Maastricht University

The nature and predictive power of preferences : global evidence

Citation for published version (APA):

Falk, A., Becker, A., Dohmen, T. J., Enke, B., Huffman, D., & Sunde, U. (2015). *The nature and predictive power of preferences : global evidence*. ROA. ROA Research Memoranda No. 012
https://doi.org/10.26481/umaror.2015012

Document status and date:
Published: 01/01/2015

DOI:
10.26481/umaror.2015012

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 20 Apr. 2024

Maastricht University

Research Centre for Education and the Labour Market | ROA

The nature and predictive power of preferences: Global evidence

Armin Falk
Anke Becker
Thomas Dohmen
Benjamin Enke
David Huffman
Uwe Sunde

ROA Research Memorandum

ROA-RM-2015/12

Researchcentrum voor Onderwijs en Arbeidsmarkt | ROA
Research Centre for Education and the Labour Market | ROA

The nature and predictive power of preferences: Global evidence

Armin Falk
Anke Becker
Thomas Dohmen
Benjamin Enke
David Huffman
Uwe Sunde

ROA-RM-2015/12*
November 2015

Research Centre for Education and the Labour Market
Maastricht University
P.O. Box 616, 6200 MD Maastricht, The Netherlands
T +31 43 3883647 F +31 43 3884914

secretary-roa-sbe@maastrichtuniversity.nl
www.roa.nl

* The ROA Research Memorandum Series was created in order to make research results available for discussion, before those results are submitted for publication in journals.

Abstract

The nature and predictive power of preferences: Global evidence**

This paper presents the Global Preference Survey, a globally representative dataset on risk and time preferences, positive and negative reciprocity, altruism, and trust. We collected these preference data as well as a rich set of covariates for 80,000 individuals, drawn as representative samples from 76 countries around the world, representing 90 percent of both the world's population and global income. The global distribution of preferences exhibits substantial variation across countries, which is partly systematic: certain preferences appear in combination, and follow distinct economic, institutional, and geographic patterns. The heterogeneity in preferences across individuals is even more pronounced and varies systematically with age, gender, and cognitive ability. Around the world, our preference measures are predictive of a wide range of individual-level behaviors including savings and schooling decisions, labor market and health choices, prosocial behaviors, and family structure. We also shed light on the cultural origins of preference variation around the globe using data on language structure.

JEL classification: D01, D03, F00
Keywords: economic preferences, cultural variation

Armin Falk
Department of Economics
University of Bonn
Adenauerallee 24-42
53113 Bonn
Germany
armin.falk@uni-bonn.de

Anke Becker
Department of Economics
University of Bonn
Adenauerallee 24-42
53113 Bonn
Germany
anke.becker@uni-bonn.de

Thomas Dohmen
Department of Economics
University of Bonn
Adenauerallee 24-42
53113 Bonn
Germany
tdohmen@uni-bonn.de

Benjamin Enke
Department of Economics
University of Bonn
Adenauerallee 24-42
53113 Bonn
Germany
benjamin.enke@uni-bonn.de

** For valuable comments and discussions we are grateful to Doug Bernheim, Johannes Hermle, Benedikt Herrmann, Fabian Kosse, Andrei Shleifer, and seminar participants at Caltech, Konstanz, and UC San Diego. Ammar Mahran provided oustanding research assistance. Armin Falk acknowledges financial support from the European Research Council through ERC # 209214.

David Huffman
Department of Economics
University of Pittsburgh
230 South Bouquet Street
Pittsburg, PA 15260
United States of America
huffmand@pitt.edu

Uwe Sunde
Department of Economics
University of Munich
Geschwister-Scholl Platz 1
80333 München
Germany
uwe.sunde@lmu.de

1 Introduction

This paper presents the Global Preference Survey (GPS), a novel and unique globally representative dataset. The data include measures of risk preference, time preference, positive and negative reciprocity, altruism, and trust that we collected for 80,000 individuals, drawn as representative samples in each of 76 countries. The coverage of countries spans all continents, a broad set of cultures, a wide range of development levels, and represents about 90 percent of both the world's population and global income, making the data also representative across countries. The underlying survey measures were selected and tested through a rigorous ex ante experimental validation procedure involving real monetary stakes, so that the survey items have a demonstrated ability to capture actual heterogeneity in state-of-the-art experiments with financial incentives (Falk et al., 2015). To ensure comparability of preference measures across countries, the elicitation followed a standardized protocol that was implemented through the professional infrastructure of the Gallup World Poll. Monetary stakes related to the elicitation involved comparable values in terms of purchasing power across countries, and the survey items were culturally neutral and translated using state-of-the-art procedures. In addition, pre-tests in 22 countries of various cultural heritage revealed the broad applicability of our survey items. In consequence, the resulting dataset provides an ideal basis for the first systematic investigation of the distribution, determinants, and predictive power of preferences around the world.

Using these data, we provide evidence for several novel findings, both at the country and at the individual level. First, for each of the six traits, we document a substantial variation not just across individuals, but also across entire countries. Second, we show that this cross-country heterogeneity is at least partly systematic and follows pronounced economic, geographic and cultural patterns. All preferences are significantly associated with important country-level variables including per capita income, democracy, inequality, redistributive policies, religion, and geographic or climatic variables. Third, the various preference measures are correlated, giving rise to distinct "preference profiles" of groups of countries. Fourth, in spite of the substantial between-country variation, most of the total individual-level variation in all preferences is due to within-country heterogeneity. Fifth, investigating the structure of this individual-level variation, we find that in the world population as a whole, all of the preferences are systematically related to individual characteristics. For instance, women tend to be less patient and more risk averse, and exhibit stronger social predispositions, than men. Patience is hump-shaped in age, while risk taking as well as positive and negative reciprocity are lower for older people. Self-reported cogni-

tive skills positively correlate with patience, risk taking, and all social preferences. Sixth, we provide evidence of heterogeneity across countries underlying the strong *average* patterns of the individual-level correlates of preferences in the world population as a whole. We show that while some relationships between preferences and sociodemographics (such as between risk aversion and gender) are common to almost all cultures, others appear more culturally or institutionally specific. For example, patience and positive reciprocity exhibit a hump-shaped relationship with age in developed countries that is almost entirely absent in developing nations. Seventh, we show that individual-level preferences are also significantly correlated with household income, subjective perceptions of safety and health, as well as religious affiliation. Eighth, we examine the predictive power of preference heterogeneity for economic behaviors. Around the world, patient individuals are more likely to save and have higher educational attainment; more risk tolerant individuals are more likely to become self-employed and to be smokers; and social preferences are highly predictive of a broad range of prosocial behaviors and outcomes such as donating, volunteering time, assisting strangers, helping friends and relatives, or family structure. Finally, we shed light on the cultural origins of the global preference variation by making use of information on language structure: people who speak languages that do not require an explit coding of the future are more patient, positively reciprocal, trusting, and altruistic, both across and within countries.

Our analysis provides the first systematic assessment of the nature and explanatory power of preference heterogeneity around the world. The underlying data are, however, well-suited for a much broader research agenda on the determinants and implications of certain preference profiles. Going forward, the data lend themselves to investigations both at the micro- and the macro-level. At the micro level, several studies have examined individual-level preference heterogeneity and the corresponding correlates, like gender, in specific samples and cultures (see, e.g., Barsky et al., 1997; Frederick, 2005; Croson and Gneezy, 2009; Dohmen et al., 2008, 2010, 2011). However, the previous lack of data has prevented systematic investigations of the cultural specificity of such findings, an issue that is relevant for understanding the cultural or biological mechanisms through which individual characteristics like age or gender might shape preferences. Our results highlight some cases in which generalizing beyond single countries can be particularly misleading, because it ignores the country and population specificity of such effects. At the same time, the data show how some relationships are close to universal. Likewise, while previous work has provided evidence that preferences are predictive of important economic decisions, it has been an open question whether preferences are uniformly predictive of

behaviors across cultures and institutional backgrounds, and to which extent they shape heterogeneity in life outcomes.[1]

The GPS data may also prove valuable for research in cultural economics and political economy (Guiso et al., 2006; Fernández, 2011; Alesina and Giuliano, forthcoming; Giuliano and Nunn, 2013). To date, empirical research into the roots of cross-country variation in preferences has been impeded by a lack of appropriate measures and representative sampling; contributions on the cross-country heterogeneity in preferences have typically made use of small and non-representative samples in a limited set of countries (Roth et al., 1991; Henrich et al., 2001; Herrmann et al., 2008). Accordingly, researchers interested in the determinants and implications of cultural variation have considered variables such as female labor force participation, fertility, individualism, and future-orientation (Giuliano, 2007; Fernández and Fogli, 2009; Gorodnichenko and Roland, 2011; Alesina et al., 2013; Chen, 2013; Alesina et al., forthcoming; Galor and Özak, 2014), but have not studied the preference component of culture. The data of the GPS, which feature 80,000 individuals from various cultural backgrounds, are likely to produce new insights in this direction.

Apart from such micro-level analyses, the representative cross-country nature of our data also permits an investigation of the relationships of preferences to aggregate economic and social outcomes across countries, which to date is uncharted territory.[2] Motivated by the strong and systematic correlations reported in this paper, the preference data may be used both in an attempt to explain cross-country differences in aggregate outcomes, and in controlling for preference differences when interest lies in identifying other relationships.

The remainder of the paper proceeds as follows. In the next section, we present the Global Preference Survey dataset. In Section 3, we describe the nature of cross-country variation in preferences. Section 4 studies the relationship between preferences and individual characteristics, while Section 5 investigates the relationships between preferences and behaviors. In Section 6, we analyze the relationship between preferences and language structure. Section 7 concludes.

[1] Time preference correlates with outcomes ranging from savings to Body Mass Index (Ventura, 2003; Kirby and Petry, 2004; Borghans and Golsteyn, 2006; Eckel et al., 2005; Chabris et al., 2008; Tanaka et al., 2010; Meier and Sprenger, 2010; Sutter et al., 2013; Golsteyn et al., 2014). Risk preferences are related to various risky decisions, including being self-employed, migrating, and holding risky assets (See, e.g., Barsky et al., 1997; Bonin et al., 2007; Guiso and Paiella, 2008; Dohmen et al., 2011). Social preferences are correlated with cooperative behaviors in various aspects of life including in the workplace (Dohmen et al., 2009; Rustagi et al., 2010; Carpenter and Seki, 2011; Kosfeld and Rustagi, 2015).

[2] An exception is the burgeneoing literature on the importance of trust, see, e.g., Knack and Keefer (1997); Guiso et al. (2009); Algan and Cahuc (2010).

2 Dataset

2.1 General Data Characteristics

The Global Preference Survey (GPS) is a new globally representative survey designed to measure respondents' time preferences, risk preferences, social preferences, and trust. The GPS data were collected within the framework of the Gallup World Poll, which surveys representative population samples in a large number of countries about social and economic issues on an annual basis. In 2012, we added the GPS to the World Poll's questionnaire in 76 countries, so that the survey items were fielded through the existing professional infrastructure of one of the world's leading global survey companies. Four noteworthy features characterize the preference data: (i) representative population samples within each country, (ii) geographical and economic representativeness in terms of countries covered, (iii) a rigorous experimental validation and selection procedure of the underlying survey items, and (iv) a standardized data collection protocol across countries. We discuss these features in the following; in addition, Appendix A contains an extensive documentation of the data-collection process as well as additional details on the survey measures.

First, we measure preferences in large representative population samples in each country.[3] The median sample size was 1,000 participants per country, in 76 countries all over the world.[4] In total, we collected preference measures for more than 80,000 participants worldwide. Respondents were selected through probability sampling; ex-post representativeness of the data can be achieved using weights provided by Gallup.[5] In sum, our data allow for valid inferences about the distribution of preferences in each country as well as about between-country differences in preferences.

Second, the data are characterized by geographical representativeness in terms of the countries being covered. The sample of 76 countries is not restricted to Western industrialized nations, but covers all continents, various cultures, and different levels of development. Specifically, our sample includes 15 countries from the Americas, 25 from Europe, 22 from Asia and Pacific, as well as 14 African countries, 11 of which are Sub-Saharan. This set of countries covers about 90% of both the world population and global income.

Third, we designed, tested, and selected the survey items of the GPS using a rig-

[3]Data sets that contain preference measures for several countries typically come from small- or medium-scale surveys or experiments and are based on student or other convenience samples (e.g., Wang et al. (2011), Rieger et al. (forthcoming), Vieider et al. (2015), Vieider et al. (2014).

[4]Notable exceptions include China (2,574 obs.), Haiti (504 obs.), India (2,539 obs.), Iran (2,507 obs.), Russia (1,498 obs.), and Suriname (504 obs.).

[5]These weights are constructed to render the observations representative in terms of age, gender, income, education, and geographic location.

orous ex-ante experimental validation and selection procedure (for details see Falk et al., 2015). While items in international surveys are frequently designed based on introspective arguments of plausibility or relevance, our items are the result of an explicit formal selection procedure, which also ensures that the resulting measures are predictive of actual preferences as measured through state-of-the-art experiments. Arguably, such an ex-ante validation of survey items constitutes a significant methodological advance over the ad-hoc selection of questions for surveys. As detailed in Falk et al. (2015), in the validation procedure, experimental subjects completed incentivized choice experiments to measure their preference parameters, and also answered a large battery of candidate survey questions. For each preference, those survey items that jointly perform best in predicting the financially incentivized behavior were selected to form the preference survey module.[6] Thus, the module does not only consist of survey questions that predict behavior, but is composed of the best behavioral predictors out of a large set of candidate measures.

In a next step, the GPS was developed for implementation in the Gallup World Poll. To this end, Gallup conducted pre-tests in 22 countries of various cultural heritage, in order to ensure the implementability of the module in the available survey time of 7 to 8 minutes, and to test whether respondents of culturally and economically heterogeneous background understand and interpret the items adequately (see Appendix A.3 for details). Other measures taken to ensure that the survey items were comparable across cultures included: (i) translation of all items back and forth in an iterative process using Gallup's regular translation scheme, and (ii) calibration of monetary values used in the survey questions according to median household income for each country.[7] Finally, the interviews for the World Poll 2012 took place face-to-face or via telephone by professional interviewers. Thus, the survey items were fielded in a comparable way using a standardized procedure across countries.

[6] We excluded quantitative measures that require long and complex instructions, or which had shorter alternative quantitative measures that were close substitutes, from the set of candidate measures before the item selection procedure was conducted.

[7] As a benchmark, we used the monetary amounts in Euro that were offered in the validation study in Germany. Since monetary amounts used in the validation study with the German sample were round numbers to facilitate easy calculations (e.g., the expected return of a lottery with equal chances of winning and losing) and to allow for easy comparisons (e.g., 100 Euro today versus 107.50 in 12 months), we also rounded monetary amounts in all other countries to the next "round" number. While this necessarily resulted in some (very minor) variations in the real stake size between countries, it minimized cross-country differences in the understanding the quantitative items due to difficulties in assessing the involved monetary amounts.

2.2 Preference Measures

For each preference, we obtain a final individual-level measure by weighing responses to multiple survey items using the weights obtained from the experimental validation procedure. These weights are based on an OLS regression of observed behavior in the financially incentivized experiments on the respective survey measures (see Falk et al., 2015, for details). We first standardize individual-level responses to all items (i.e., compute z-scores) and then weigh these standardized responses using the OLS weights to derive the best predictor of observed experimental behavior. Finally, for ease of interpretation, each preference measure is again standardized at the individual level, so that, by construction, each preference has a mean of zero and a standard deviation of one in the individual-level world sample.

The GPS contains twelve items which are summarized in Table 1. For most preferences, the set of questions consists of a combination of qualitative items, which are more abstract, and quantitative questions, which put the respondent into precisely defined hypothetical choice scenarios.[8]

Patience. Our measure of patience is derived from the combination of responses to two survey measures, one with a quantitative and one with a qualitative format. The quantitative survey measure consists of a series of five interdependent hypothetical binary choices between immediate and delayed financial rewards, a format commonly referred to as "staircase" (or "unfolding brackets") procedure (Cornsweet, 1962). In each of the five questions, participants had to decide between receiving a payment

Table 1: Survey items of the GPS

Preference	**Item Description**	Weight
Patience	Intertemporal choice sequence using staircase method	0.71
	Self-assessment: Willingness to wait	0.29
Risk taking	Lottery choice sequence using staircase method	0.47
	Self-assessment: Willingness to take risks in general	0.53
Positive reciprocity	Self-assessment: Willingness to return a favor	0.48
	Gift in exchange for help	0.52
Negative reciprocity	Self-assessment: Willingness to take revenge	0.37
	Self-assessment: Willingness to punish unfair behavior towards self	0.265
	Self-assessment: Willingness to punish unfair behavior towards others	0.265
Altruism	Donation decision	0.54
	Self-assessment: Willingness to give to good causes	0.46
Trust	Self-assessment: People have only the best intentions	1

Notes. See Appendix A.6 for the wording of the questions and Appendix A.7.2 for a discussion of the weights.

[8] Under certain assumptions, the quantitative items allow the computation of quantitative measures such as a CRRA coefficient or an internal rate of return.

today or larger payments in 12 months:

> *Suppose you were given the choice between receiving a payment today or a payment in 12 months. We will now present to you five situations. The payment today is the same in each of these situations. The payment in 12 months is different in every situation. For each of these situations we would like to know which one you would choose. Please assume there is no inflation, i.e., future prices are the same as today's prices. Please consider the following: Would you rather receive amount x today or y in 12 months?*

The immediate payment x remained constant in all subsequent four questions, but the delayed payment y was increased or decreased depending on previous choices (see Appendix A.6.1 for an exposition of the entire sequence of binary decisions). In essence, by adjusting the delayed payment according to previous choices, the questions "zoom in" around the respondent's point of indifference between the smaller immediate and the larger delayed payment and make efficient use of limited and costly survey time. The sequence of questions has 32 possible ordered outcomes. In the international survey, monetary amounts x and y were expressed in the respective local currency, scaled relative to median household income in the given country. Notably, this measure not only resembles standard experimental procedures of eliciting time preferences, but it is also precisely defined, arguably making it less prone to culture-dependent interpretations. This makes the quantitative patience measure well-suited for a multinational study like the present one.

The qualitative measure of patience is given by the respondents' self-assessment regarding their willingness to wait on an 11-point Likert scale, asking "how willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?" As discussed above, the two items were first standardized and then combined linearly to form the final measure of patience, which was then standardized again at the individual level in the world sample. The quantitative measure obtained a weight of 71%.

Risk Taking. Risk preferences were also elicited through a series of related quantitative questions as well as one qualitative question. Just as with patience, the quantitative measure consists of a series of five binary choices between a fixed lottery and varying sure payments, hence making use of the advantages of precisely defined, quantitative survey items in culturally and economically heterogeneous samples:

> *Please imagine the following situation. You can choose between a sure payment of a particular amount of money, or a draw, where you would*

have an equal chance of getting amount x or getting nothing. We will present to you five different situations. What would you prefer: a draw with a 50 percent chance of receiving amount x, and the same 50 percent chance of receiving nothing, or the amount of y as a sure payment?

The questions are again interdependent in the sense that the choice of the lottery results in an increase of the sure amount being offered in the next question, and vice versa. Appendix A.6.2 contains an exposition of the entire sequence of survey items. The qualitative item asks for the respondents' self-assessment of their willingness to take risks on an eleven-point scale ("In general, how willing are you to take risks?"). This qualitative subjective self-assessment has previously been shown to be predictive of risk-taking behavior in the field in a representative sample (Dohmen et al., 2011) as well as of incentivized experimental risk-taking across countries in student samples (Vieider et al., 2014). The qualitative item and the outcome of the quantitative staircase measure were combined through roughly equal weights.

Positive Reciprocity. People's propensity to act in a positively reciprocal way was also measured using one qualitative item and one question with a quantitative component. First, respondents were asked to provide a self-assessment about how willing they are to return a favor on an 11-point Likert scale. Second, participants were presented a choice scenario in which they were asked to imagine that they got lost in an unfamiliar area and that a stranger – when asked for directions – offered to take them to their destination. Participants were then asked which out of six presents (worth between 5 and 30 euros in 5 euros intervals) they would give to the stranger as a "thank you". These two items receive roughly equal weights.

Negative Reciprocity. Negative reciprocity was elicited through three self-assessments. First, people were asked how willing they are to take revenge if they are treated very unjustly, even if doing so comes at a cost (0-10). The second and third item probed respondents about their willingness to punish someone for unfair behavior, either towards *themselves* or towards a *third person*.[9] This last item captures prosocial punishment and hence a concept akin to norm enforcement. These three items receive weights of about one third each.

[9]In the original validation study, the second and third item were collapsed into one question which asked people how willing they are to punish others, without specifying *who* was treated unfairly (Falk et al., 2015). However, in the cross-country pre-test, a number of respondents indicated that this lack of specificity confused them, so that we broke this survey item up into two questions. Accordingly, the weights for deriving an individual-level index of negative reciprocity are determined by dividing the OLS weight for the original item by two.

Altruism. Altruism was measured through a combination of one qualitative and one quantitative item, both of which are related to donation. The qualitative question asked people how willing they would be to give to good causes without expecting anything in return on an 11-point scale. The quantitative scenario depicted a situation in which the respondent unexpectedly received 1,000 euros and asked them to state how much of this amount they would donate. These two items were weighted about equally.

Trust. To measure trust, we used one item, which asked people whether they assume that other people only have the best intentions (Likert scale, 0-10).[10]

2.3 Further Variables of Interest

The GPS data include a wide range of individual-level background variables which can be linked to the preference measures. These background variables include the core items of the Gallup World Poll such as (i) extensive sociodemographic information (e.g., age, gender, family structure, country of birth, religious affiliation, location of residence, or migration background including country of origin), (ii) a variety of self-reported behaviors and economic outcome variables including income, educational attainment, savings, labor market decisions, health, and behavior in social interactions, and (iii) opinions and attitudes about issues such as local and global politics, local institutional quality, economic prospects, safety, or happiness. We also elicited a self-reported proxy for cognitive skills by asking people to assess themselves regarding the statement "I am good at math" on an 11-point Likert scale. Finally, the data contain regional identifiers (usually at the state or province level), hence allowing for cross-regional analyses within countries.

