assem, and Dennie van Dolder, "Risky Choice in the Limelight," this REVIEW 98 (2016), 318–332.

Becker, Anke, Thomas Dohmen, Benjamin Enke, and Armin Falk, "The Ancient Origins of the Cross-Country Heterogeneity in Risk Preferences," University of Bonn working paper (2015).

Berger, Loïc, Han Bleichrodt, and Louis Eeckhoudt, "Treatment Decisions under Ambiguity," *Journal of Health Economics* 32 (2013), 559–569.

Binmore, Ken, Lisa Stewart, and Alex Voorhoeve, "How Much Ambiguity Aversion?" *Journal of Risk and Uncertainty* 45 (2012), 215–238.

Borghans, Lex, James J. Heckman, Bart H. H. Golsteyn, and Huub Meijers, “Gender Differences in Risk Aversion and Ambiguity Aversion,” *Journal of the European Economic Association* 7 (2009), 649–658.

Bouchouicha, Ranoua, Peter Martinsson, Haileselassie Medhin, and Ferdinand Vieider, “Stake Effects on Ambiguity Attitudes for Gains And Losses,” *Theory and Decision* 83 (2017), 19–35.

Bouchouicha, Ranoua, and Ferdinand M. Vieider, “Growth, Entrepreneurship, and Risk Tolerance: A Risk-Income Paradox,” University of Reading working paper (2017).

Brislin, Richard W., “Back-Translation for Cross-Cultural Research,” *Journal of Cross-Cultural Psychology* 1 (1970), 185–216.

Bryan, Gharad, “Ambiguity and Insurance,” London School of Economics and Political Science working paper (2010).

Charness, Gary, Edi Karni, and Dan Levin, “Ambiguity Attitudes and Social Interactions: An Experimental Investigation,” *Journal of Risk and Uncertainty* 46 (2013), 1–25.

Chow, Clare Chua, and Rakesh K. Sarin, “Comparative Ignorance and the Ellsberg Paradox,” *Journal of Risk and Uncertainty* 22 (2001), 129–139.

——— “Known, Unknown, and Unknowable Uncertainties,” *Theory and Decision* 52 (2002), 127–138.

Cohen, Michele, Jean-Yves Jaffray, and Tanios Said, “Experimental Comparison of Individual Behavior under Risk and Uncertainty for Gains and for Losses,” *Organizational Behavior and Human Decision Processes* 39 (1987), 1–22.

Dimmock, Stephen G., Roy Kouwenberg, Olivia S. Mitchell, and Kim Peijnenburg, “Estimating Ambiguity Preferences and Perceptions in Multiple Prior Models: Evidence from the Field,” *Journal of Risk and Uncertainty* 51 (2015), 219–244.

Dimmock, Stephen G., Roy Kouwenberg, and Peter P. Wakker, “Ambiguity Attitudes in a Large Representative Sample,” *Management Science* 62 (2015), 1363–1380.

Easly, David, and Maureen O’Hara, “Ambiguity and Nonparticipation: The Role of Regulation,” *Review of Financial Studies* 22 (2009), 1817–1843.

Ellsberg, Daniel, “Risk, Ambiguity and the Savage Axioms,” *Quarterly Journal of Economics* 75 (1961), 643–669.

Etchart-Vincent, Nathalie, and Olivier l’Haridon, “Monetary Incentives in the Loss Domain and Behavior toward Risk: An Experimental Comparison of Three Reward Schemes Including Real Losses,” *Journal of Risk and Uncertainty* 42 (2011), 61–83.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, “The Nature and Predictive Power of Preferences: Global Evidence,” University of Bonn working paper (2015).

Fox, Craig R., and Amos Tversky, “Ambiguity Aversion and Comparative Ignorance,” *Quarterly Journal of Economics* 110 (1995), 585–603.

Friedl, Andreas, Katharina Lima de Miranda, and Ulrich Schmidt, “Insurance Demand and Social Comparison: An Experimental Analysis,” *Journal of Risk and Uncertainty* 48 (2014), 97–109.

Frisch, Deborah, and Jonathan Baron, “Ambiguity and Rationality,” *Journal of Behavioral Decision Making* 1 (1988), 149–157.

Giné, Xavier, Robert Townsend, and James Vickery, “Patterns of Rainfall Insurance Participation in Rural India,” *World Bank Economic Review* 22 (2008), 539–566.

Hofstede, Geert H., *Culture’s Consequences: International Differences in Work-Related Values* (Thousand Oaks, CA: Sage, 1980).