3 Cross-Country Analysis

The analysis begins with an investigation of the heterogeneity of preferences around the world. Figure 1 shows how the country averages for each (standardized) preference compare to the world average. The figure reveals that preferences vary substantially across countries, by at least one standard deviation for each preference (see figure notes on color coding).[11] Most country differences displayed in Figure 1 are statistically significant. Calculating t-tests of all possible (2,850) pairwise comparisons

[10] Given the existence of the World Values Survey data, we can perform a first plausibility check on our data by showing that our trust measure is correlated with the WVS data ($\rho = 0.53$, $p < 0.01$).

[11] Appendix A.8 provide an alternative way to visualize the heterogeneity, with histograms of preferences at the country and individual levels.

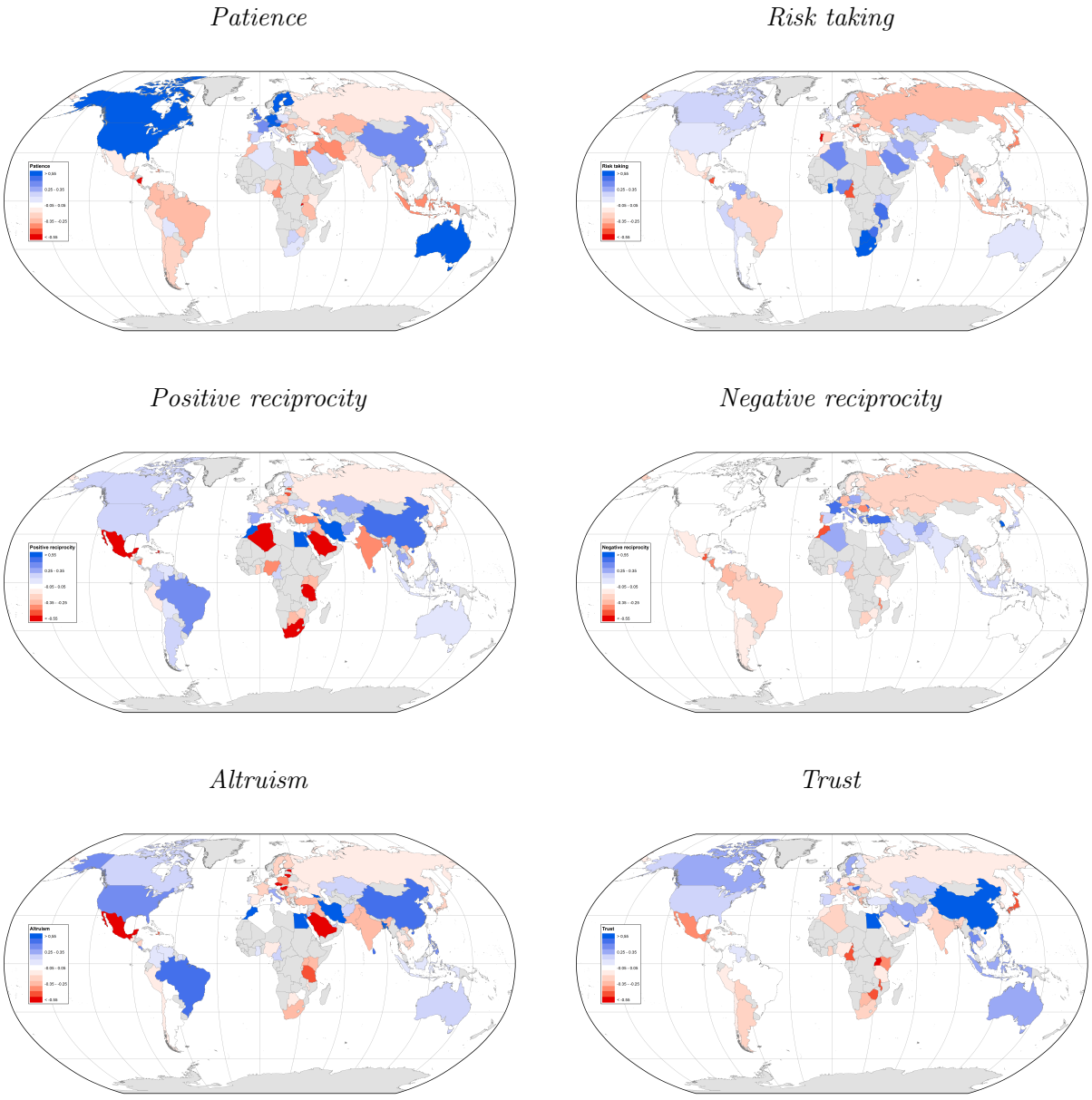

Figure 1: World maps of preferences. In each figure, white denotes the world average. Darker blue indicates higher values of a given trait, while darker red colors indicate lower values, all of which are measured in standard deviations from the world mean. Grey indicates missings.

for each preference, the fraction of significant (1-percent level) country differences are: 78% for risk, 83% for patience, 80% for altruism, 81% for positive reciprocity, 79% for negative reciprocity, and 78% for trust, respectively.

To provide a complementary perspective on the geographic and cultural variation in aggregate preferences, Figures 9a and 9b in Appendix C group countries into six world regions: Western and "Neo" Europe (i.e., the US, Canada, and Australia), Former Communist Eastern Europe, Asia, North Africa and Middle East,

Sub-Saharan Africa, and Southern America. For each region, we present two scatter plots which illustrate the distribution of patience, risk taking, negative reciprocity, and "prosociality"[12] within each region, relative to the world mean of the respective preference. Populations in Western and "Neo" Europe tend to be substantially more patient than the world mean. In fact, all of the ten most patient countries in the world are either located in Western Europe or part of the English-speaking world, with the Northern European countries exhibiting particularly high levels of patience. Western European countries are also notable for negative reciprocity; eight out of the ten most negatively reciprocal countries are located in Europe. Three of the five most negatively reciprocal countries in our sample – Turkey, Greece, and South Korea – have been found in previous research (Herrmann et al., 2008) to be particularly prone to retaliatory (anti-social) punishments in incentivized social dilemma games.

To the East, the former communist Eastern European countries are on average rather risk averse and not very patient, but the patterns are less clear compared to their Western European counterparts. Similar patterns obtain for East and South Asia, where most populations except the Confucian ones (China, Japan, South Korea) are relatively impatient.

Middle Eastern and North African populations have in common relatively high levels of risk tolerance and low levels of patience. Prosociality and negative reciprocity of this group of countries are fairly diverse. Notably, all of the ten most risk tolerant countries in our sample are located in the Middle East or Africa; in addition, *all* sub-Saharan populations are on average less prosocial than the world mean and are rather impatient.

Finally, in the Southern Americas, most populations appear impatient. They also have low levels of negative reciprocity and intermediate values in risk taking and prosociality. In sum, these results highlight that different types of preferences are spatially and culturally concentrated.

To begin to open the black box of cross-country variation in preferences, we proceed by relating preferences to country-level characteristics. In a first step, we seek to understand which fraction of the between-country variation can be explained by commonly employed variables that are plausibly exogenous to preferences. To this end, Figure 2 plots the R-squared of OLS regressions of each preference on a set of variables which proxy for geography, climate, diversity, and religion. Geography includes distance to the equator, longitude, and the fraction of the population at risk of contracting malaria; climate includes average precipitation and average temperature

[12]Given the high correlations between altruism, positive reciprocity, and trust (see below), we define prosociality as the unweighted average of these three measures. Very similar results obtain if we run a factor analysis and use the first factor of the three measures.

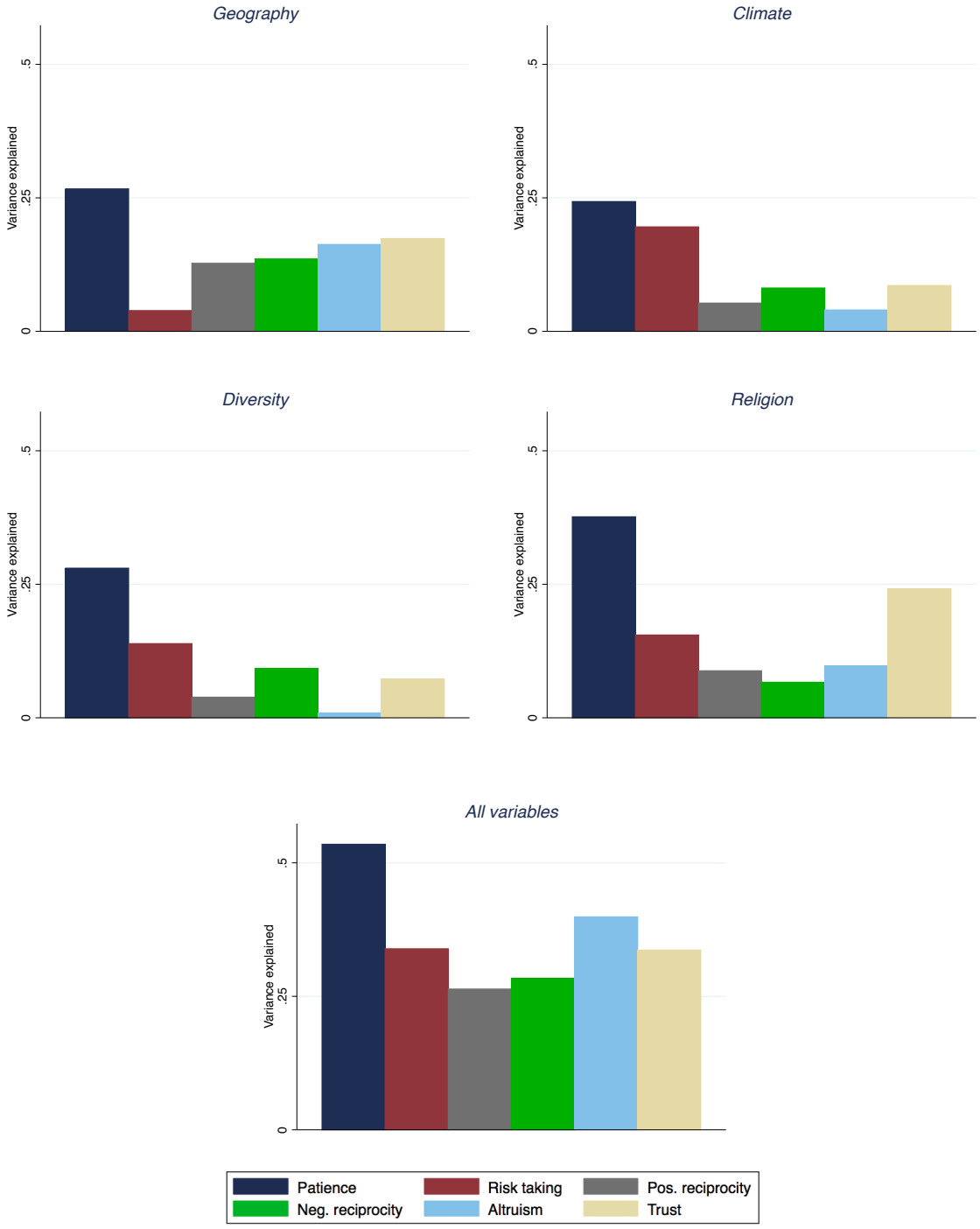

Figure 2: Cross-country variance explained. The figures plot the R^2 of an OLS regression of the respective preference on absolute latitude, longitude, and the fraction of the population at risk of contracting malaria (geography), average temperature, average precipitation and the fraction of the population living in the (sub)-tropics (climate), religious and ethnic fractionalization as well as linguistic diversity (diversity), and the share of protestants, catholics, muslims, jews, hinduists, and buddhists (religion), as well as all of these variables. All figures contain 73 countries. See Appendix D for the regression coefficients and Appendix I for details on all covariates.

as well as the fraction of the population living in the (sub)-tropics; diversity includes ethnic and religious fractionalization as well as linguistic diversity (Fearon, 2003); and religion includes the share of the population adhering to catholic, protestant, muslim, buddhist, hinduist, or jewish beliefs.

As Figure 2 shows, the variance explained by categories differs, and it differs by preference. For example, geography explains 20% of the variation in patience, but is largely unrelated to risk. Climate, by contrast, is strongly related to both preferences. The diversity indices explain more than a quarter of the varience in patience, but virtually no variation in altruism or positive reciprocity. Religion shares explain a considerable fraction of the variation in patience and trust, but little of the other social preferences. All variables in combination can account for between 26% and 53% of the between-country variation in preferences, with patience generally exhibiting the highest share of variance explained. Appendix D presents the regression coefficients underlying each figure.

In a next step, we study the correlations between preferences and country-level outcomes or characteristics which may be endogenous to preferences. Table 2 presents Pearson correlations between all preferences and a set of economic, political, and health variables.[13] Overall, the country-level preference measures vary systematically with important aggregate economic and social outcomes. In particular, we find that a population's average patience is strongly correlated with the degree of economic and institutional development, as exemplified by high income per capita, a high democracy index, and high average life expectancy at birth. Average risk taking exhibits significant correlations with a set of variables which are connected to the "riskiness" of the respective economic, social, and health environment. Risk tolerant populations are those with low life expectancy, high economic inequality, low redistribution, low labor protection, and more intentional homicides. Thus, overall, risk averse populations seem to be more protected and have stronger social safety nets, consistent with the view that the riskiness of the overall environment is influenced by, or influences, people's risk attitudes.

Regarding social preferences, positive reciprocity tends to be higher in environments with high life expectancy and in countries with low crime. Populations with strong negatively reciprocal inclinations are richer, have lower inequality, redistribute more, and exhibit lower numbers of homicides. Finally, trust is positively correlated with national income, but not with democracy. Higher inequality and a higher homicide rate are associated with lower trust. These intriguing correlations raise the question of whether preferences might shape, be shaped by, or co-evolve with

[13]Computing Spearman rank correlations yields very similar results.

Table 2: Pairwise correlations between preferences and country-level variables

	Log [GDP p/c PPP]	Democracy index	Life expectancy	Gini coefficient	Redistribution (% of GDP)	Labor regulation	Log [# of homicides]
Patience	0.630***	0.449***	0.425***	0.064	0.512***	-0.003	-0.459***
Risk taking	-0.103	-0.167	-0.394***	0.318**	-0.228*	-0.367**	0.210*
Positive reciprocity	0.078	-0.051	0.284**	-0.217	0.111	0.066	-0.231**
Negative reciprocity	0.230**	0.091	0.212*	-0.291**	0.264**	0.002	-0.400***
Altruism	-0.091	-0.227*	0.075	0.011	-0.204*	-0.058	-0.051
Trust	0.282***	-0.089	0.382***	-0.297**	0.144	0.188	-0.266**
Observations	76	74	76	50	70	56	76

Notes. Pairwise Pearson correlations between average preferences and other variables at country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

cross-country variation in development, inequality, or institutions, pointing towards an important direction for future research.

While various preferences exhibit economic and geographic patterns, preferences may also be correlated amongst each other, giving rise to country-level preference profiles. To investigate the relationship among different preferences, Table 3 shows Pearson correlations of preferences together with levels of significance.[14] The significant correlations indicate that preferences are not distributed independently of one another. One set of traits that goes together is risk tolerance and patience, as shown by the positive and statistically significant correlation at the country level. This is in spite of the special case of Sub-Saharan African countries, which tend to be risk seeking and impatient, as discussed above.[15] Another grouping of positively correlated traits involves prosociality, i.e., the traits of positive reciprocity, altruism and trust. While trust constitutes a belief rather than a preference, all of these traits share in common that they describe positive behavioral dispositions towards others. The correlation between altruism and positive reciprocity is particularly high, and trust also tends to be higher where people are positively reciprocal. This is intuitive as it is hard to imagine stable and high levels of trust in environments absent positive reciprocity, i.e., trust rewarding behaviors.[16] Despite being related to the social domain, negative reciprocity is not at all correlated with prosociality. We report the correlation structure among preferences at the individual level in Appendix B.

Evidence that preference dispositions vary substantially across countries does not imply that cross-country or cultural differences are the primary source of preference

Table 3: Pairwise correlations between preferences at country level

	Patience	Risk taking	Pos. reciprocity	Neg. reciprocity	Altruism	Trust
Patience	1					
Risk taking	0.231**	1				
Positive reciprocity	0.0202	-0.256**	1			
Negative reciprocity	0.262**	0.193*	-0.154	1		
Altruism	-0.00691	-0.0155	0.711***	-0.132	1	
Trust	0.186	-0.0613	0.363***	0.160	0.272**	1

Notes. Pairwise Pearson correlations between average preferences at country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[14]The results are similar when computing Spearman correlations.

[15]Excluding African countries, the positive correlation between risk taking and patience increases to 0.30, while other correlations remain largely the same. The correlation between the staircase risk and patience items is 0.19, while that between the two qualitative risk and patience items is 0.55.

[16]Given that our survey item for trust measures only the belief-component of trust (as opposed to first-mover behavior in trust games, which is also affected by risk preferences), the low correlation between trust and risk taking is consistent with previous within-country findings.

variation in the world. Table 4 shows results from a total variance decomposition, which reveals that the within-country variation in preferences is actually larger than the between-country variation, an observation that varies only minimally by preference. Part of the within-country variation might reflect measurement error, so that the variation in true preferences is overstated.[17] However, the available evidence on the size of test-retest correlations and measurement error suggests that it is highly unlikely that measurement error alone produces the fact that within-country variation dominates between-country variation, see Appendix H for details.

The relative importance of within-country variation does not imply that country differences are negligible or irrelevant. It does, however, suggest that individual characteristics contribute relatively more to the formation of human preferences than national borders.

Table 4: Between- vs. within-country variation

Preference	Between-country variation (%)	Within-country variation (%)
Patience	13.5	86.5
Risk taking	9.0	91.0
Positive reciprocity	12.0	88.0
Negative reciprocity	7.0	93.0
Altruism	12.3	87.7
Trust	8.2	91.8

Notes. Results from a variance decomposition in which the total individual-level variation in the respective preference is decomposed into the variance of the average preference across countries and the average of the within-country variance. Formally, the between-country variation corresponds to the R^2 of an OLS regression of all individual-level observations on a set of country dummies in which all observations are weighted by the sampling weights provided by Gallup to achieve (ex post) representativeness.

[17] In fact, comparing the between- and within-country variation across survey items reveals that the between-country variation tends to be relatively larger for the quantitative survey items. For example, in the case of patience, the quantitative staircase procedure exhibits a between-country variation of 15.7%, while the qualitative patience measure has a between-country variation of 7.3%.

4 Preferences and Individual Characteristics

4.1 Age, Gender, and Cognitive Ability

The pronounced within-country heterogeneity calls for a better understanding of the individual-level determinants of preferences. The analysis focuses on three main characteristics: age, gender and cognitive ability, taking self-reported math skills as a proxy for the latter.[18] Gender, age, and cognitive ability are interesting to study, and have received particular attention in previous research on preferences (e.g., Barsky et al., 1997; Donkers et al., 2001; Croson and Gneezy, 2009; Frederick, 2005; Sutter and Kocher, 2007; Dohmen et al., 2010, 2011; Benjamin et al., 2013), for two main reasons. First, they are associated with important differences in economic outcomes. If preferences vary with these traits, they could be part of the explanation. Second, these traits are plausibly exogenous to preferences. The previous literature has proposed various mechanisms through which gender, age, and cognitive ability might be related to preferences, ranging from biological to purely social (Croson and Gneezy, 2009; Dohmen et al., 2011; Benjamin et al., 2013). There is limited knowledge, however, about the relative importance of these different types of mechanisms. For example, alternative explanations for age effects include an influence of idiosyncratic historical and cultural environments on the one hand, to biological aspects of the aging process on the other hand. The ability to examine how different preferences vary with characteristics across countries with diverse historical experiences can shed light on such questions.

Table 5 reports regressions of each preference on age, age squared, gender and math skills in the full world sample, conditional on country fixed effects. For each preference, the second column contains additional covariates to be discussed below. The variables are standardized, so the coefficients show the change in the dependent variable (respective preference) in standard deviation units, for a one unit change in an independent variable (individual characteristics).

The estimates in table 5 reveal that, in the world population as a whole, preferences vary significantly with gender, age, and cognitive ability. Specifically, for gender, the strongest relationship is for risk preference: women are relatively more risk averse than men. Women are also significantly more prosocial, i.e., they tend to be more altruistic, positively reciprocal, and trusting, and are less negatively reciprocal. Women are slightly more impatient than men. In terms of age, the

[18] This proxy may tend to capture the numeracy aspect of cognitive skills. Subjective assessments of ability are correlated with measured cognitive ability, and have predictive power for academic achievement (Spinath et al., 2006). While such relative self-assessment might be interpreted in different ways across countries, we only use self-reported cognitive skills in within-country analyses.

Table 5: Correlates of preferences at individual level

	Dependent variable:											
	Patience		Risk taking		Pos. reciprocity		Neg. reciprocity		Altruism		Trust	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age	0.72***	0.69***	-0.083	0.46**	1.02***	0.79***	-0.36*	-0.18	-0.0060	0.024	0.37*	0.071
	(0.17)	(0.18)	(0.20)	(0.19)	(0.17)	(0.18)	(0.19)	(0.19)	(0.14)	(0.14)	(0.21)	(0.16)
Age squared	-1.45***	-1.34***	-1.20***	-1.75***	-1.17***	-0.90***	-0.45**	-0.69***	0.015	-0.13	0.032	0.24
	(0.20)	(0.21)	(0.21)	(0.19)	(0.18)	(0.20)	(0.18)	(0.20)	(0.15)	(0.15)	(0.20)	(0.17)
1 if female	-0.056***	-0.033***	-0.17***	-0.16***	0.049***	0.057***	-0.13***	-0.12***	0.10***	0.10***	0.066***	0.059***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Subj. math skills	0.028***	0.025***	0.046***	0.043***	0.038***	0.037***	0.040***	0.043***	0.044***	0.042***	0.056***	0.060***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log [Household income p/c]		0.038***		0.063***		0.042***		0.019*		0.051***		-0.013*
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)
Subj. law and order index		0.058***		0.044*		0.021		-0.066**		0.037*		0.22***
		(0.02)		(0.02)		(0.03)		(0.03)		(0.02)		(0.03)
Subj. health index		0.097***		0.12***		0.12***		-0.056**		0.10***		0.052**
		(0.02)		(0.03)		(0.04)		(0.03)		(0.02)		(0.03)
1 if christian		-0.082**		-0.12***		-0.012		-0.13***		0.13***		0.10***
		(0.03)		(0.02)		(0.03)		(0.03)		(0.03)		(0.02)
1 if muslim		-0.13**		-0.098**		-0.0000071		-0.11**		0.19***		0.18***
		(0.05)		(0.05)		(0.05)		(0.05)		(0.05)		(0.05)
1 if hinduist		-0.12*		-0.16***		0.0091		-0.18***		0.13**		0.16***
		(0.06)		(0.04)		(0.05)		(0.05)		(0.05)		(0.05)
1 if buddhist		0.10		-0.19***		0.014		-0.062		0.16**		-0.061
		(0.09)		(0.05)		(0.06)		(0.05)		(0.06)		(0.12)
1 if jew		0.40***		-0.11**		-0.024		0.20***		0.29***		0.12**
		(0.06)		(0.04)		(0.05)		(0.06)		(0.04)		(0.05)
1 if other religion		-0.00022		-0.18***		-0.028		-0.060		0.14**		0.0053
		(0.06)		(0.04)		(0.08)		(0.05)		(0.05)		(0.09)
Constant	-0.37***	-0.58***	0.21***	-0.32***	-0.079**	-0.40***	0.37***	0.41***	-0.064**	-0.67***	-0.078**	-0.29***
	(0.04)	(0.09)	(0.04)	(0.09)	(0.04)	(0.09)	(0.05)	(0.10)	(0.03)	(0.06)	(0.04)	(0.08)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	78501	58479	78445	58437	78869	58733	77521	57731	78632	58532	77814	57952
R^2	0.165	0.154	0.167	0.185	0.128	0.121	0.112	0.114	0.135	0.130	0.111	0.100

Notes. OLS estimates, standard errors (clustered at country level) in parentheses. Coefficients are in terms of units of standard deviations of the respective preference (relative to the individual world mean). For the purposes of this table, age is divided by 100. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

regression results indicate that, on average: Young individuals are relatively more willing to take risks, and punish; the middle aged are especially positively reciprocal and patient; the elderly have the strongest risk aversion and are relatively trusting. Preferences are also significantly related to self-reported cognitive ability: high cognitive ability individuals are more patient, less risk averse, more positively and negatively reciprocal, more trusting, and more altruistic.

We next exploit the ability to study the relationship of preferences to characteristics separately, for 76 different countries, to understand the extent to which the relationship of preferences to characteristics is culturally specific. Figure 3 addresses this question for gender differences. For each country, we regress each preference on age, age squared, gender, and cognitive ability. Figure 3 plots the resulting (conditional) gender coefficients. Each dot represents a country, the respective coefficient and the respective level of significance (green not significantly different from zero, pink at 10, blue at 5, and red at 1 percent level of significance, respectively). To ease reading, each panel also contains a horizontal line at zero.

Figure 3 shows that greater risk aversion among women is common to most countries. In 95 percent of countries, the gender coefficient is non-zero and in the direction of greater risk aversion among women. Of these negative coefficients, 82 percent are statistically significant at least at the 10-percent level. The gender difference in negative reciprocity is next in terms of universality: in 89 percent of countries is the female population on average less willing to reciprocate negatively than men. The impression that the majority of countries have a similar qualitative relationship between gender and preferences extends to the other social preferences, although there is more heterogeneity. For altruism, positive reciprocity, and trust about 79 percent, 71 percent, and 71 percent of countries have women being more pro-social than men, respectively. Patience is the most variable in terms of gender differences, but still, 68 percent of countries have non-zero coefficients of the same sign. These findings show that there are striking commonalities in terms of how gender and preferences are related, across a large and diverse set of cultural backgrounds. At the same time, there is substantial heterogeneity in the magnitude of the relationships across countries.

Figure 4 explores how the relationship between age and preferences varies across countries. This figure is divided by whether countries are OECD-members or not, in order to show some of the most salient cross-country differences: OECD-members exhibit a hump-shaped pattern for both positive reciprocity and patience, that is almost entirely absent in non-OECD countries. In Appendix E.1, we also provide the age profiles for the six world regions defined above, i.e., Western and "Neo"

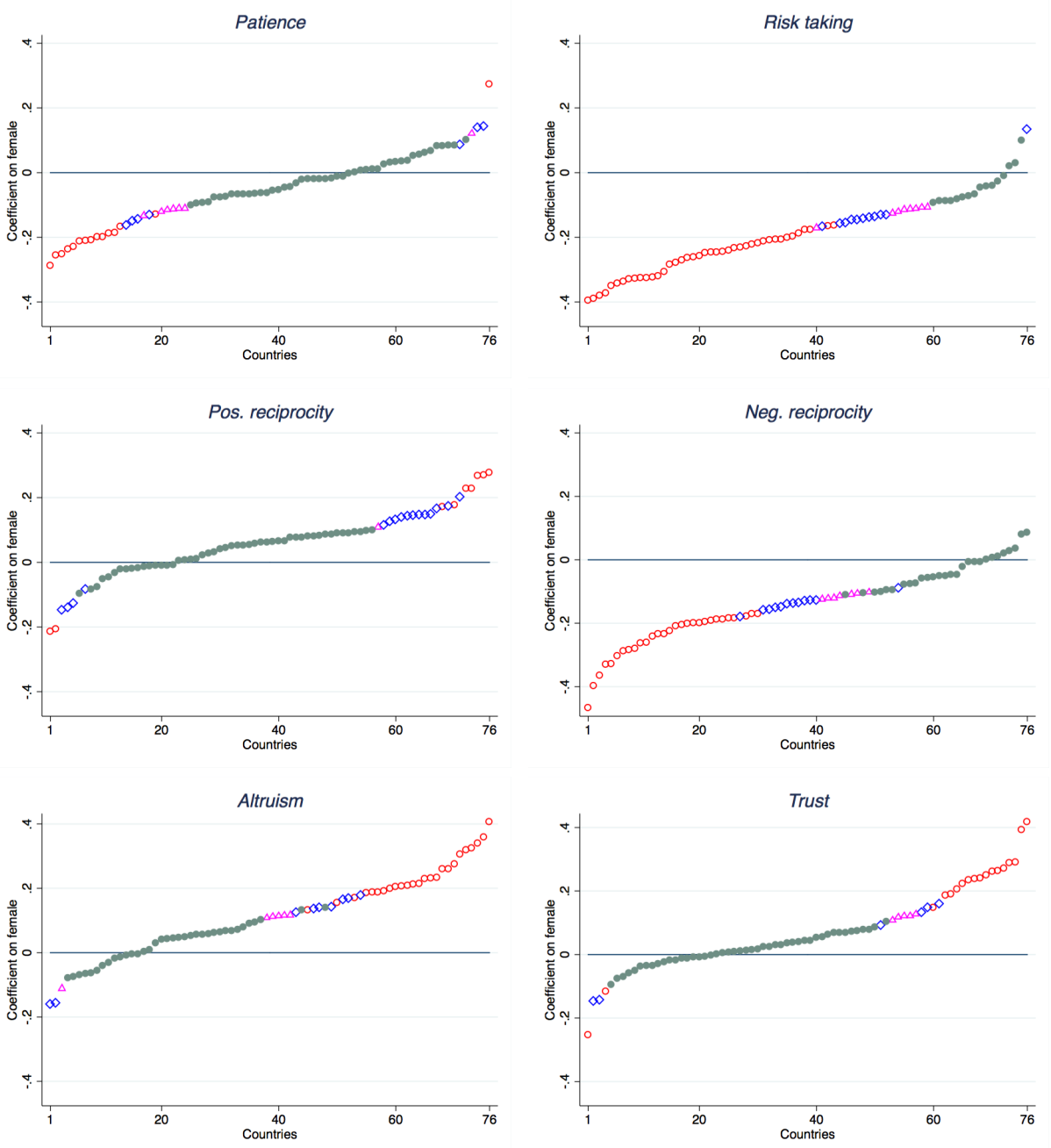

Figure 3: Gender correlations separately by country. Each panel plots the distribution of gender correlations. That is, for each country, we regress the respective preference on gender, age and its square, and subjective math skills, and plot the resulting gender coefficients as well as their significance level. In order to make countries comparable, each preference was standardized (z-scores) within each country before computing the coefficients. Green dots indicate countries in which the gender correlation is not statistically different from zero at the 10% level, while red / blue / pink dots denote countries in which the effect is significant at the 1% / 5% / 10% level, respectively. Positive coefficients imply that women have higher values in the respective preference.

Europe, Former Communist Eastern Europe, Asia, North Africa and Middle East, Sub-Saharan Africa, and Southern America. These figures show that patience and positive reciprocity have hump-shaped age profiles mainly in Western and Neo Europe, but also reveal that age profiles for altruism are variable across world regions. By contrast, the age profiles for risk taking, negative reciprocity, and trust are more universal.[19]

Intuitively, the relationship between cognitive ability and preferences might not be as strongly tied to socially constructed roles as gender, or age, and might hence have an even more universal relationship to preferences independent of cultural background. Indeed, we find that self-reported cognitive ability is related to each preference in a strikingly similar way across countries. For every preference, at least 93 percent of countries have the same qualitative relationship between the preference and cognitive ability, for all preferences. Similarly to Figure 3, Appendix E.2 shows graphs of the cognitive ability coefficients, by preference and country.

In sum, the findings reveal that there are at least some mechanisms linking preferences to gender, age, and cognitive ability that are universal features of humans, possibly due to biological or psychological mechanisms. At the same time, the results also show a significant amount of variation in the quantitative magnitude and sometimes even the direction of relationships with individual characteristics. Thus, the results provide an important caveat to studies on more specific samples, which sometimes produce highly variable or even contradictory results.

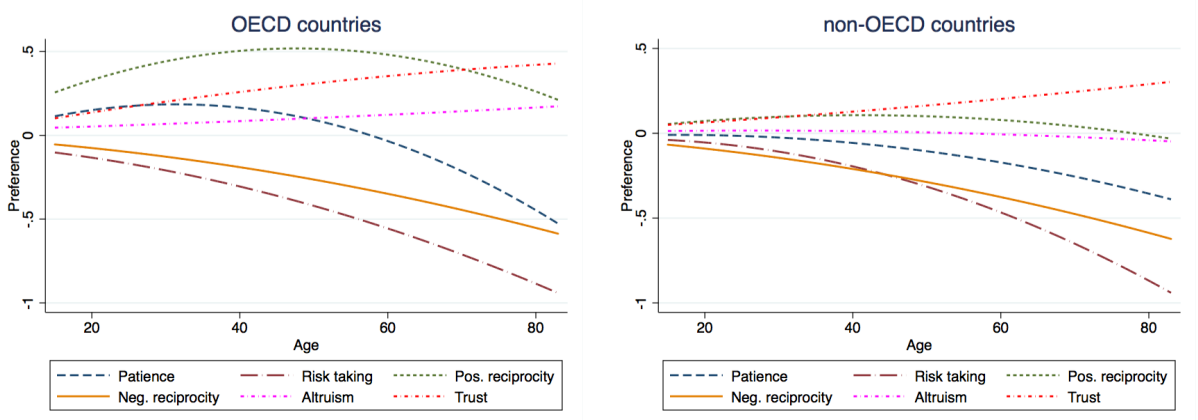

Figure 4: Age profiles by OECD membership. The figures depict the relationship between preferences and age conditional on country fixed effects, gender, and subjective math skills. These are augmented component plus residuals plots, in which the vertical axis represents the component of the preference that is predicted by age and its square plus the residuals from the regression in the second column of Table 5. The horizontal axis represents age, winsorized at 83 (99th percentile).

[19]Similar results on the (partially non-linear) age profiles and their relationship to the degree of development obtain when we depict the age profile for all 76 countries separately; these figures are available upon request.

4.2 Further Correlates of Preferences

The analysis proceeds by investigating the relationship between preferences on the one hand and income, religious affiliation, physical health, and subjective safety perceptions, on the other hand. While these characteristics and attitudes are not as exogenous to preferences as age, gender, and cognitive ability, they might nevertheless plausibly affect the formation of preferences.

Health and safety perceptions are evaluated using two indices that Gallup provides by aggregating several survey items. For example, the subjective law and order index includes a question on whether respondents recently had money or property stolen. The subjective health index makes use of an item asking people whether they have health problems that are atypical given their age, among others (see Appendix I for details).

The results are reported in columns (2), (4), (6), (8), (10), and (12) of Table 5. The table shows that household income is positively correlated with all preferences, but not with trust. Subjective perceptions of law and order are positively correlated with patience and trust, and negatively associated to negative reciprocity. One possible interpretation is that if people feel safe, they trust others more and are more willing to postpone financial rewards. Negative reciprocity might be less prevalent if an individual believes institutions already provide mechanisms for strong formal sanctions. Previous work finds that retaliatory punishment is more pronounced in countries where people perceive the rule of law as weak (Herrmann et al., 2008).

Subjective physical health perceptions are positively correlated with all traits, except for negative reciprocity. Turning to religious affiliation, the regressions investigate the correlation between six religion dummies (christian, muslim, hinduist, buddhist, jew, other), taking seculars as baseline category. Again note that the religious affiliation coefficients are entirely identified from within-country variation in religious denomination and preferences. We find a very consistent pattern: relative to seculars, christians, muslims, and hinduists are less patient, more risk averse, less negatively reciprocal as well as more altristic and trusting. Similar patterns hold for buddhists. Jews, on the other hand, are more patient, altruistic, and trusting, and have strong negatively reciprocal inclinations. Overall, these results suggest that, relative to atheists, being religious of any denomination tends to be associated with relatively high risk aversion, high values on the prosocial dimensions altruism and trust, and low negative reciprocity (albeit with a few notable exceptions).

5 Preferences and Individual Behaviors

We now turn to investigating the relationships of preferences to individual behaviors and outcomes. Understanding the relationship between our preference measures and individual-level economic and social decisions is important in two respects. First, such analyses provide insights into the role of heterogeneity in underlying preference parameters for generating observed choice behavior, on a global scale. Second, it allows us to evaluate the meaningfulness and behavioral relevance of our items in a culturally and economically highly heterogeneous sample.[20]

5.1 Accumulation Decisions

We evaluate the explanatory power of the GPS patience measure by relating it to the accumulation of physical and human capital. Table 6 presents estimates of OLS regressions of different outcomes on patience. Columns (1) and (2) display the results of a linear probability model, in which we employ as dependent variable a binary indicator for whether the respondent saved in the previous year. Patience is correlated with savings behavior both with and without country fixed effects, and conditional on socioeconomic covariates such as age, gender, income, cognitive ability, and religion. The point estimate implies that a one standard deviation increase in patience is associated with a roughly 20% increase of the probability of saving relative to the baseline probability of 26.7%. Columns (3) and (4) establish that patience is also significantly related to educational attainment; these estimates are based on a three-step categorical variable (roughly: primary, seondary, and tertiary education).[21]

In Appendix F.1, we show that the significant relationship between our patience variable and accumulation processes is not driven by only a few countries. Specifically, by plotting the distribution of point estimates and their significance level across countries, we show that the coefficient of patience is positive in the more than 90% of countries for both savings and education, and mostly statistically significant. For instance, the correlation between patience and education is in most cases statistically significant at least at the 5% level in 74% of all countries.

[20] Throughout this section, the respective dependent variables are sometimes only available for a subset of countries because the respective question was not part of Gallup's core questionnaire.

[21] Appendix F.2 presents robustness checks on all results in this section using (ordered) probit estimations.

Table 6: Patience and accumulation decisions, risk preferences and risky choices

	Dependent variable:									
	Accumulation decisions				Risky choices					
	Saved last year		Education level		Own business		Plan to start business		Smoking intensity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Patience	0.050***	0.025***	0.12***	0.033***						
	(0.01)	(0.01)	(0.02)	(0.00)						
Risk taking					0.031***	0.022***	0.033***	0.017***	0.050***	0.023*
					(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Age		0.025		0.99***		1.55***		0.56***		2.55***
		(0.23)		(0.26)		(0.12)		(0.09)		(0.31)
Age squared		-0.21		-1.83***		-1.54***		-0.69***		-2.86***
		(0.24)		(0.25)		(0.12)		(0.10)		(0.31)
1 if female		-0.0010		-0.016		-0.053***		-0.018***		-0.58***
		(0.01)		(0.01)		(0.01)		(0.00)		(0.03)
Subj. math skills		0.012***		0.041***		0.0056***		0.0030***		-0.011***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Log [Household income p/c]		0.11***		0.14***		0.020***		-0.0072**		-0.010
		(0.01)		(0.01)		(0.00)		(0.00)		(0.01)
Constant	0.27***	-0.37***	1.87***	0.24***	0.14***	-0.38***	0.11***	0.038	0.38***	0.39***
	(0.03)	(0.09)	(0.03)	(0.07)	(0.01)	(0.05)	(0.01)	(0.03)	(0.03)	(0.07)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Religion FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	15260	14459	79357	69272	72839	62985	57072	51489	15309	14490
R^2	0.011	0.132	0.030	0.329	0.008	0.104	0.011	0.120	0.005	0.198

OLS estimates, standard errors (clustered at country level) in parentheses. For the purposes of this table, age is divided by 100. Saved last year is a binary indicator, while education level is measured in three categories (roughly elementary, secondary, and tertiary education, see Appendix I). Self-employment and planned self-employment are binary, while smoking intensity is measured in three categories (never, occasionally, frequently). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.2 Risky Choices

To investigate whether risk preferences are related to important risky decisions in life, we build on previous within-country findings, which have found a relationship of risk attitudes to self-employment and health behavior (Dohmen et al., 2011). As columns (5) and (6) of Table 6 establish, our preference measure predicts actual self-employment both across and within countries. The same pattern holds when considering individuals' intention to start their own business, conditional on not being self-employed (columns (7)-(8)).

Columns (9) and (10) relate risk preferences to the respondent's smoking intensity, measured on a three-point scale (never, occasionally, and frequently). We find that more risk-tolerant people are more likely to smoke, both with and without country fixed effects, and conditional on a large set of covariates. Appendix F.1 shows that the correlations between risk preferences and labor market or health decisions are not restricted to a particular set of countries. Rather, risk preferences are related to risky behaviors in a qualitatively similar way around the world, although quantitative magnitudes of the relationships do vary. For example risk taking is significantly positively related to planned self-employment at least at the 10% level in about 90% of countries in the sample.

5.3 Social Interactions

We analyze the relationships of the social preference measures to behaviors and outcomes in the social domain.[22] Table 7 summarizes the results. Columns (1)-(8) show that altruism is significantly related to a broad range of prosocial behaviors including donating, volunteering time, helping strangers, or sending money or goods to other people in need. Across the different behavioral categories, the point estimate is very consistent and implies that an increase in altruism by one standard deviation is correlated with an increase in the probability of engaging in prosocial activities of 3.5–6.5 percentage points, which corresponds to an increase of roughly 15–20% compared to the respective baseline probabilities.[23] Positive reciprocity is a significant correlate of helping people in need (columns (5) through (8)), perhaps a manifestation of generalized reciprocity in the sense that reciprocal people who have been helped before are also willing to help others. In contrast, the negative reciprocity variable is virtually uncorrelated with all of the prosocial activities in the first eight columns.

[22] Since trust constitutes a belief rather than a preference, we do not incorporate it in the discussion here. However, all results are robust to controlling for trust.

[23] These baseline probabilities are 31.8%, 21.6%, 48.3%, and 23.7%, repectively (see Table 7 for the order of variables).

Table 7: Social preferences and social interactions

	Dependent variable:													
	Donated money		Volunteered time		Helped stranger		Sent money / goods to other individual		Voiced opinion to official		Have friends / relatives I can count on		In a relationship	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Altruism	0.064***	0.061***	0.046***	0.038***	0.064***	0.052***	0.035***	0.033***	0.028***	0.025***	0.0079	0.016***	0.0048	0.0013
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Positive reciprocity	0.0055	-0.00037	-0.000074	0.0049	0.035***	0.033***	0.018***	0.017***	-0.0021	-0.0025	0.015*	0.017***	0.022***	0.0072***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Negative reciprocity	0.0040	-0.0042	-0.0033	-0.0035	-0.0026	-0.0032	0.011**	0.0051	0.016***	0.016***	0.012***	0.0037	-0.0034	-0.00074
	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age		0.62***		0.42***		0.73***		0.26***		1.02***		-0.90***		5.58***
		(0.08)		(0.08)		(0.07)		(0.07)		(0.08)		(0.09)		(0.16)
Age squared		-0.47***		-0.47***		-0.90***		-0.29***		-1.00***		0.76***		-5.42***
		(0.09)		(0.08)		(0.07)		(0.08)		(0.09)		(0.09)		(0.17)
1 if female		0.013**		-0.017***		-0.014**		-0.00057		-0.045***		0.013***		-0.024***
		(0.01)		(0.01)		(0.01)		(0.00)		(0.01)		(0.00)		(0.01)
Subj. math skills		0.0091***		0.0074***		0.0091***		0.0070***		0.0091***		0.0053***		0.0024**
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Log [Household income p/c]		0.031***		0.0042		0.021***		0.034***		0.012***		0.036***		-0.036***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Constant	0.32***	-0.022	0.22***	0.15***	0.49***	0.37***	0.24***	-0.17***	0.22***	-0.084**	0.82***	0.48***	0.58***	-0.32***
	(0.02)	(0.04)	(0.01)	(0.05)	(0.02)	(0.06)	(0.02)	(0.04)	(0.01)	(0.03)	(0.02)	(0.04)	(0.01)	(0.05)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Religion FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	58229	53439	58213	53430	55991	53226	56253	53559	55944	53174	65986	59209	77881	68176
R^2	0.020	0.192	0.012	0.089	0.028	0.093	0.012	0.124	0.006	0.062	0.003	0.118	0.002	0.218

OLS estimates, standard errors (clustered at country level) in parentheses. For the purposes of this table, age is divided by 100. See Appendix I for details on all dependent variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

As columns (9) and (10) show, however, negative reciprocity is a significant predictor of whether people are willing to voice their opinion to a public official.