Huang, Rocco, “Distance and Trade: Disentangling Unfamiliarity Effects and Transport Cost Effects,” *European Economic Review* 51 (2007), 161–181.

Kunreuther, Howard, Jacqueline Meszaros, Robin M. Hogarth, and Mark Spranca, “Ambiguity and Underwriter Decision Processes,” *Journal of Economic Behavior and Organization* 26 (1995), 337–352.

La Porta, Rafael, Florencio Lopez-De-Silanes, Andrei Shleifer, and Robert W. Vishny, “Legal Determinants of External Finance,” *Journal of Finance* 52 (1997), 1131–1150.

l’Haridon, Olivier, and Ferdinand M. Vieider, “All over the Map: Heterogeneity of Risk Preferences across Individuals, Contexts, and Countries,” *Quantitative Economics*, in press.

Mukerji, Sujoy, “Ambiguity Aversion and Incompleteness of Contractual Form,” *American Economic Review* 88 (1998), 1207–1231.

Muthukrishnan, A. V., Luc Wathieu, and Jing Alison Xu, “Ambiguity Aversion and the Preference for Established Brands,” *Management Science* 55 (2009), 1933–1941.

Oechssler, Jörg, and Alex Roomets, “A Test of Mechanical Ambiguity,” *Journal of Economic Behavior and Organization* 119 (2015), 153–162.

Powell, Melanie, and David Ansic, “Gender Differences in Risk Behaviour in Financial Decision-Making: An Experimental Analysis,” *Journal of Economic Psychology* 18 (1997), 605–628.

Rieger, Marc Oliver, Mei Wang, and Thorsten Hens, “Risk Preferences around the World,” *Management Science* 61 (2014), 637–648.

Snijders, Tom A. B., and Roel J. Bosker, *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling,* 2nd ed. (London: Sage, 2012).

Spolaore, Enrico, and Romain Wacziarg, “The Diffusion of Development,” *Quarterly Journal of Economics* 124 (2009), 469–529.

Stahl, Dale O., “Heterogeneity of Ambiguity Preferences,” this REVIEW 96 (2014), 609–617.

Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler, and Stefan T. Trautmann, “Impatience and Uncertainty: Experimental Decisions Predict Adolescents’ Field Behavior,” *American Economic Review* 103 (2013), 510–531.

Trautmann, Stefan T., and Gijs van de Kuilen, “Ambiguity Attitudes,” in Gideon Keren and George Wu, eds., *The Wiley Blackwell Handbook of Judgment and Decision Making* (Hoboken, NJ: Wiley Blackwell, 2015).

Trautmann, Stefan T., Ferdinand M. Vieider, and Peter P. Wakker, “Preference Reversals for Ambiguity Aversion,” *Management Science* 57 (2011), 1320–1333.

Tversky, Amos, and Craig R. Fox, “Weighing Risk and Uncertainty,” *Psychological Review* 102 (1995), 269–283.

Vieider, Ferdinand M., “Moderate Stake Variations for Risk and Uncertainty, Gains and Losses: Methodological Implications for Comparative Studies,” *Economics Letters* 117 (2012), 718–721.

Vieider, Ferdinand M., Abebe Beyene, Randall A. Bluffstone, Sahan Dissanayake, Zenebe Gebreegziabher, Peter Martinsson, and Alemu Mekonnen, “Measuring Risk Preferences in Rural Ethiopia,” *Economic Development and Cultural Change* 66 (2018), 417–446.

Vieider, Ferdinand M., Thorsten Chmura, Tylor Fisher, Takao Kusakawa, Peter Martinsson, Frauke Mattison Thompson, and Adewara Sunday, “Within- versus Between-Country Differences in Risk Attitudes: Implications for Cultural Comparisons,” *Theory and Decision* 78 (2015), 209–218.

Vieider, Ferdinand M., Mathieu Lefebvre, Ranoua Bouchouicha, Thorsten Chmura, Rustamdjan Hakimov, Michal Krawczyk, and Peter Martinsson, “Common Components of Risk and Uncertainty Attitudes across Contexts and Domains: Evidence from 30 Countries,” *Journal of the European Economic Association* 13:3 (2015), 421–452.

von Gaudecker, Hans-Martin, Arthur van Soest, and Erik Wengström, “Heterogeneity in Risky Choice Behaviour in a Broad Population,” *American Economic Review* 101 (2011), 664–694.