Columns (11) through (14) examine the relationship between social preferences and respondents' family and friendship status. We find that more altruistic and more positively reciprocal people are more likely to have friends they can count on when in need, and that positive reciprocity correlates with being in a relationship.[24]

The overall pattern in Table 7 highlights that preferences are predictive of a wide range of behaviors in the social domain and that the preference measures are of sufficiently high quality to discriminate between different types of social behavior. As Appendix F.1 shows, these relationships are not restricted to a small set of countries, but instead hold for most countries separately. For instance, the correlation between altruism and donating is statistically significant at the 5% level in 80% of all countries.

In sum, all of the GPS preference measures are significantly related to a broad range of economic and social behaviors in a way one would intuitively expect. This indicates that preference heterogeneity is important for understanding variation in economic outcomes worldwide. In addition, the fact that the correlations are qualitatively similar across cultural backgrounds and development levels provides reassuring evidence that the GPS survey items do indeed capture the relevant underlying preferences even in a very heterogeneous sample in terms of cultural background and economic development. In this sense, the correlations provide an important out-of-context validation check for the survey module.

6 The Cultural Origins of Preference Variation

The previous analyses have documented a large amount of preference heterogeneity both across and within countries. An immediate question is whether this variation has deep cultural origins. While heterogeneity in different preferences may have myriad of potentially different historical or cultural sources, we conclude the paper by drilling deeper into the potential cultural origins of preference differences. We focus on one particular proxy for cultural variation, i.e., linguistic features. Language has been used as proxy for cultural variation in many previous studies (e.g., Fearon, 2003; Desmet et al., 2009, 2012; Spolaore and Wacziarg, 2015). In deriving specific testable hypotheses on the relationship between preferences and linguistic or cultural variation, we follow the work of Chen (2013) and investigate the relationship between economic preferences and a structural feature of languages called future-time reference (FTR). As discussed in detail by Chen (2013), languages differ in whether

[24] Also see Dohmen et al. (2009).

or not they require people to grammatically mark future events: some languages require people to explicitly distinguish between present and future by making use of constructions such as "I *will* go to school tomorrow" (strong FTR), while others allow speakers to talk about the future in present tense (weak FTR). Chen (2013) argues that strong FTR languages may make the future feel more distant, potentially resulting in less future-oriented behavior. In empirically testing this proposition, he develops a binary FTR classification of languages and shows that – both across and within countries – people who speak weak FTR languages save more, are less likely to smoke, and have better health.[25]

Building on this insight, we investigate the relationship between preferences and future-time reference. Our analysis serves two purposes. First, our patience measure allows for a direct replication and extension of Chen's results on future-oriented behavior. In particular, our patience measure arguably constitutes a more fundamental and direct proxy for how people trade off current and future rewards than, e.g., medical obesity. Second, our data allow for a systematic investigation of whether the cultural trait captured by FTR is also related to other preferences rather than just time preference. In particular, it is conceivable that people for whom the future seems less distant are more likely to invest resources today to reap social benefits in the future. Thus, traits that are related to cooperation, repeated interaction, and reputation building should be more pronounced in weak FTR languages. Our data on positive reciprocity, trust, and altruism provide natural candidates for such an investigation. Our analysis proceeds in two steps. First, we investigate the relationship between average preferences and FTR at the country level. Then, we exploit within-country variation in preferences and FTR.

6.1 Cross-Country Evidence

To study the relationship between preferences and FTR, we employ Gallup's interview language as a proxy for the language respondents speak in their daily lives.[26] We apply Chen's classification of languages to our dataset, which results in a set of 55 coded languages. In addition, we were able to code an additional four languages ourselves using the methodology outlined in Chen (2013).[27] In sum, we have access

[25] Sutter et al. (2014) show that the same relationship exists for children in a bilingual city in Italy.

[26] Correspondence with Gallup suggests that, naturally, in some countries interview language is an imperfect proxy for the language people are most familiar with. Thus, proxying people's daily language with their interview language results in measurement error and hence attenuation bias, which works against finding statistically significant effects in our analyses.

[27] These languages are: Fulfulde (weak FTR), Khmer (strong FTR), Moroccan Arabic (weak FTR), and Dari (strong FTR). In addition, we changed one of Chen's classifications after corre-

to 59 classified languages for a total of 75,224 respondents.[28] All results are robust to only making use of the languages coded by Chen.

After classifying each respondent, we compute the country-level fraction of people whose language corresponds to weak as opposed to strong FTR. Then, we regress average preferences in a given country on this fraction. To take into account that we can classify only a subset of respondents in some countries (making the fraction speaking weak FTR languages a less precise estimate of the true population counterpart), our regressions weigh all observations by the fraction of people whose language can be classified. Thus, countries in which we can classify a larger fraction of respondents receive higher weight, as should be the case from a measurement error perspective.[29]

Table 8 presents the results. For each preference, we report two specifications, one without covariates and one with control variables commonly employed in cross-country regressions, i.e., continent fixed effects, (log) per capita income, distance to the equator, longitude, the fraction of the population that is at risk of contracting malaria, and average precipitation. Results show that, across countries, weak FTR is significantly correlated with average patience (columns (1)-(2)). As columns (5)-(6) and (11)-(12) show, similar patterns obtain for positive reciprocity and trust. In contrast, altruism, risk taking, and negative reciprocity are not significantly correlated with the fraction speaking weak FTR languages.

6.2 Within-Country Evidence

In a second step of the analysis, we exploit within-country variation in preferences and FTR. Such analyses are arguably better suited to identify cultural origins of preferences because they can account for unobserved heterogeneity at the country level.

In many countries in our sample, we observe some variation in interview languages. However, variation in language does not necessarily mean variation in FTR. In fact, only in Estonia, Nigeria, and Switzerland (2,925 respondents in total) do interview languages vary across respondents such that we observe within-country variation in FTR. Thus, we proceed by regressing individual-level preferences on a dummy for whether a respondent speaks a weak or strong FTR language, conditional

sponding with him. He classified Persian as strong FTR, while it is in fact weak FTR. None of our results depend on how we code Persian.

[28] We could not classify 23 languages, which are mostly spoken by small minorities (5,113 respondents in total).

[29] Appendix G confirms that virtually identical results are obtained when running unweighted OLS regressions, suggesting that measurement error in the fraction speaking weak FTR languages is weak.

Table 8: Preferences and FTR: Cross-country results

	Dependent variable:											
	Patience		Risk taking		Pos. reciprocity		Neg. reciprocity		Altruism		Trust	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fraction of population speaking weak FTR	0.37***	0.24***	-0.14*	-0.0080	0.13*	0.16**	-0.024	-0.088	0.043	0.099	0.17**	0.18**
	(0.13)	(0.09)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.09)	(0.09)	(0.10)	(0.08)	(0.08)
Log [GDP p/c PPP]		0.15***		0.032		-0.072*		0.058		-0.078*		-0.00072
		(0.04)		(0.03)		(0.04)		(0.04)		(0.04)		(0.03)
Distance to equator		0.010*		0.0017		-0.0069		-0.0081		-0.0022		-0.0072
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.00)
Longitude		-0.0019		0.0023		0.0022		0.00091		0.0025		-0.000039
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
% at risk of malaria		0.25		-0.14		-0.27		-0.13		-0.71**		-0.16
		(0.21)		(0.23)		(0.29)		(0.17)		(0.27)		(0.19)
Average precipitation		-0.00024		-0.00092		0.00065		-0.0011		0.0031**		-0.0011
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Constant	-0.055	-1.42***	0.034	-0.041	-0.047	1.39***	0.020	-0.13	-0.049	1.27**	-0.047	0.56
	(0.04)	(0.51)	(0.04)	(0.36)	(0.05)	(0.43)	(0.04)	(0.50)	(0.05)	(0.50)	(0.04)	(0.38)
Continent FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	75	74	75	74	75	74	75	74	75	74	75	74
R^2	0.146	0.636	0.031	0.442	0.022	0.253	0.001	0.271	0.002	0.334	0.053	0.408

WLS estimates, robust standard errors in parentheses. All observations are weighted by the fraction of respondents whose language can be classified as weak or strong FTR. The regressions exclude Haiti for which no respondent could be classified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

on country fixed effects and age, age squared, gender, and our cognitive skills proxy. Columns (1), (3), (5), (7), (9), and (11) of Table 9 present the results. Consistent with the cross-country evidence, we find that individuals speaking weak FTR languages are more patient, more positively reciprocal, and more trusting. In addition, these people are also significantly more altruistic.[30] We do not find significant relationships between FTR and risk taking or negative reciprocity. For each preference, a second column adds further controls, i.e., regional (state or province) fixed effects, religion fixed effects, household income, health, and subjective institutional quality. Despite this comprehensive set of covariates, and only exploiting within-region variation in FTR and preferences, we obtain almost identical results.

In sum, the results at the subnational level closely mirror those obtained in cross-country analyses, the one exception being altruism.[31] Thus, across levels of aggregation, weak FTR is predictive of higher patience and higher levels of the prosocial traits positive reciprocity, altruism, and trust. These results are supportive of the hypothesis that cultural variation might coevolve with preferences. In particular, just as in the case of the work by Chen (2013), our results lend themselves to two interesting interpretations. First, speaking a weak FTR language may actually *cause* patience and cooperation-enhancing prosociality; second, the historical evolution of linguistic features and the formation preferences may both be a product of some other very deep cultural trait. Regardless of the precise interpretation adopted, our results highlight that the contemporary preference variation may have very deep historical roots,[32] and that the GPS data are well-suited to identify such effects.

7 Discussion, Applications, and Outlook

Many theories about human behavior in economics and other fields assume that a fundamental set of preferences drives decision-making of individual agents. These include preferences about risk, timing of rewards, and in the social domain, reciprocity, altruism, and trust. Despite their importance, empirical evidence on the extent and nature of preference heterogeneity has been restricted to varying measures available

[30] When we restrict the sample to those countries with within-country variation in FTR and regress the respective preference only on the FTR indicator as well as country fixed effects, the resulting coefficient is always positive and statistically significat at the 10% level for patience and at the 1% level for positive reciprocity, trust, and altruism. In Appendix G, we report the coefficient on FTR separately for each country in which we observe within-country variation.

[31] Note that the correspondence between within- and across-country results is in no way mechanical, i.e., it need not necessarily be the case that individual- and country-level correlations are aligned.

[32] As discussed by Chen (2013), variation in future-time reference is at least several hundred years old.

Table 9: Preferences and FTR: Individual-level results

	Dependent variable:											
	Patience		Risk taking		Pos. reciprocity		Neg. reciprocity		Altruism		Trust	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1 if weak FTR	0.095**	0.053***	0.067	0.0079	0.18***	0.11***	-0.073*	-0.0037	0.24***	0.19**	0.33***	0.32**
	(0.05)	(0.02)	(0.04)	(0.07)	(0.04)	(0.04)	(0.04)	(0.09)	(0.04)	(0.09)	(0.04)	(0.13)
Age	0.76***	0.80***	-0.098	0.49**	1.07***	0.92***	-0.39***	-0.25	0.032	0.074	0.42***	0.045
	(0.10)	(0.18)	(0.10)	(0.21)	(0.10)	(0.17)	(0.10)	(0.20)	(0.10)	(0.16)	(0.10)	(0.16)
Age squared	-1.51***	-1.48***	-1.21***	-1.80***	-1.22***	-1.09***	-0.44***	-0.63***	-0.017	-0.21	-0.0044	0.27
	(0.11)	(0.21)	(0.11)	(0.21)	(0.11)	(0.18)	(0.11)	(0.20)	(0.11)	(0.17)	(0.11)	(0.16)
1 if female	-0.061***	-0.040***	-0.18***	-0.17***	0.045***	0.054***	-0.13***	-0.13***	0.100***	0.093***	0.069***	0.053***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Subj. math skills	0.028***	0.023***	0.045***	0.040***	0.039***	0.039***	0.039***	0.036***	0.044***	0.040***	0.056***	0.056***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log [Household income p/c]		0.037***		0.052***		0.049***		0.012		0.046***		-0.0083
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)
Subj. health index		0.10***		0.14***		0.062***		-0.028		0.084***		0.053***
		(0.02)		(0.03)		(0.02)		(0.03)		(0.02)		(0.02)
Subj. law and order index		0.065***		0.054**		0.0041		-0.075***		0.022		0.19***
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)
Constant	-0.49***	-0.88***	0.15***	-0.79***	-0.13**	-0.51***	0.60***	0.16	-0.17***	-0.40***	-0.46***	-0.74***
	(0.06)	(0.08)	(0.06)	(0.10)	(0.06)	(0.09)	(0.07)	(0.13)	(0.06)	(0.13)	(0.06)	(0.16)
Country FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Region FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Religion FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	73460	52628	73414	52610	73811	52862	72501	52003	73580	52675	72811	52159
R^2	0.166	0.215	0.172	0.254	0.127	0.230	0.117	0.200	0.137	0.199	0.113	0.167

OLS estimates, robust standard errors (clustered at country level) in parentheses. For the purposes of this table, age is divided by 100. See Table 5 for the religion categories. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

for a limited set of countries, and typically non-representative samples. This paper has presented the first assessment of the distribution and nature of these fundamental traits on a globally representative basis using a novel dataset, which includes behaviorally validated survey measures of preferences. The findings in this paper are clearly only a first step towards tapping the potential of the GPS. The cross-cultural dimension of the data and the representative sampling design allow entirely new perspectives and levels of analysis. We illustrate this by discussing three broad directions for future research: the mechanisms underlying the relationship between preferences and individual characteristics, the deeper causes of cross-country variation in preferences, and the potential consequences of certain country-level preference profiles.

First, the data vastly expand the amount of information available for understanding the relationship between individual-level characteristics and preferences. The precise ways in which the strength and direction of preference differences vary across different environments, locations, and institutions, may shed further light on the mechanisms underlying preference differences. For example, if gender differences in preferences are correlated within countries, this would suggest some deep mechanisms that extend across preference domains.

Second, there is much more that can be done to investigate the ultimate origins of the cross-country variation in preference. While our analysis of the relationship between FTR and preferences has provided a first step in this direction, other cultural proxies or historical events might likewise generate differences in preferences. For example, other linguistic structures, such as politeness in pronoun usage (Helmbrecht, 2003), or the role of gender in the language might be related to preference differences (Corbett, 1991). The correlation structure of preferences may also be informative for understanding the ultimate sources of preference differences. Traits may also coevolve, to the extent that they are complementary in contributing to evolutionary fitness. In this regard, it is suggestive that the groupings of positively correlated preferences that we find are plausibly complementary, in the context of theories about the human ability to sustain cooperation (e.g., high patience and strong negative reciprocity).

A third direction for future research is a detailed investigation of the link between aggregate outcomes and preferences at the country level. Given the previous lack of representative preference data, this is completely uncharted territory. Exploring in detail the many ramifications of preference differences for explaining outcomes is beyond the scope of this paper. However, to illustrate the potential power of the GPS data in understanding cross-country variation in the economic and social domain,

we conclude with two examples. First, a tendency to retaliate could exacerbate conflicts, so that negative reciprocity may be relevant for explaining the occurrence of conflict or war. Second, a large body of dynamic theories of comparative development and growth highlight the crucial role of time preference for aggregate accumulation processes. Consistent with such theories, in a follow-up paper, Dohmen et al. (2015), we find that patience is not only predictive of future-oriented decisions and income at the individual level, but also across regions within countries, and even across entire populations: more patient countries have higher savings rates, invest more into education as well as into the stock of ideas and knowledge, grow faster, and are wealthier.

References

Alesina, Alberto and Paola Giuliano, "Culture and Institutions," *Journal of Economic Literature*, forthcoming.

_ , **Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg**, "Fractionalization," *Journal of Economic Growth*, 2003, *8* (2), 155–194.

_ , **Paola Giuliano, and Nathan Nunn**, "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, 2013, *128* (2), 469–530.

_ , **Yann Algan, Pierre Cahuc, and Paola Giuliano**, "Family Values and the Regulation of Labor," *Journal of the European Economic Association*, forthcoming.

Algan, Yann and Pierre Cahuc, "Inherited Trust and Growth," *American Economic Review*, 2010, pp. 2060–2092.

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

Barro, Robert, "Religion Adherence Data," http://scholar.harvard.edu/barro/publications/religion-adherence-data 2003.

Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics*, 1997, *112*, 537–579.

Beauchamp, Jonathan, David Cesarini, and Magnus Johannesson, "The Psychometric Properties of Measures of Economic Risk Preferences," *Working Paper*, 2011.

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro, "Who is' Behavioral'? Cognitive Ability and Anomalous Preferences," *Journal of the European Economic Association*, 2013, *11*, 1231–1255.

Bonin, Holger, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde, "Cross-Sectional Earnings Risk and Occupational Sorting: The Role of Risk Attitudes," *Labour Economics*, 2007, *14* (6), 926–937.

Borghans, Lex and Bart H.H. Golsteyn, "Time Discounting and the Body Mass Index: Evidence from the Netherlands," *Economics & Human Biology*, 2006, *4* (1), 39–61.

Botero, Juan, Simeon Djankov, Rafael LaPorta, Florencio Lopez de Silanes, and Andrei Shleifer, "The Regulation of Labor," *Quarterly Journal of Economics*, 2004, *119* (4), 1339–1382.

Carpenter, Jeffrey and Erika Seki, "Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay," *Economic Inquiry*, 2011, *49* (2), 612–630.

Chabris, Christopher F., David Laibson, Carrie L. Morris, Jonathan P. Schuldt, and Dmitry Taubinsky, "Individual Laboratory-Measured Discount Rates Predict Field Behavior," *Journal of Risk and Uncertainty*, 2008, *37* (2-3), 237–269.

Chen, M. Keith, "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets," *American Economic Review*, 2013, *103* (2), 690–731.

Corbett, Greville G., *Gender*, Cambridge University Press, 1991.

Cornsweet, Tom N., "The Staircase-Method in Psychophysics," *American Journal of Psychology*, 1962, *75* (3), 485–491.

Croson, Rachel and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature*, 2009, *47*, 448–474.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Romain Wacziarg, "The Political Economy of Linguistic Cleavages," *Journal of development Economics*, 2012, *97* (2), 322–338.

_ , Shlomo Weber, and Ignacio Ortuño-Ortín, "Linguistic Diversity and Redistribution," *Journal of the European Economic Association*, 2009, *7* (6), 1291–1318.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Representative Trust and Reciprocity: Prevalence and Determinants," *Economic Inquiry*, 2008, *46* (1), 84–90.

_ , _ , _ , and _ , "Homo Reciprocans: Survey Evidence on Behavioural Outcomes," *Economic Journal*, 2009, *119* (536), 592–612.

_ , _ , _ , and _ , "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 2010, *100* (3), 1238–1260.

_ , _ , _ , _ , **Jürgen Schupp, and Gert G. Wagner**, "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," *Journal of the European Economic Association*, 2011, *9* (3), 522–550.

_ , **Benjamin Enke, Armin Falk, David Huffman, and Uwe Sunde**, "Patience and the Wealth of Nations," *Working Paper*, 2015.

Donkers, Bas, Bertrand Melenberg, and Arthur Van Soest, "Estimating Risk Attitudes Using Lotteries: A Large Sample Approach," *Journal of Risk and Uncertainty*, 2001, *22* (2), 165–195.

Eckel, Catherine, Cathleen Johnson, and Claude Montmarquette, "Savings Decisions of the Working Poor: Short- and Long-Term Horizons," in Jeff Carpenter, Glenn Harrison, and John List, eds., *Research in Experimental Economics. Vol. 10: Field Experiments in Economics*, Amsterdam: Elsevier, 2005.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde, "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," *Working Paper*, 2015.

Fearon, James D, "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, 2003, *8* (2), 195–222.

Fernández, Raquel, "Does Culture Matter?," in Jesse Benhabib, Matthew O. Jackson, and Alberto Bisin, eds., *Handbook of Social Economics*, North Holland, 2011.

_ **and Alessandra Fogli**, "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 146–177.

Frederick, Shane, "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 2005, *19* (4), 25–42.

Gallup, John L., Jeffrey D. Sachs et al., "The Economic Burden of Malaria," *American Journal of Tropical Medicine and Hygiene*, 2000, *64* (1-2 Suppl), 85–96.

Galor, Oded and Ömer Özak, "The Agricultural Origins of Time Preference," *Working Paper*, 2014.

Giuliano, Paola, "Living Arrangements in Western Europe: Does Cultural Origin Matter?," *Journal of the European Economic Association*, 2007, *5* (5), 927–952.

_ **and Nathan Nunn**, "The Transmission of Democracy: From the Village to the Nation-State," *American Economic Review: Papers & Proceedings*, 2013, *103* (3), 86.

Golsteyn, Bart H.H., Hans Grönqvist, and Lena Lindahl, "Adolescent Time Preferences Predict Lifetime Outcomes," *Economic Journal*, 2014.

Gorodnichenko, Yuriy and Gerard Roland, "Individualism, Innovation, and Long-Run Growth," *Proceedings of the National Academy of Sciences*, 2011, *108* (Supplement 4), 21316–21319.

Guiso, Luigi and Monica Paiella, "Risk Aversion, Wealth, and Background Risk," *Journal of the European Economic Association*, 2008, *6* (6), 1109–1150.

_ **, Paola Sapienza, and Luigi Zingales**, "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, 2006, *20* (2), 23–48.

_ **,** _ **, and** _ , "Cultural Biases in Economic Exchange?," *Quarterly Journal of Economics*, 2009, *124* (3), 1095–1131.

Helmbrecht, Johannes, "Politeness Distinctions in Second Person Pronouns," *Deictic Conceptualisation of Space, Time and Person*, 2003, *112*, 168.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath, "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies," *American Economic Review*, 2001, *91* (2), 73–78.

Herrmann, Benedikt, Christian Thöni, and Simon Gächter, "Antisocial Punishment Across Societies," *Science*, 2008, *319* (5868), 1362–1367.

Kirby, Kris N. and Nancy M. Petry, "Heroin and Cocaine Abusers Have Higher Discount Rates for Delayed Rewards than Alcoholics or Non-Drug-Using Controls," *Addiction*, 2004, *99* (4), 461–471.

Knack, Stephen and Philip Keefer, "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *Quarterly Journal of Economics*, 1997, pp. 1251–1288.

Kosfeld, Michael and Devesh Rustagi, "Leader Punishment and Cooperation in Groups: Experimental Field Evidence from Commons Management in Ethiopia," *American Economic Review*, 2015, *105* (2), 747–783.

Meier, Stephan and Charles Sprenger, "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2010, pp. 193–210.

Michalopoulos, Stelios, "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 2012, *102* (4), 1508–1539.

Nordhaus, William D., "Geography and Macroeconomics: New Data and New Findings," *Proceedings of the National Academy of Sciences of the United States of America*, 2006, *103* (10), 3510–3517.

Rieger, Marc-Oliver, Mei Wang, and Thorsten Hens, "Risk Preferences Around the World," *Management Science*, forthcoming.

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir, "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study," *American Economic Review*, 1991, *81* (5), 1068–1095.

Rustagi, Devesh, Stefanie Engel, and Michael Kosfeld, "Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management," *Science*, 2010, *330* (6006), 961–965.

Spinath, Birgit, Frank M. Spinath, Nicole Harlaar, and Robert Plomin, "Predicting School Achievement From General Cognitive Ability, Self-Perceived Ability, and Intrinsic Value," *Intelligence*, 2006, *34* (4), 363–374.

Spolaore, Enrico and Romain Wacziarg, "Ancestry, Language and Culture," *Working Paper*, 2015.

Sutter, Matthias and Martin G. Kocher, "Trust and Trustworthiness Across Different Age Groups," *Games and Economic Behavior*, 2007, *59* (2), 364–382.

_ , _ , Daniela Glätzle-Rützler, and Stefan T. Trautmann, "Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior," *American Economic Review*, 2013, *103* (1), 510–531.

_ , Silvia Angerer, Daniela Glätzle-Rützler, and Philipp Lergetporer, "The Effects of Language on Children's Intertemporal Choices," *Working Paper*, 2014.

Tanaka, Tomomi, Colin F Camerer, and Quang Nguyen, "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *American Economic Review*, 2010, *100* (1), 557–571.

Ventura, Luigi, "Direct Measures of Time Preference," *Economic and Social Review*, 2003, *34* (3), 293–310.

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*, 2014, *13*.

_ , **Thorsten Chmura, Tyler 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*, 2015, *78* (2), 209–218.

Wang, Mei, Marc Oliver Rieger, and Thorsten Hens, "How Time Preferences Differ: Evidence from 45 Countries," *Working Paper*, 2011.

APPENDIX

A Global Preference Survey

A.1 Overview

The cross-country dataset measuring risk aversion, patience, positive and negative reciprocity, altruism, and trust, was collected through the professional infrastructure of the Gallup World Poll 2012. The data collection process consisted of three steps. First, an experimental validation procedure was conducted to select the survey items. Second, there was a pre-test of the selected survey items in a variety of countries to ensure implementability in a culturally diverse sample. Third, the final data set was collected through the regular professional data collection efforts in the framework of the World Poll 2012.

A.2 Experimental Validation

To ensure the behavioral validity of the preference measures, all underlying survey items were selected through an experimental validation procedure (see Falk et al. (2015) for details). To this end, a sample of 409 German undergraduates completed standard state-of-the-art financially incentivized laboratory experiments designed to measure risk aversion, patience, positive and negative reciprocity, altruism, and trust. The same sample of subjects then completed a large battery of potential survey items. In a final step, for each preference, those survey items were selected which jointly performed best in explaining the behavior under real incentives observed in the choice experiments.

A.3 Pre-Test and Adjustment of Survey Items

Prior to including the preference module in the Gallup World Poll 2012, it was tested in the field as part of the World Poll 2012 pre-test, which was conducted at the end of 2011 in 22 countries. The main goal of the pre-test was to receive feedback on each item from various cultural backgrounds in order to assess potential difficulties in understanding and differences in the respondents' interpretation of items. Based on respondents' feedback and suggestions, minor modifications were made to several items before running the survey as part of the World Poll 2012.

The pre-test was run in 10 countries in central Asia (Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Russia, Tajikistan, Turkmenistan, Uzbekistan) 2 countries in South-East Asia (Bangladesh and Cambodia), 5 countries in

Southern and Eastern Europe (Croatia, Hungary, Poland, Romania, Turkey), 4 countries in the Middle East and North Africa (Algeria, Jordan, Lebanon, and Saudi-Arabia), and 1 country in Eastern Africa (Kenya). In each country, the sample size was 10 to 15 people. Overall, more than 220 interviews were conducted. In most countries, the sample was mixed in terms of gender, age, educational background, and area of residence (urban / rural).

Participants in the pre-test were asked to state any difficulties in understanding the items and to rephrase the meaning of items in their own words. If they encountered difficulties in understanding or interpreting items, respondents were asked to make suggestions on how to modify the wording of the item in order to attain the desired meaning.

Overall, the understanding of both the qualitative items and the quantitative items was satisfactory. In particular, no interviewer received any complaints regarding difficulties in assessing the quantitative questions or understanding the meaning of the probability used in the hypothetical risky choice items. When asked about rephrasing the qualitative items in their own words, most participants seemed to have understood the items in exactly the way that was intended. Nevertheless, some (sub-groups of) participants suggested adjustments to the wording of some items. This resulted in minor changes to four items, relative to the "original" experimentally validated items:

1. The use of the term "lottery" in hypothetical risky choices was troubling to some Muslim participants. As a consequence, we dropped the term "lottery" and replaced it with "draw".

2. The term "charity" caused confusion in Eastern Europe and Central Asia, so it was replaced it with "good cause".

3. Some respondents asked for a clarification of the question asking about one's willingness to punish unfair behavior. This feedback lead to splitting the question into two separate items, one item asking for one's willingness to punish unfair behavior towards others, and another asking for one's willingness to punish unfair behavior towards oneself.

4. When asked about hypothetical choices between monetary amounts today versus larger amounts one year later, some participants, especially in countries with current or relatively recent phases of volatile and high inflation rates, stated that their answer would depend on the rate of inflation, or said that they would always take the immediate payment due to uncertainty with respect to future inflation. Therefore, we decided to add the following phrase

to each question involving hypothetical choices between immediate and future monetary amounts: "Please assume there is no inflation, i.e., future prices are the same as today's prices."

A.4 Selection of Countries

The goal when selecting countries was to ensure representative coverage of the global population. Thus, countries from each continent and each region within continents were chosen. Another goal was to maximize variation with respect to observables, such as GDP per capita, language, historical and political characteristics, or geographical location and climatic conditions. Accordingly, the selection process favored non-neighboring and culturally dissimilar countries. This procedure resulted in the following sample of 76 countries:
East Asia and Pacific: Australia, Cambodia, China, Indonesia, Japan, Philippines, South Korea, Thailand, Vietnam
Europe and Central Asia: Austria, Bosnia and Herzegovina, Croatia, Czech Republic, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Italy, Kazakhstan, Lithuania, Moldova, Netherlands, Poland, Portugal, Romania, Russia, Serbia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom
Latin America and Caribbean: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Guatemala, Haiti, Mexico, Nicaragua, Peru, Suriname, Venezuela
Middle East and North Africa: Algeria, Egypt, Iran, Iraq, Israel, Jordan, Morocco, Saudi Arabia, United Arab Emirates
North America: United States, Canada
South Asia: Afghanistan, Bangladesh, India, Pakistan, Sri Lanka
Sub-Saharan Africa: Botswana, Cameroon, Ghana, Kenya, Malawi, Nigeria, Rwanda, South Africa, Tanzania, Uganda, Zimbabwe

A.5 Sampling and Survey Implementation

A.5.1 Background

Since 2005, the international polling company Gallup has conducted an annual World Poll, in which it surveys representative population samples in almost every country around the world on, e.g., economic, social, political, and environmental issues. The collection of our preference data was embedded into the regular World Poll 2012 and hence made use of the pre-existing polling infrastructure of one of the largest

professional polling institutes in the world.[33]

Selecting Primary Sampling Units

In countries in which face-to-face interviews are conducted, the first stage of sampling is the identification of primary sampling units (PSUs), consisting of clusters of households. PSUs are stratified by population size and / or geography and clustering is achieved through one or more stages of sampling. Where population information is available, sample selection is based on probabilities proportional to population size. If population information is not available, Gallup uses simple random sampling.

In countries in which telephone interviews are conducted, Gallup uses a random-digit-dialing method or a nationally representative list of phone numbers. In countries with high mobile phone penetration, Gallup uses a dual sampling frame.

Selecting Households and Respondents

Gallup uses random route procedures to select sampled households. Unless an outright refusal to participate occurs, interviewers make up to three attempts to survey the sampled household. To increase the probability of contact and completion, interviewers make attempts at different times of the day, and when possible, on different days. If the interviewer cannot obtain an interview at the initially sampled household, he or she uses a simple substitution method.

In face-to-face and telephone methodologies, random respondent selection is achieved by using either the latest birthday or Kish grid methods.[34] In a few Middle East and Asian countries, gender-matched interviewing is required, and probability sampling with quotas is implemented during the final stage of selection. Gallup implements quality control procedures to validate the selection of correct samples and that the correct person is randomly selected in each household.

[33] See http://www.gallup.com/strategicconsulting/156923/worldwide-research-methodology.aspx

[34] The latest birthday method means that the person living in the household whose birthday among all persons in the household was the most recent (and who is older than 15) is selected for interviewing. With the Kish grid method, the interviewer selects the participants within a household by using a table of random numbers. The interviewer will determine which random number to use by looking at, e.g., how many households he or she has contacted so far (e.g., household no. 8) and how many people live in the household (e.g., 3 people, aged 17, 34, and 36). For instance, if the corresponding number in the table is 7, he or she will interview the person aged 17.

Sampling Weights

Ex post, data weighting is used to ensure a nationally representative sample for each country and is intended to be used for calculations within a country. These sampling weights are provided by Gallup. First, base sampling weights are constructed to account for geographic oversamples, household size, and other selection probabilities. Second, post-stratification weights are constructed. Population statistics are used to weight the data by gender, age, and, where reliable data are available, education or socioeconomic status.

A.5.2 Translation of Items

The items of the preference module were translated into the major languages of each target country. The translation process involved three steps. As a first step, a translator suggested an English, Spanish or French version of a German item, depending on the region. A second translator, being proficient in both the target language and in English, French, or Spanish, then translated the item into the target language. Finally, a third translator would review the item in the target language and translate it back into the original language. If differences between the original item and the back-translated item occurred, the process was adjusted and repeated until all translators agreed on a final version.

A.5.3 Adjustment of Monetary Amounts in Quantitative Items

All items involving hypothetical monetary amounts were adjusted for each country in terms of their real value. Monetary amounts were calculated to represent the same share of a country's median income in local currency as the share of the amount in Euro of the German median income since the validation study had been conducted in Germany. Monetary amounts used in the validation study with the German sample were "round" numbers to facilitate easy calculations (e.g., the expected return of a lottery with equal chances of winning and losing) and to allow for easy comparisons (e.g., 100 Euro today versus 107.50 in 12 months). To proceed in a similar way in all countries, monetary amounts were always rounded to the next "round" number. For example, in the quantitative items involving choices between a lottery and varying safe options, the value of the lottery was adjusted to a round number. The varying safe options were then adjusted proportionally as in the original version. While this necessarily resulted in some (very minor) variations in the real stake size between countries, it minimized cross-country differences in the understanding the quantitative items due to difficulties in assessing the involved monetary amounts.

A.6 Wording of Survey Items

In the following, "willingness to act" indicates the following introduction: *We now ask for your willingness to act in a certain way in four different areas. Please again indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing to do so". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.*

Similarly, "self-assessments" indicate that the respective statement was preceded by the following introduction: *How well do the following statements describe you as a person? Please indicate your answer on a scale from 0 to 10. A 0 means "does not describe me at all" and a 10 means "describes me perfectly". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.*

A.6.1 Patience

1. (Sequence of five interdependent quantitative questions:) *Suppose you were given the choice between receiving a payment today or a payment in 12 months. We will now present to you five situations. The payment today is the same in each of these situations. The payment in 12 months is different in every situation. For each of these situations we would like to know which you would choose. Please assume there is no inflation, i.e, future prices are the same as today's prices. Please consider the following: Would you rather receive 100 Euro today or x Euro in 12 months?*

   The precise sequence of questions was given by the "tree" logic in Figure S5.

2. (Willingness to act:) *How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?*

A.6.2 Risk Taking

1. *(Similar to self-assessment:) Please tell me, in general, how willing or unwilling you are to take risks. Please use a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means you are "very willing to take risks". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.*

2. (Sequence of five interdependent quantitative questions:) *Please imagine the following situation. You can choose between a sure payment of a particular*

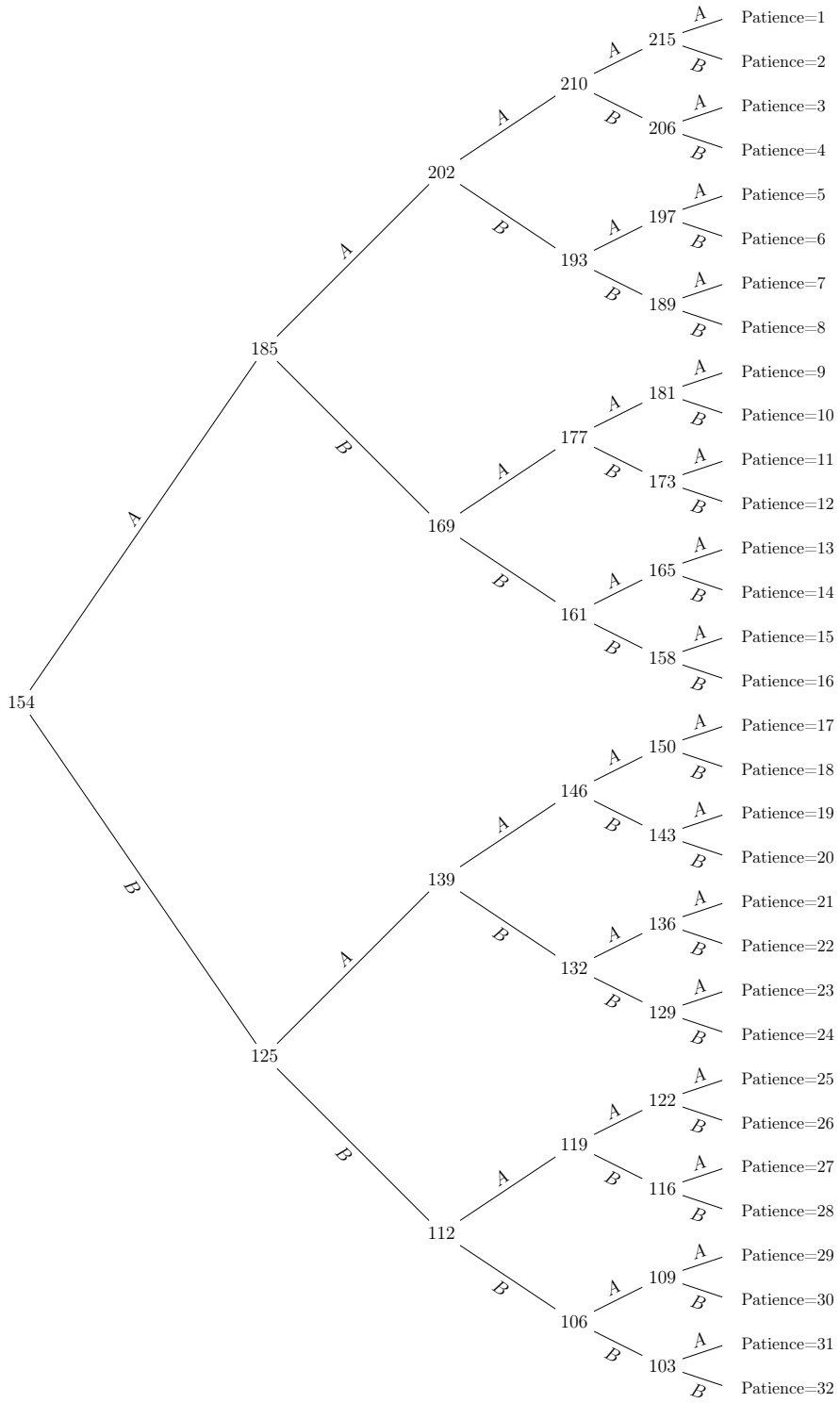

Figure 5: Tree for the staircase time task (numbers = payment in 12 months, A = choice of "100 euros today", B = choice of "x euros in 12 months". The staircase procedure worked as follows. First, each respondent was asked whether they would prefer to receive 100 euros today or 154 euros in 12 months from now (leftmost decision node). In case the respondent opted for the payment today ("A"), in the second question the payment in 12 months was adjusted upwards to 185 euros. If, on the other hand, the respondent chose the payment in 12 months, the corresponding payment was adjusted down to 125 euros. Working further through the tree follows the same logic.

amount of money, or a draw, where you would have an equal chance of getting amount x or getting nothing. We will present to you five different situations. What would you prefer: a draw with a 50 percent chance of receiving amount x, and the same 50 percent chance of receiving nothing, or the amount of y as a sure payment? The precise sequence of questions was given by the "tree" logic in Figure S6.

A.6.3 Positive Reciprocity

1. (Self-assessment:) *When someone does me a favor I am willing to return it.*

2. *(Hypothetical situation:) Please think about what you would do in the following situation. You are in an area you are not familiar with, and you realize you lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 Euro in total. However, the stranger says he or she does not want any money from you. You have six presents with you. The cheapest present costs 5 Euro, the most expensive one costs 30 Euro. Do you give one of the presents to the stranger as a "thank-you"-gift? If so, which present do you give to the stranger? No present / The present worth 5 / 10 / 15 / 20 / 25 / 30 Euro.*

A.6.4 Negative Reciprocity

1. (Self-assessment:) *If I am treated very unjustly, I will take revenge at the first occasion, even if there is a cost to do so.*

2. (Willingness to act:) *How willing are you to punish someone who treats you unfairly, even if there may be costs for you?*

3. (Willingness to act:) *How willing are you to punish someone who treats others unfairly, even if there may be costs for you?*

A.6.5 Altruism

1. *(Hypothetical situation:) Imagine the following situation: Today you unexpectedly received 1,000 Euro. How much of this amount would you donate to a good cause? (Values between 0 and 1000 are allowed.)*

2. (Willingness to act:) *How willing are you to give to good causes without expecting anything in return?*

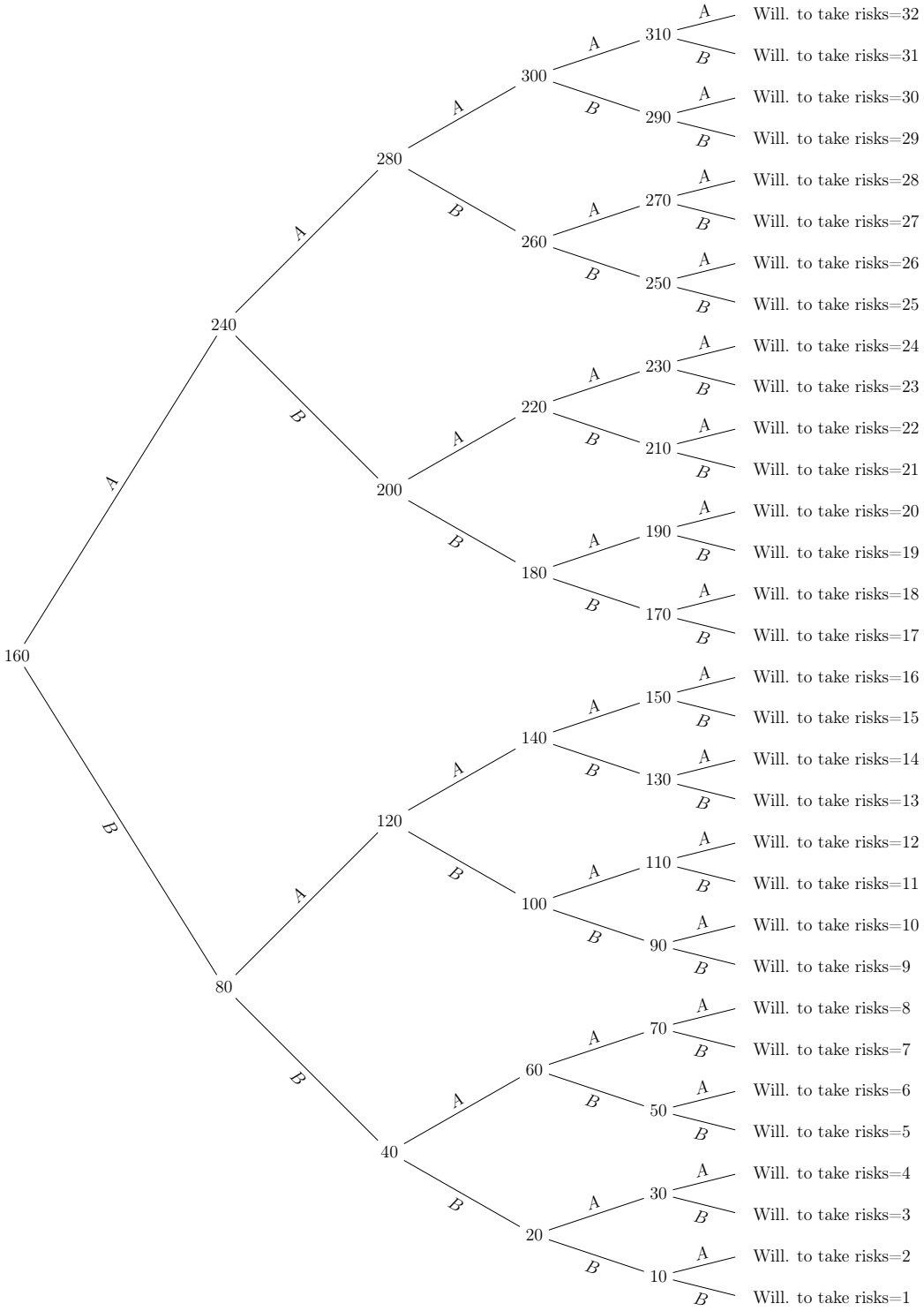

Figure 6: Tree for the staircase risk task (numbers = sure payment, A = choice of sure payment, B = choice of lottery). The staircase procedure worked as follows. First, each respondent was asked whether they would prefer to receive 160 euros for sure or whether they preferred a 50:50 chance of receiving 300 euros or nothing. In case the respondent opted for the safe choice ("B"), the safe amount of money being offered in the second question decreased to 80 euros. If, on the other hand, the respondent opted for the gamble ("A"), the safe amount was increased to 240 euros. Working further through the tree follows the same logic.

A.6.6 Trust

(Self-assessment:) *I assume that people have only the best intentions.*

A.7 Computation of Preference Measures

A.7.1 Cleaning and Imputation of Missings

In order to efficiently use all available information in our data, missing survey items were imputed based on the following procedure:

- If one (or more) survey items for a given preference were missing, then the missing items were predicted using the responses to the available items. The procedure was as follows:
  - Suppose the preference was measured using two items, call them a and b. For those observations with missing information on a, the procedure was to predict its value based on the answer to b and its relationship to a, which was estimated by regressing b on a for the sub-sample of subjects who had nonmissing information on both, a and b (on the world sample).
  - For the unfolding-brackets time and risk items, the imputation procedure was similar, but made additional use of the informational content of the responses of participants who started but did not finish the sequence of the five questions. Again suppose that the preference is measured using two items and suppose that a (the staircase measure) is missing. If the respondent did not even start the staircase procedure, then imputation was done using the methodology described above. On the other hand, if the respondent answered between one and four of the staircase questions, a was predicted using a different procedure. Suppose the respondent answered four items such that his final staircase outcome would have to be either x or y. A probit was run of the "x vs. y" decision on b, and the corresponding coefficients were used to predict the decision for all missings (note that this constitutes a predicted probability). The expected staircase outcome was then obtained by applying the predicted probabilities to the respective staircase endpoints, i.e., in this case x and y. If the respondent answered three (or less) questions, the same procedure was applied, the only difference being that in this case the obtained predicted probabilities were applied to the expected values of the staircase outcome conditional on reaching the respective node. Put differently, the procedure outlined above was applied recursively by working backwards through the

"tree" logic of the staircase procedure, resulting in an expected value for the outcome node.

  - If all survey items for a given preference were missing, then no imputation took place.

- Across the 12 survey items, between 0% and 8% of all responses had to be imputed.

A.7.2 Computation of Preference Indices at the Individual Level

For each of the traits (risk preferences, time preferences, positive reciprocity, negative reciprocity, altruism, and trust), an individual-level index was computed that aggregated responses across different survey items. Each of these indices was computed by (i) computing the z-scores of each survey item at the individual level and (ii) weighing these z-scores using the weights resulting from the experimental validation procedure of Falk et al. (2015). Formally, these weights are given by the coefficients of an OLS regression of observed behavior in the experimental validation study on responses to the respective survey items, such that the weights sum to one. In practice, for almost all preferences, the coefficients assign roughly equal weight to all corresponding survey items. The weights are given by:

- Patience:

$$\text{Patience} = 0.7115185 \times \text{Staircase time} + 0.2884815 \times \text{Will. to give up sth. today}$$

- Risk taking:

$$\text{Risk taking} = 0.4729985 \times \text{Staircase risk} + 0.5270015 \times \text{Will. to take risks}$$

- Positive reciprocity:

$$\text{Pos. reciprocity} = 0.4847038 \times \text{Will. to return favor} + 0.5152962 \times \text{Size of gift}$$

- Negative reciprocity:

$$\begin{aligned}\text{Neg. reciprocity} = {} & 0.5261938/2 \times \text{Will. to punish if oneself treated unfairly} \\ & + 0.5261938/2 \times \text{Will. to punish if other treated unfairly} \\ & + 0.3738062 \times \text{Will. to take revenge}\end{aligned}$$

As explained above, in the course of the pre-test, the negative reciprocity survey item asking people for their willingness to punish others was split up into two questions, one asking for the willingness to punish if oneself was treated unfairly and one asking for the willingness to punish if someone was treated unfairly. In order to apply the weighting procedure from the validation procedure to these items, the weight of the original item was divided by two and these modified weights were assigned to the new questions.

- Altruism:

  $$\text{Altruism } = 0.5350048\times \text{ Will. to give to good causes } +0.4649952\times \text{ Hypoth. donation}$$

- Trust: The survey included only one corresponding item.

A.7.3 Computation of Country Averages

In order to compute country-level averages, individual-level data were weighted with the sampling weights provided by Gallup, see Section S 1.5.1. These sampling weights ensure that our measures correctly represent the population at the country level.

A.8 Histograms by Preference

A.8.1 Individual Level

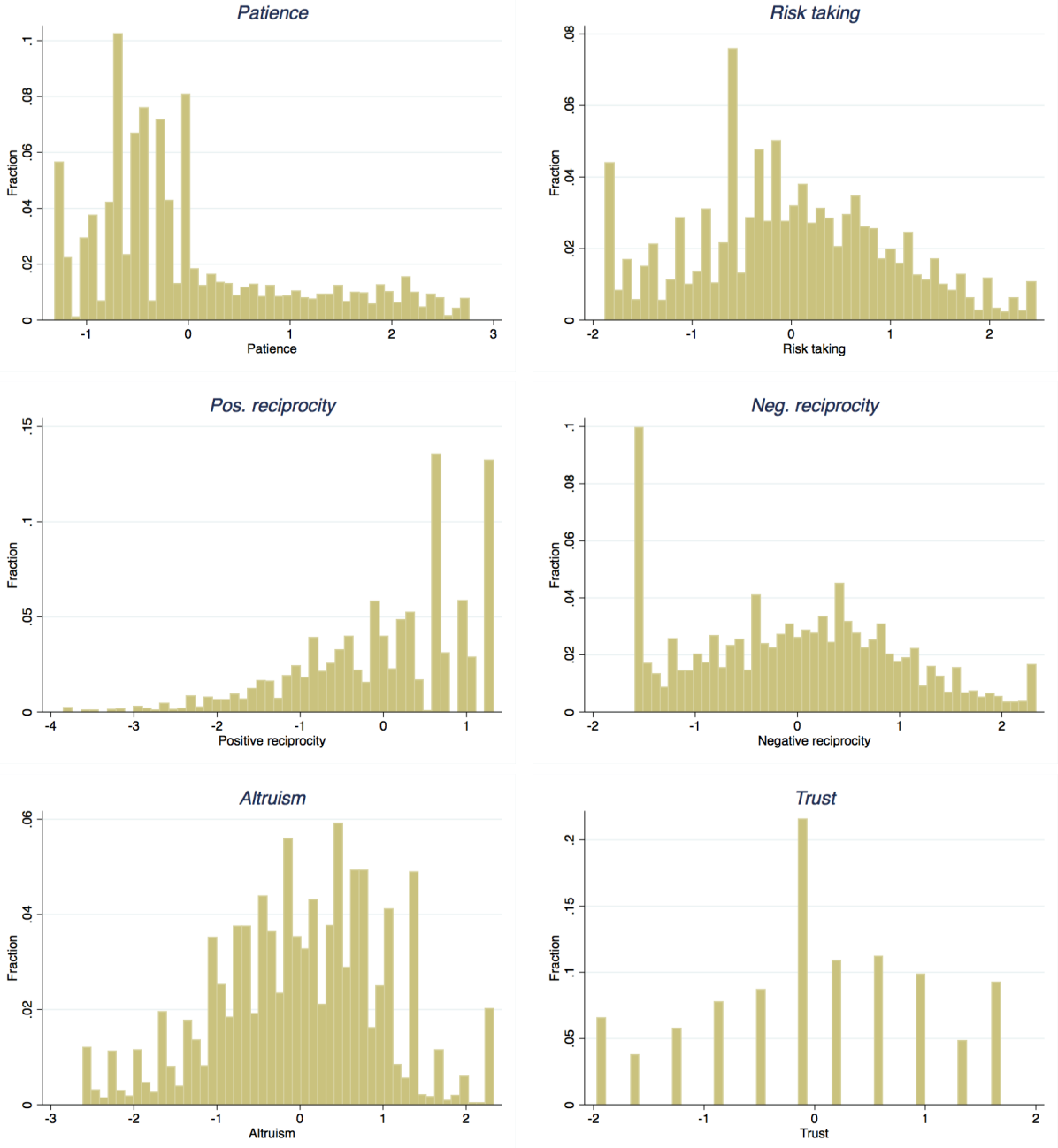

Figure 7: Distribution of preferences at individual level. The figure plots the distribution of standardized preference measures at the individual level. All data are standardized at the level of the individual in the full sample.

A.8.2 Country Level

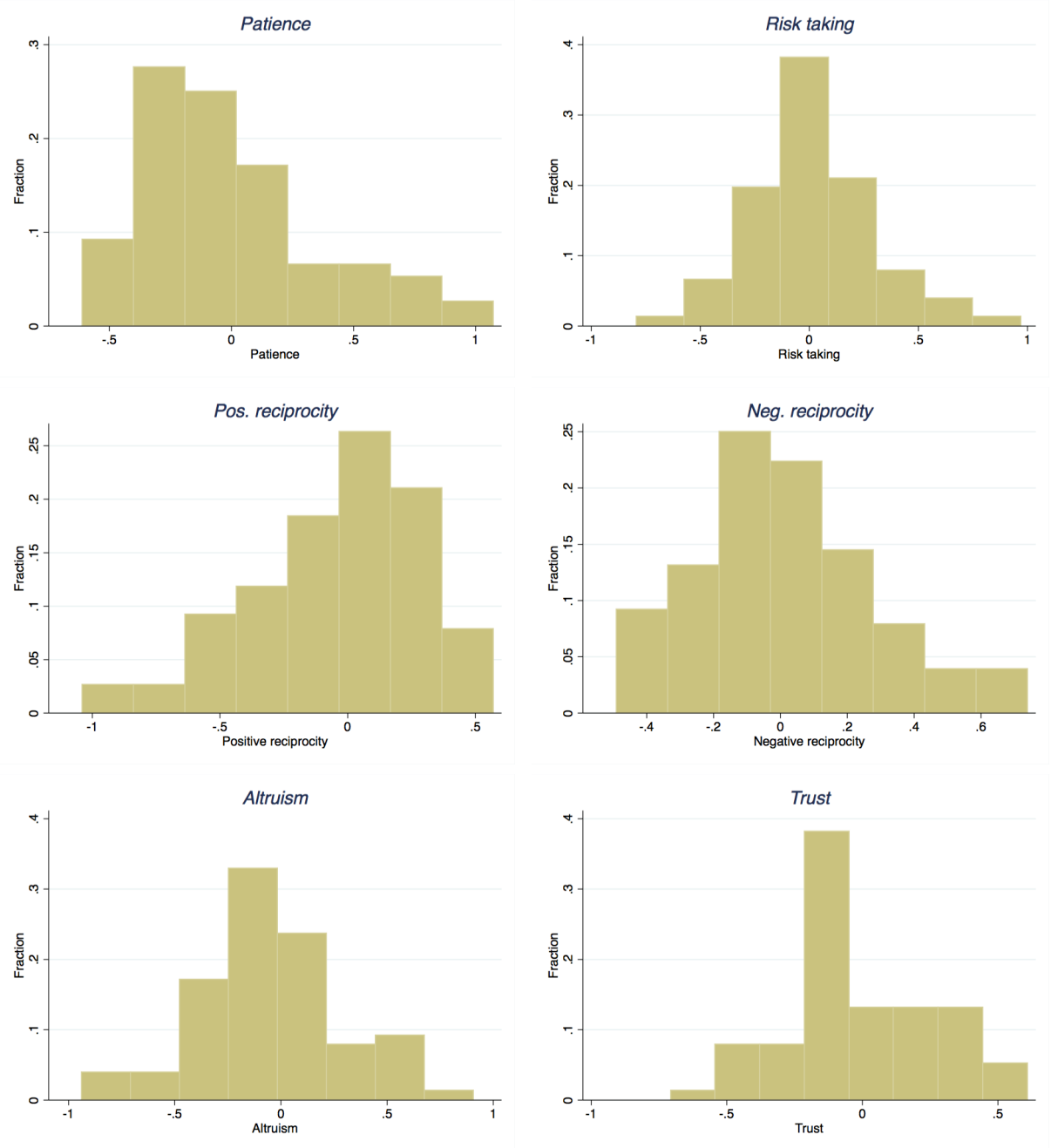

Figure 8: Distribution of preferences at country level. The figure plots the distribution of country averages of standardized preferences. All data are standardized at the level of the individual using the full sample.

B Correlations Among Individual-Level Preferences

Table 10 reports the correlation structure among preferences at the individual level. The correlations are computed conditional on country fixed effects to ensure that level differences in preferences across countries do not spuriously generate the results. At the same time, the correlation structure without country fixed effects is quantitatively very similar and is available upon request.

Table 10: Partial correlations between preferences at individual level conditional on country fixed effects

	Patience	Risk taking	Pos. reciprocity	Neg. reciprocity	Altruism	Trust
Patience	1					
Risk taking	0.210***	1				
Pos. reciprocity	0.084***	0.068***	1			
Neg. reciprocity	0.112***	0.228***	0.010***	1		
Altruism	0.098***	0.106***	0.329***	0.067***	1	
Trust	0.044***	0.047***	0.114***	0.075***	0.151***	1

Notes. Pairwise partial correlations between preferences at individual level, conditional on country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The next step in the analysis shows that the significant individual-level correlations among preferences in the world sample are not driven by a few outlier countries only. To this end, Table 11 shows the number of countries in which each pair of preferences is significantly correlated at the 1% level. The results show that in most cases the correlations are significant in a large fraction of the 76 countries.

Table 11: Number of countries in which preferences are significantly correlated

	Patience	Risk taking	Pos. reciprocity	Neg. reciprocity	Altruism	Trust
Patience						
Risk taking	71					
Pos. reciprocity	40	30				
Neg. reciprocity	53	73	19			
Altruism	47	50	76	32		
Trust	21	24	54	37	62	

Notes. Number of countries for which a given pair of preferences is significantly correlated at the 1% level.

C Scatter Plots of Preferences by World Region

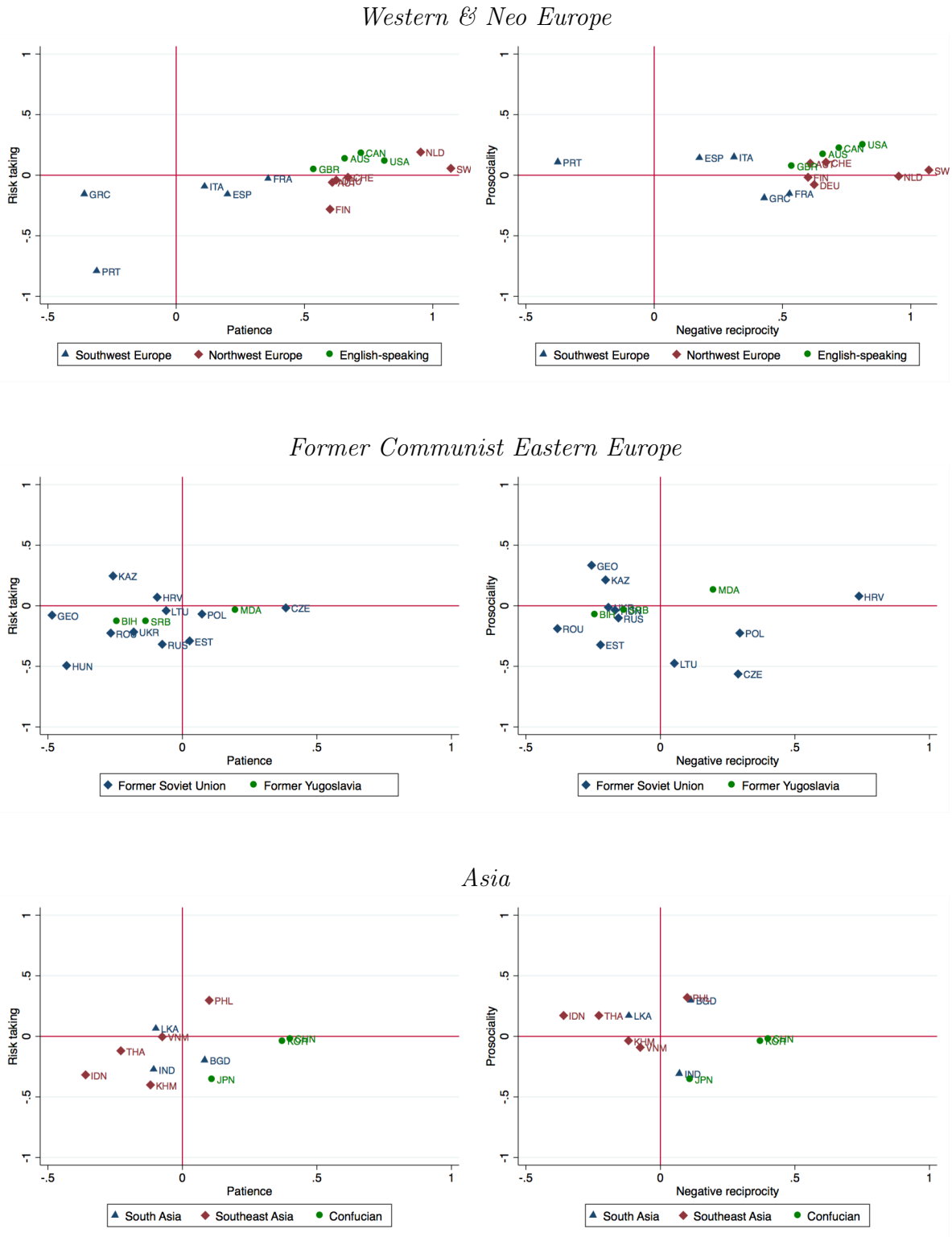

Figure 9a: Risk, time, and social preferences by world region (1/2). Each subpanel (row) plots risk taking, patience, negative reciprocity, and prosociality of all countries within a given world region. The prosociality score is computed as the average of altruism, positive reciprocity, and trust.

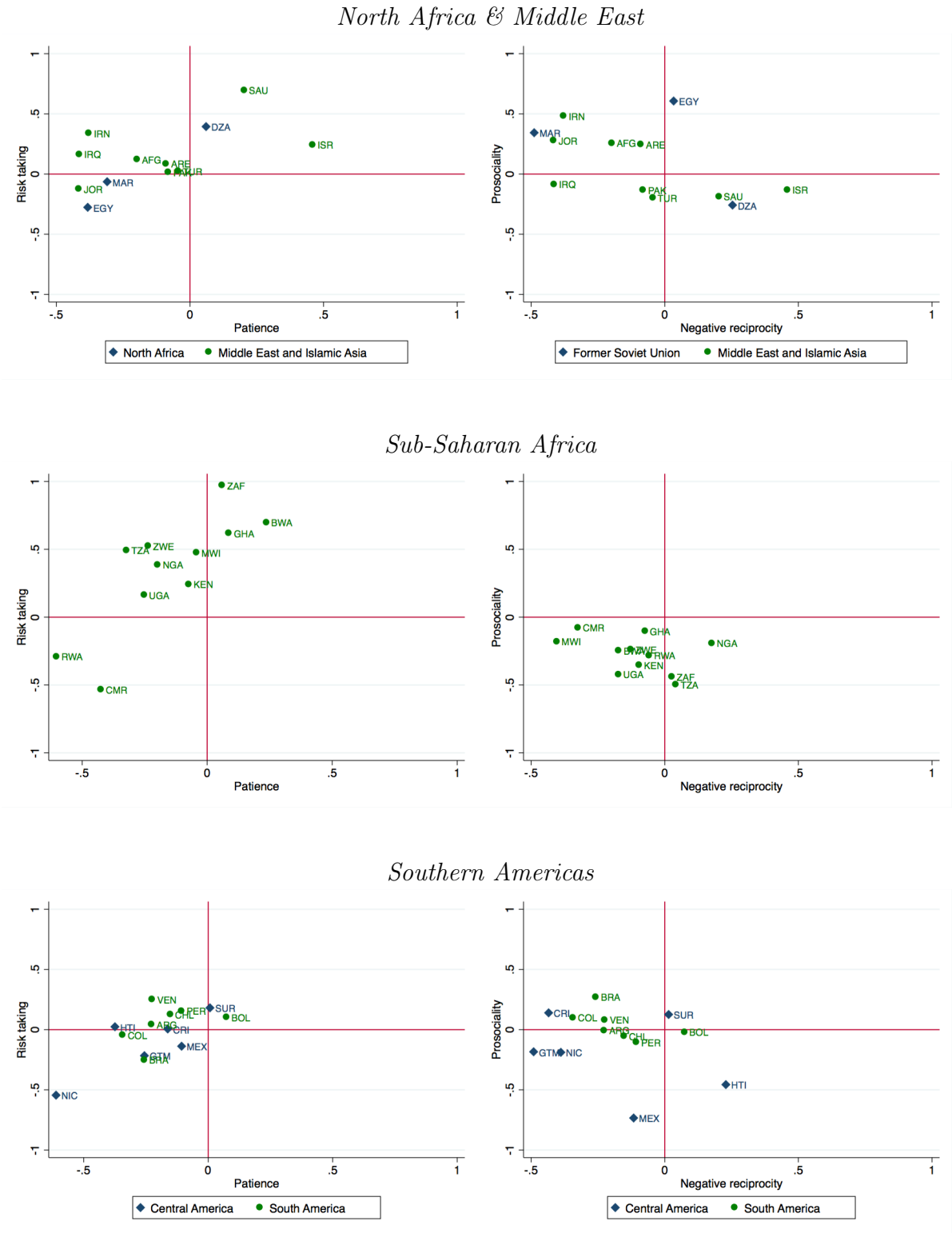

Figure 9b: Risk, time, and social preferences by world region (2/2). Each subpanel (row) plots risk taking, patience, negative reciprocity, and prosociality of all countries within a given world region. The prosociality score is computed as the average of altruism, positive reciprocity, and trust.

D Regressions for Variance Explained Across Countries

Table 12: Preferences and geographic, climatic, religious, and diversity variables (1/2)

	Dependent variable:														
	Patience					Risk taking					Pos. reciprocity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Distance to equator	0.013***				0.0047	-0.0014				-0.0020	-0.0040				-0.0024
	(0.00)				(0.01)	(0.00)				(0.01)	(0.00)				(0.01)
Longitude	0.000048				-0.00015	-0.000082				-0.00016	0.0013*				0.0017*
	(0.00)				(0.00)	(0.00)				(0.00)	(0.00)				(0.00)
% at risk of malaria	0.093				-0.17	0.12				-0.15	-0.44***				-0.43**
	(0.13)				(0.15)	(0.15)				(0.19)	(0.14)				(0.20)
Average precipitation		0.00050			0.00035		-0.0033***			-0.0027**		0.0019			0.00095
		(0.00)			(0.00)		(0.00)			(0.00)		(0.00)			(0.00)
Average temperature		-0.019**			-0.0010		0.0072			0.0067		-0.0022			-0.014
		(0.01)			(0.01)		(0.01)			(0.01)		(0.01)			(0.01)
% living in (sub-)tropical zones		-0.11			-0.069		0.27			0.26		-0.26			0.13
		(0.17)			(0.23)		(0.16)			(0.24)		(0.21)			(0.32)
Linguistic diversity			0.50**		0.22			0.081		0.20			-0.31		-0.53*
			(0.22)		(0.25)			(0.28)		(0.34)			(0.25)		(0.29)
Religious fractionalization			0.57***		0.43*			0.30*		0.41**			-0.22		-0.066
			(0.15)		(0.24)			(0.16)		(0.18)			(0.19)		(0.22)
Ethnic fractionalization			-0.95***		-0.39			0.26		-0.12			0.22		0.62**
			(0.21)		(0.24)			(0.23)		(0.29)			(0.18)		(0.24)
Constant	-0.41***	0.31**	0.00024	-0.31**	-0.45	0.036	0.089	-0.24**	0.25	0.085	0.14	-0.080	0.049	-0.12	0.068
	(0.11)	(0.13)	(0.09)	(0.13)	(0.44)	(0.10)	(0.08)	(0.10)	(0.18)	(0.40)	(0.11)	(0.10)	(0.10)	(0.10)	(0.47)
Religion shares	No	No	No	Yes	Yes	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73
R^2	0.267	0.243	0.280	0.377	0.535	0.039	0.196	0.139	0.155	0.339	0.128	0.053	0.039	0.088	0.264

Notes. OLS estimates, robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Preferences and geographic, climatic, religious, and diversity variables (2/2)

	Dependent variable:														
	Neg. reciprocity					Altruism					Trust				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Distance to equator	0.0029				0.0051	-0.010***				-0.012**	0.00013				0.00024
	(0.00)				(0.00)	(0.00)				(0.01)	(0.00)				(0.01)
Longitude	0.0013***				0.0012	0.0016**				0.0011	0.0012*				0.0012
	(0.00)				(0.00)	(0.00)				(0.00)	(0.00)				(0.00)
% at risk of malaria	-0.052				0.060	-0.51***				-0.74***	-0.30**				-0.29**
	(0.13)				(0.16)	(0.12)				(0.15)	(0.13)				(0.14)
Average precipitation		-0.00023			-0.00034		0.0016			0.00051		0.00046			-0.00033
		(0.00)			(0.00)		(0.00)			(0.00)		(0.00)			(0.00)
Average temperature		0.0040			-0.0014		0.0032			-0.020**		0.0035			0.0012
		(0.00)			(0.01)		(0.01)			(0.01)		(0.01)			(0.01)
% living in (sub-)tropical zones		-0.21			-0.034		-0.12			0.51**		-0.29			0.069
		(0.15)			(0.21)		(0.17)			(0.25)		(0.19)			(0.29)
Linguistic diversity			0.43**		0.36			-0.21		-0.31			0.16		-0.14
			(0.18)		(0.26)			(0.25)		(0.31)			(0.22)		(0.24)
Religious fractionalization			-0.024		0.20			-0.069		0.34			-0.23		-0.028
			(0.13)		(0.20)			(0.18)		(0.21)			(0.14)		(0.18)
Ethnic fractionalization			-0.51***		-0.39			0.13		-0.073			-0.27		0.074
			(0.17)		(0.28)			(0.21)		(0.29)			(0.20)		(0.23)
Constant	-0.099	0.028	0.11	-0.0023	-0.13	0.35***	-0.18	-0.00074	0.073	0.66*	0.016	-0.023	0.14*	-0.41***	-0.32
	(0.09)	(0.08)	(0.08)	(0.15)	(0.37)	(0.12)	(0.12)	(0.10)	(0.12)	(0.38)	(0.10)	(0.09)	(0.08)	(0.07)	(0.36)
Religion shares	No	No	No	Yes	Yes	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73
R^2	0.136	0.081	0.092	0.067	0.284	0.163	0.040	0.009	0.097	0.399	0.174	0.086	0.072	0.242	0.337

Notes. OLS estimates, robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Individual-Level Characteristics and Preferences

E.1 Age Profiles Separately by World Region

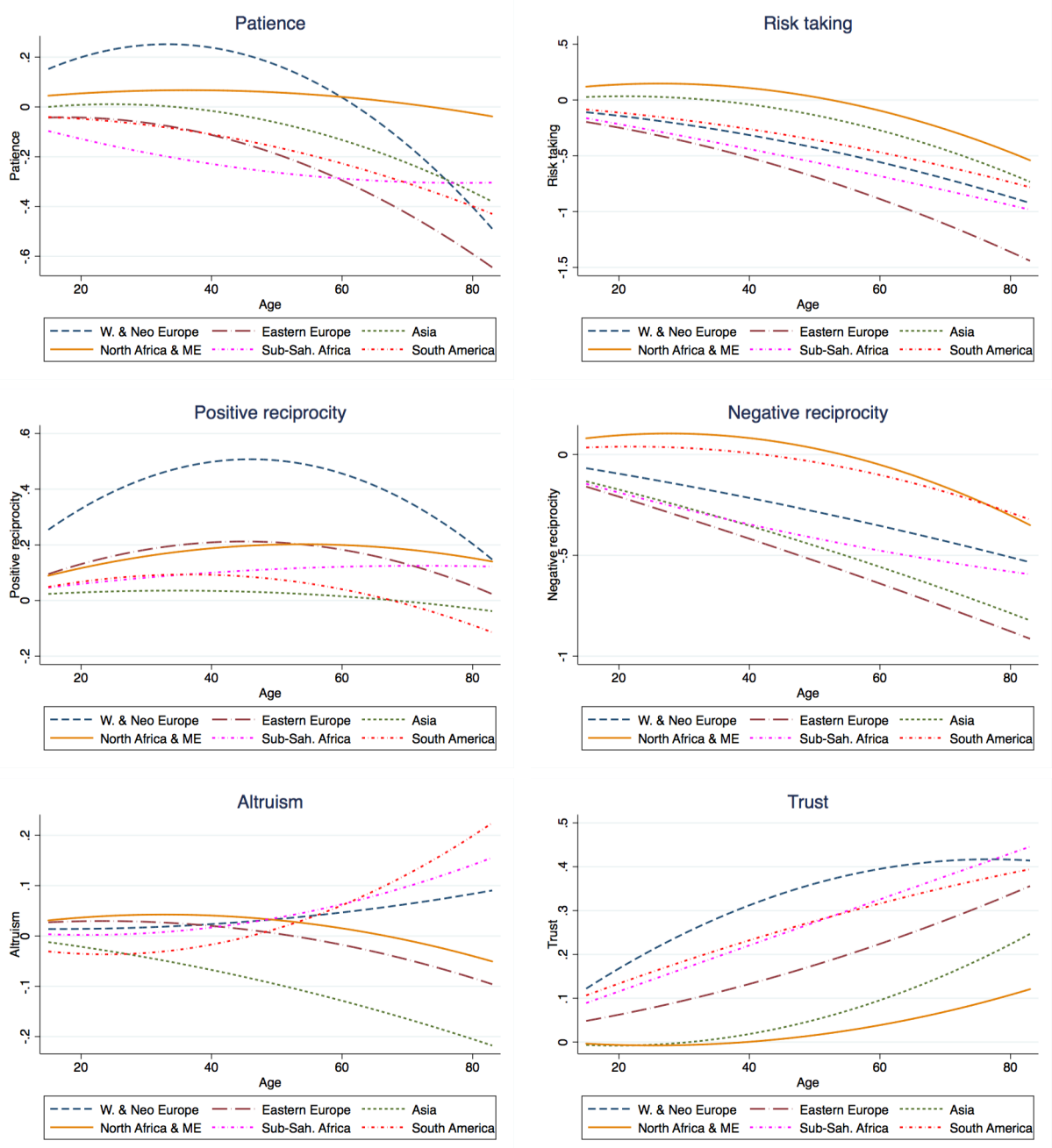

Figure 10: Age profiles separately by continent.

E.2 Cognitive Ability and Preferences by Country

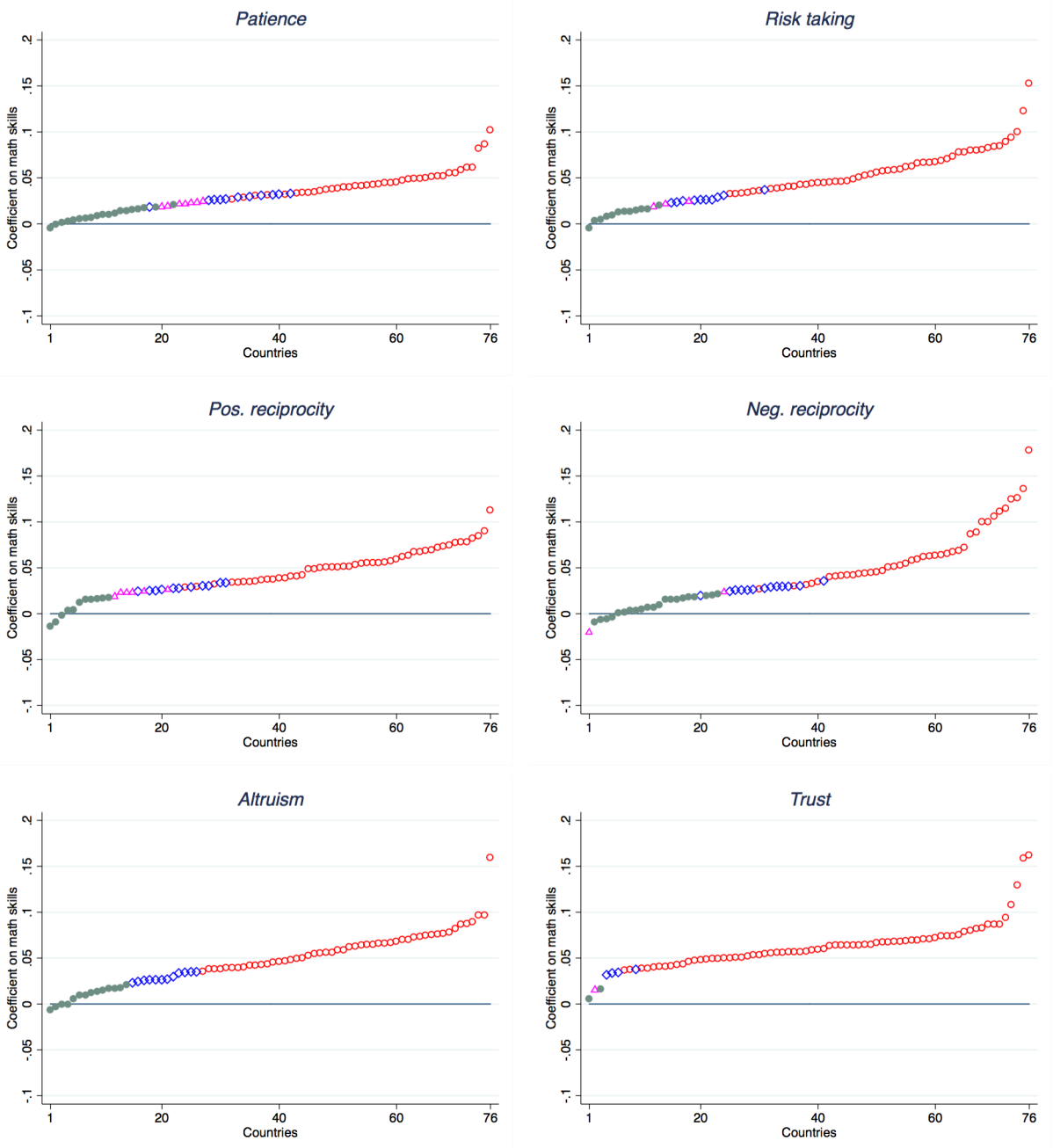

Figure 11: Cognitive ability correlations separately by country. Each panel plots the distribution of cognitive ability correlations. That is, for each country, we regress the respective preference on gender, age and its square, and subjective math skills, and plot the resulting math skill coefficients as well as their significance level. In order to make countries comparable, each preference was standardized (z-scores) within each country before computing the coefficients. Green dots indicate countries in which the cognitive ability effect is not statistically different from zero at the 10% level, while red / blue / pink dots denote countries in which the effect is significant at the 1% / 5% / 10% level, respectively. Positive coefficients imply that higher cognitive ability people have higher values in the respective preference.

F Individual-Level Behaviors

F.1 Distribution of Coefficients Across Countries

This section shows that the conditional correlations on the relationships between preferences and individual-level behaviors that we reported on the global level in the main text, are not due to a few outlier countries only. Instead, the results suggest that our preference measures predict behavior across a broad set of countries. To show this, we regress the behaviors discussed in Section 5 on the respective preference, separately for each country, and then plot the distribution and statistical significance of the resulting coefficients. For instance, the top left panel in Figure 12 shows that the positive correlation between patience and savings holds in virtually all countries in our sample.

While Figure 12 reports the results for patience and risktaking, Figure 13 visualizes the relationships between altruism and behaviors. Finally, Figure 14 presents the correlations between positive and negative reciprocity and the behaviors discussed in Section 5 of the main text.

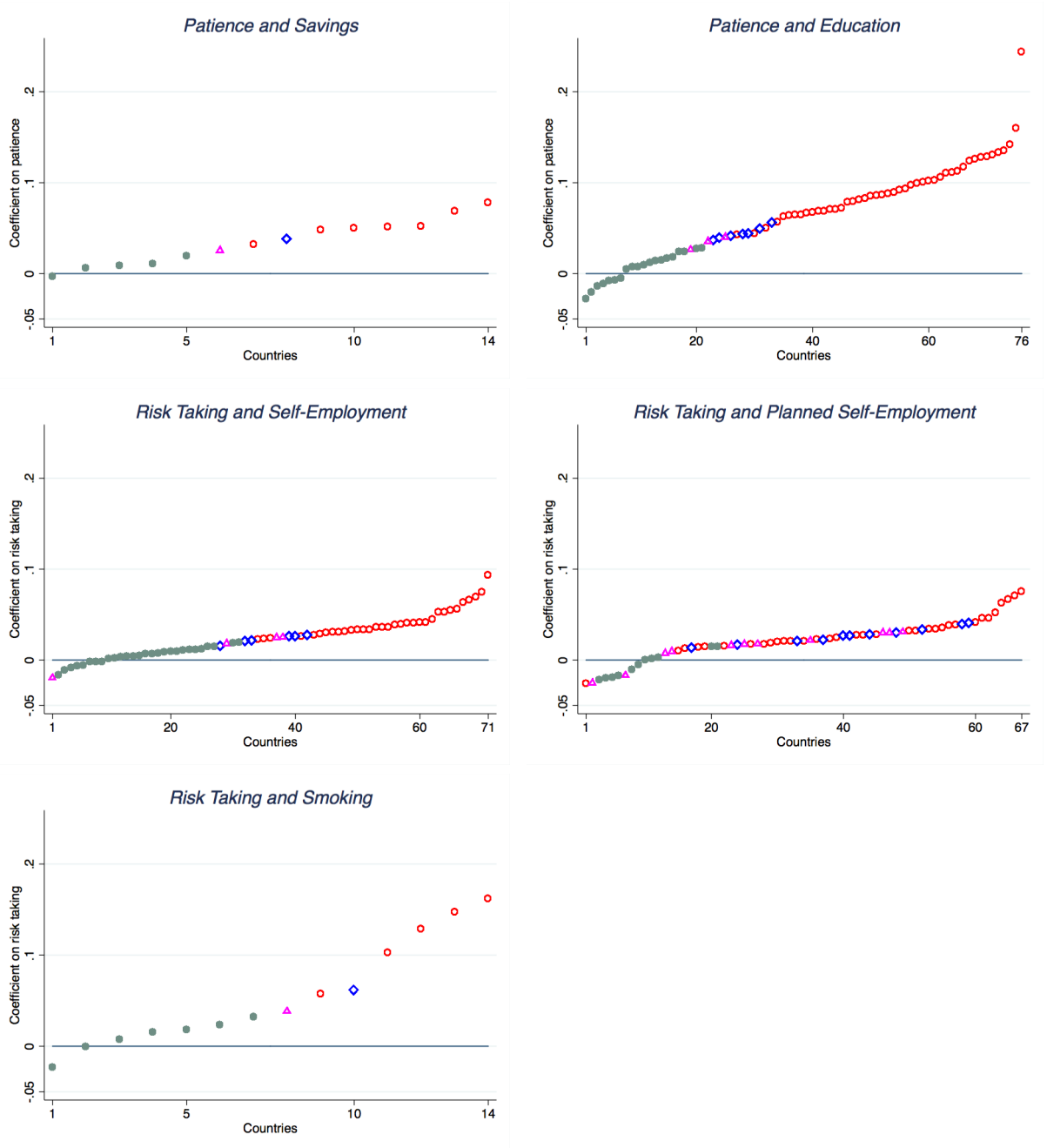

Figure 12: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable, each preference was standardized (z-scores) within each country before computing the coefficients. Green dots indicate countries in which the correlation is not statistically different from zero at the 10% level, while red / blue / pink dots denote countries in which the correlation is significant at the 1% / 5% / 10% level, respectively. Positive coefficients imply that a higher preference measure is related to a higher outcome measure.

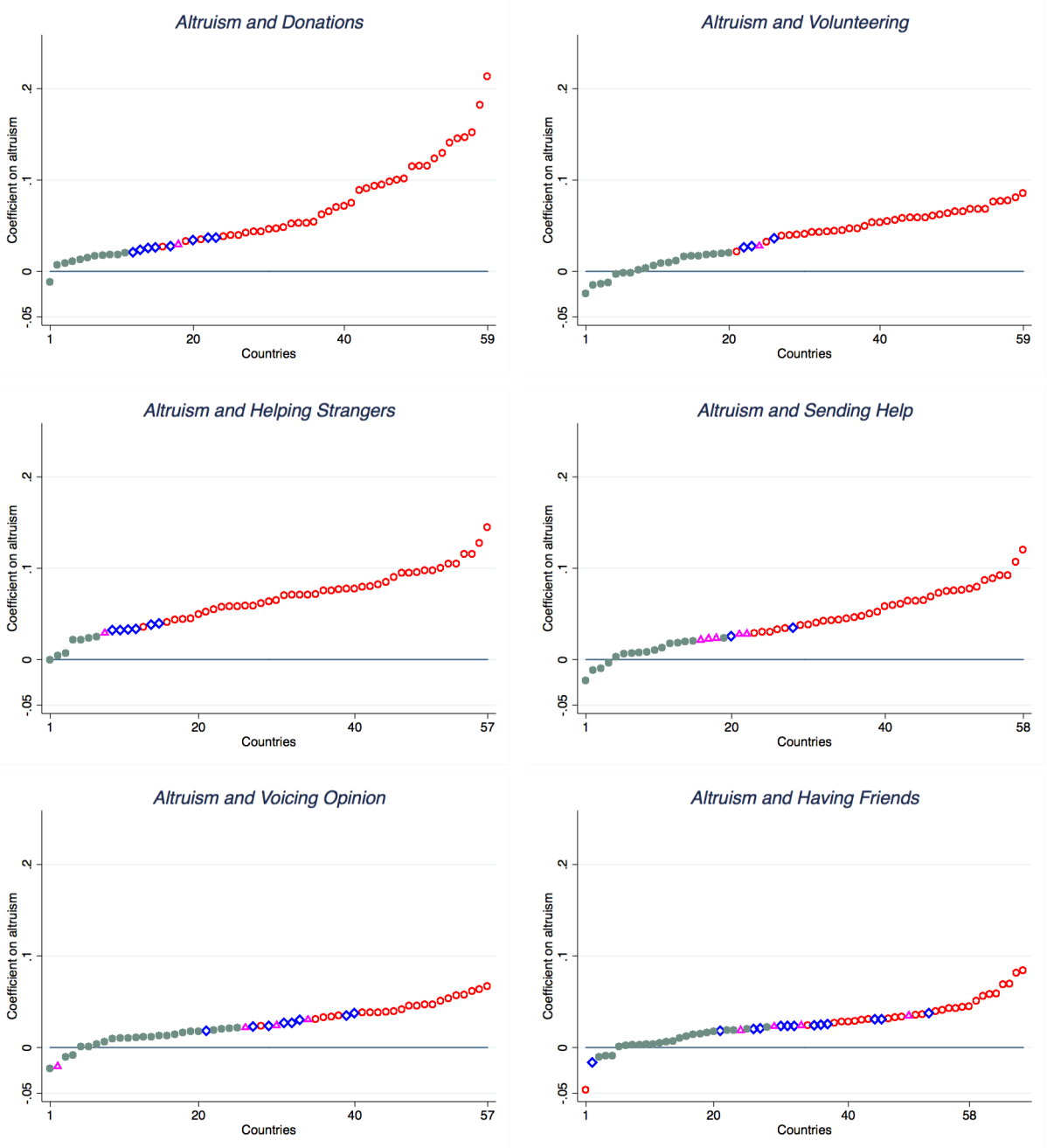

Figure 13: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable, each preference was standardized (z-scores) within each country before computing the coefficients. Green dots indicate countries in which the correlation is not statistically different from zero at the 10% level, while red / blue / pink dots denote countries in which the correlation is significant at the 1% / 5% / 10% level, respectively. Positive coefficients imply that a higher preference measure is related to a higher outcome measure.

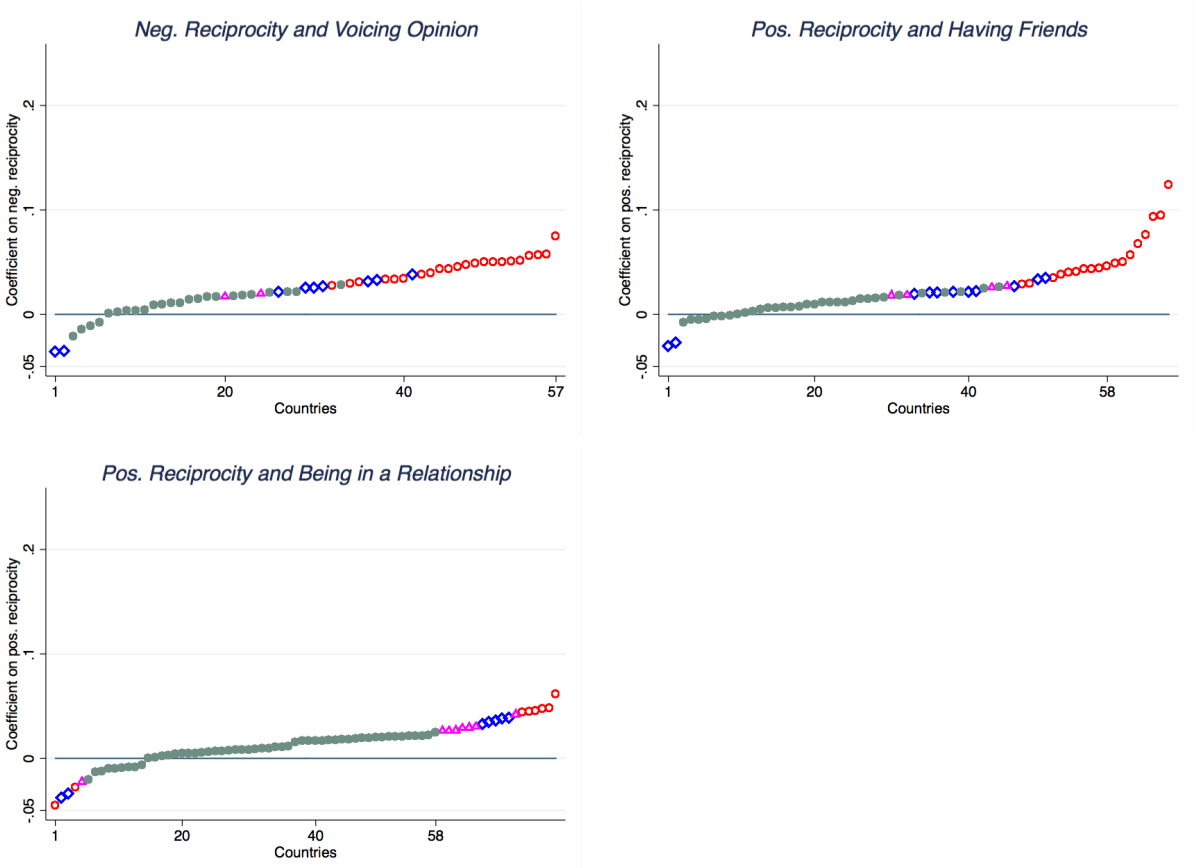

Figure 14: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable, each preference was standardized (z-scores) within each country before computing the coefficients. Green dots indicate countries in which the correlation is not statistically different from zero at the 10% level, while red / blue / pink dots denote countries in which the correlation is significant at the 1% / 5% / 10% level, respectively. Positive coefficients imply that a higher preference measure is related to a higher outcome measure.

F.2 Robustness Checks

This appendix reports robustness checks on the relationship between preferences and behaviors at the individual level. Specifically, while Section 5 of the main text reported the results of OLS estimations, we now re-estimate all specifications using probit or ordered probit regressions. As Tables 14 and 15 show, the results are unchanged.

Table 14: Patience and accumulation decisions, risk preferences and risky choices

	Dependent variable:									
	Accumulation decisions				Risky choices					
	Saved last year		Education level		Own business		Plan to start business		Smoking intensity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Patience	0.15***	0.079***	0.20***	0.070***						
	(0.03)	(0.02)	(0.03)	(0.01)						
Risk taking					0.14***	0.11***	0.17***	0.11***	0.097***	0.050**
					(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Age		0.021		2.19***		9.23***		5.71***		6.30***
		(0.76)		(0.56)		(0.53)		(0.43)		(0.75)
Age squared		-0.63		-4.08***		-9.50***		-7.50***		-7.13***
		(0.78)		(0.54)		(0.57)		(0.52)		(0.72)
1 if female		-0.0036		-0.036		-0.27***		-0.13***		-1.32***
		(0.04)		(0.03)		(0.05)		(0.03)		(0.13)
Subj. math skills		0.043***		0.091***		0.029***		0.018***		-0.022***
		(0.01)		(0.00)		(0.00)		(0.00)		(0.01)
Log [Household income p/c]		0.36***		0.31***		0.092***		-0.040**		-0.026
		(0.04)		(0.02)		(0.01)		(0.02)		(0.02)
Constant	-0.61***	-2.86***	1.00***	4.94***	-1.07***	-4.18***	-1.24***	-2.07***	1.06***	1.37***
	(0.09)	(0.28)	(0.05)	(0.18)	(0.04)	(0.21)	(0.06)	(0.14)	(0.07)	(0.17)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Religion FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	15260	14459	79357	69272	72839	62985	57072	51489	15309	14490
Pseudo R^2	0.010	0.122	0.015	0.202	0.009	0.134	0.016	0.170	0.004	0.167

(Ordered) probit estimates, standard errors (clustered at country level) in parentheses. For the purposes of this table, age is divided by 100. Saved last year is a binary indicator, while education level is measured in three categories (roughly elementary, secondary, and tertiary education, see Appendix I). Self-employment and planned self-employment are binary, while smoking intensity is measured in three categories (never, occasionally, frequently). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 15: Social preferences and social interactions

	Dependent variable:													
	Donated money		Volunteered time		Helped stranger		Sent money / goods to other individual		Voiced opinion to official		Have friends or relatives I can count on in need		In a relationship	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Altruism	0.19***	0.20***	0.16***	0.14***	0.16***	0.14***	0.12***	0.12***	0.097***	0.093***	0.029	0.066***	0.012	0.0040
	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Positive reciprocity	0.015	0.0029	-0.00046	0.017	0.091***	0.091***	0.062***	0.064***	-0.0071	-0.011	0.053*	0.067***	0.056***	0.023***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)
Negative reciprocity	0.013	-0.014	-0.0097	-0.0089	-0.0062	-0.0080	0.036**	0.022*	0.056***	0.060***	0.045***	0.011	-0.0088	-0.00076
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Age		2.13***		1.57***		2.01***		0.94***		3.85***		-3.76***		16.4***
		(0.26)		(0.29)		(0.20)		(0.24)		(0.27)		(0.38)		(0.59)
Age squared		-1.63***		-1.76***		-2.47***		-1.04***		-3.84***		3.11***		-15.9***
		(0.29)		(0.30)		(0.20)		(0.28)		(0.30)		(0.36)		(0.59)
1 if female		0.044**		-0.064***		-0.040**		0.0018		-0.16***		0.051***		-0.086***
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)
Subj. math skills		0.030***		0.027***		0.025***		0.026***		0.033***		0.022***		0.0067**
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Log [Household income p/c]		0.10***		0.016*		0.056***		0.12***		0.042***		0.14***		-0.11***
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.02)		(0.01)
Constant	-0.46***	-1.71***	-0.78***	-1.09***	-0.018	-0.36**	-0.70***	-2.19***	-0.77***	-1.91***	0.90***	-0.11	0.21***	-2.37***
	(0.07)	(0.12)	(0.05)	(0.15)	(0.04)	(0.16)	(0.06)	(0.14)	(0.04)	(0.12)	(0.06)	(0.15)	(0.04)	(0.15)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Religion FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	58229	53439	58213	53430	55991	53226	56253	53559	55944	53174	65986	59209	77881	68176
Pseudo R^2	0.016	0.161	0.012	0.086	0.020	0.070	0.011	0.115	0.006	0.061	0.004	0.121	0.002	0.174

Probit estimates, standard errors (clustered at country level) in parentheses. For the purposes of this table, age is divided by 100.See Appendix I for details on all dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G Details for Relationship Between Preferences and FTR

G.1 Individual-Level Regressions Separately by Country

Table 16: Preferences and FTR: Within-country results

Country	Weak FTR	Strong FTR	Patience	Pos. reciprocity	Trust	Altruism
Estonia	Estonian	Russian	0.05	0.13*	0.38***	0.45***
Nigeria	Yoruba	English, Hausa, Igbo	-0.08	0.54***	0.63***	-0.11
Switzerland	German	French, Italian	0.17**	0.14**	0.28***	0.30***

OLS estimates, robust standard errors. The regressions report the coefficient on FTR in univariate regressions for each country in which we observe within-country variation in FTR. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G.2 Country-Level Regressions: Robustness

While the main text reported WLS estimates, Table 17 reports OLS estimates.

Table 17: Preferences and FTR: Cross-country results

	Dependent variable:											
	Patience		Risk taking		Pos. reciprocity		Neg. reciprocity		Altruism		Trust	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fraction of population speaking weak FTR	0.36***	0.23**	-0.11	0.015	0.15*	0.17**	-0.018	-0.082	0.061	0.11	0.19**	0.19**
	(0.13)	(0.09)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.09)	(0.09)	(0.09)	(0.08)	(0.07)
Log [GDP p/c PPP]		0.16***		0.032		-0.073*		0.052		-0.077*		-0.0055
		(0.04)		(0.03)		(0.04)		(0.04)		(0.04)		(0.03)
Distance to equator		0.010*		0.0015		-0.0078		-0.0054		-0.0033		-0.0057
		(0.01)		(0.00)		(0.01)		(0.01)		(0.01)		(0.00)
Longitude		-0.0018		0.0024		0.0021		0.00043		0.0028		0.000065
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
% at risk of malaria		0.25		-0.15		-0.33		-0.089		-0.72***		-0.092
		(0.19)		(0.24)		(0.28)		(0.17)		(0.26)		(0.19)
Average precipitation		-0.00013		-0.00081		0.00055		-0.0010		0.0031***		-0.0011
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Constant	-0.067	-1.42**	0.034	-0.51	-0.053	0.56	0.014	-0.019	-0.043	0.25	-0.058	0.39
	(0.04)	(0.56)	(0.04)	(0.45)	(0.05)	(0.67)	(0.04)	(0.50)	(0.05)	(0.61)	(0.04)	(0.48)
Continent FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	75	74	75	74	75	74	75	74	75	74	75	74
R^2	0.141	0.641	0.021	0.381	0.029	0.280	0.001	0.246	0.005	0.356	0.072	0.420

OLS estimates, robust standard errors in parentheses. The regressions exclude Haiti for which no respondent could be classified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

H Discussion of Measurement Error and Within- versus Between-Country Variation

In the presence of measurement error, a simple variance decomposition as shown in Table 2 tends to overstate the relative importance of within-country variation in preferences. This is because measurement error would be part of the within-country variation, whereas the aggregation to country averages mitigates measurement error and thus removes this source of variation. This section provides evidence that measurement error is unlikely to be large enough to drive the result.

To illustrate the impact of measurement error, recall that a simple regression of an individual-level preference measure M on a matrix of country dummies D yields

$$M = D'\gamma + \epsilon$$

In a setting without measurement error ϵ would be interpreted as individual specific effects that are not explained by the variation between countries. The total variance of M is given by

$$Var(M) = Var(\delta) + Var(\epsilon) + 2cov(\delta, \epsilon)$$

where $\delta = D'\gamma$. Note that the R^2 from a regression of M on the country dummies (i.e., $Var(\delta)/Var(M)$) could be interpreted as the between country-variation, i.e., the fraction of total variation explained by country dummies, if individual effects are unrelated to country effects.

If, however, the preference measure M measures the true preference parameter P with error, denoted e, the residual variation of the regression above does not only capture individual effects. Assume that M is a linear function of P and e, i.e.,

$$M = P + e,$$

such that we can rewrite

$$P + e = \delta + \epsilon$$

The total variance of the preference is hence

$$Var(P) = Var(\delta) + Var(\epsilon) - Var(e),$$

assuming that $\epsilon \perp \delta$ and $e \perp \epsilon, \delta, P$.

The regression model still allows identifying $Var(\delta)$, but the share of preference variation that is truly explained by the between-country variation is no longer

given by the R^2, $Var(\delta)/Var(M)$, but rather by $Var(\delta)/Var(P)$. To assess whether between-country or within-country effects explain a larger share of total variation, one needs to compare $Var(\delta)/Var(P)$ to $Var(\epsilon)/Var(P)$. Since $Var(P) = Var(M) - Var(e)$, $Var(e)$ needs to be determined.

The variance of measurement error, $Var(e)$, is not directly observable, but estimates of test-retest correlations of relevant preference measures are available which can be used to gauge the size of $Var(e)$. Based on arguments of plausibility, the variance of the measurement error does not appear to be large enough to invalidate the claim that the within-country variation is smaller than the between-country variation. Consider how large the proportion of measurement error in the total variation of M can be, with between-country effects still explaining a smaller share of variation than individual-specific effects. Note that between- and within-country variation add up to total variation in preferences absent measurement error: $Var(\delta)/Var(P) =$ 1- $Var(\epsilon)/Var(P)$. Thus, between-country effects explain a relatively smaller share of total variation if $Var(\delta)/Var(P) < 0.50$. Letting q with $0 < q \leq 1$ be the fraction of measurement error in M, this condition can be evaluated by scaling up the R^2 from a regression of M on the set of country dummies by $1/(1-q)$. I.e., if $Var(\delta)/(Var(M)(1-q)) < 0.5$, the between-country variation is smaller than the within-country variation, even accounting for measurement error.

Take, as an example, the estimate for risk-taking in Table 4, for which the regression of the risk measure on the set of country dummies yields an R^2 of 0.09. Solving $R^2 < 0.5(1-q)$ for q shows that as long as $q < 0.828$, the within country variation exceeds the between country variation. Previous work has shown that the test-retest correlation of the single components of this particular risk measure is around 0.6 (Beauchamp et al., 2011). This implies that, in order for measurement error alone to be able to explain the greater variation of preferences within-country than between-country, measurement error would have to be twice as large as existing evidence suggests.

I Description and Data Sources of Outcome Variables

I.1 Individual-Level Variables

Subjective law and order index. Included in Gallup's Background data (0-1). Derived from responses to three questions: "In the city or area where you live, do you have confidence in the local police force?"; "Do you feel safe walking alone at night in the city or area where you live?"; "Within the last 12 months, have you had money or property stolen from you or another household member?".

Subjective physical health index. Included in Gallup's Background data (0-1). Derived from responses to five questions: "Do you have any health problems that prevent you from doing any of the things people your age normally can do?"; "Now, please think about yesterday, from the morning until the end of the day. Think about where you were, what you were doing, who you were with, and how you felt. Did you feel well-rested yesterday?"; "Did you experience the following feelings during a lot of the day yesterday? How about physical pain?"; "Did you experience the following feelings during a lot of the day yesterday? How about worry?"; "Did you experience the following feelings during a lot of the day yesterday? How about sadness?".

Household income per capita. Included in Gallup's background data. To calculate income, respondents are asked to report their household income in local currency. Those respondents who have difficulty answering the question are presented a set of ranges in local currency and are asked which group they fall into. Income variables are created by converting local currency to International Dollars (ID) using purchasing power parity (PPP) ratios. Log household income is computed as log (1+ household income).

Education level. Included in Gallup's background data. Level 1: Completed elementary education or less (up to 8 years of basic education). Level 2: Secondary - 3 year tertiary education and some education beyond secondary education (9-15 years of education). Level 3: Completed four years of education beyond high school and / or received a 4-year college degree.

Subjective self-assessment of math skills. *How well do the following statements describe you as a person? Please indicate your answer on a scale from 0 to 10. A 0 means "does not describe me at all" and a 10 means "describes me perfectly".*

You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. I am good at math.

Saved last year. Binary variable capturing whether the respondent saved any money in the previous year. Included in Gallup's background data.

Own business. Binary variable capturing whether the respondent is self-employed. Included in Gallup's background data.

Plan to start business. Binary variable capturing whether the respondent is planning to start their own business (only asked of those who are not self-employed). Included in Gallup's background data.

Smoking intensity. Variable capturing how frequently a respondent smokes (0=never, 1=occasionally, 2=frequently). Included in Gallup's background data.

Donated money. Binary variable capturing whether the respondent donated money in the previous month. Included in Gallup's background data.

Volunteered time. Binary variable capturing whether the respondent volunteered time to an organization in the previous month. Included in Gallup's background data.

Helped stranger. Binary variable capturing whether the respondent helped a stranger who needed help in the previous month. Included in Gallup's background data.

Sent help to individual. Binary variable capturing whether the respondent sent help (money or goods) to another individual in the previous year. Included in Gallup's background data.

Voiced opinion to official. Binary variable capturing whether the respondent voiced their opinion to a public official in the previous month. Included in Gallup's background data.

Donated money. Binary variable capturing whether the respondent has relatives or friends they can count on to help them whenever needed. Included in Gallup's background data.

In a relationship. Binary variable coded as zero if the respondents is single, separated, divorced, or widowed, and as 1 if respondent is married or has a domestic partner. Included in Gallup's background data.

I.2 Country-Level Variables

Distance to equator, longitude. Source: the CEPII geo database.

Land suitability for agriculture. Index of the suitability of land for agriculture based on ecological indicators of climate suitability for cultivation, such as growing degree days and the ratio of actual to potential evapotranspiration, as well as ecological indicators of soil suitability for cultivation, such as soil carbon density and soil pH, taken from Michalopoulos (2012).

Temperature. Average monthly temperature of a country in degree Celsius, 1961-1990, taken from Ashraf and Galor (2013). Data originally based on geospatial average monthly temperature data for this period reported by the G-ECON project (Nordhaus, 2006).

Precipitation. Average monthly precipitation of a country in mm per month, 1961-1990, taken from Ashraf and Galor (2013). Data originally based on geospatial average monthly precipitation data for this period reported by the G-ECON project (Nordhaus, 2006).

Predicted genetic diversity. Predicted genetic diversity of the contemporary population, adjusted for post-Columbian migration flows and genetic distance between ethnic groups. See Ashraf and Galor (2013).

GDP per capita. Average annual GDP per capita over the period 2003 – 2012, in 2005US$. Source: World Bank Development Indicators.

Democracy index. Index that quanties the extent of institutionalized democracy, as reported in the Polity IV dataset. Average from 2003 to 2012.

Percentage at risk of malaria. The percentage of population in regions of high malaria risk (as of 1994), multiplied by the proportion of national cases involving the fatal species of the malaria pathogen, P. falciparum. This variable was originally constructed by Gallup et al. (2000) and is part of Columbia University's Earth Institute data set on malaria. Data taken from Ashraf and Galor (2013).

Percentage in (sub-)tropical zones. Percentage of area within a country which forms part of each of the tropical or sub-tropical climatic zones. Data taken from John Luke Gallup, http://www.pdx.edu/econ/jlgallup/country-geodata.

Life expectancy. Average life expectancy at birth, average from 2003 to 2012, taken from World Bank Development Indicators.

Gini coefficient. Average from 2003 to 2012, taken from World Bank Development Indicators.

Redistribution (% of GDP). Government transfers as a fraction of national income. Average from 2003 to 2012, taken from World Bank Development Indicators.

Ethnic and religious fractionalization. Indices due to Alesina et al. (2003) capturing the probability that two randomly selected individuals from the same country will be from different ethnic (religious) groups.

Linguistic diversity. Index due to Fearon (2003) capturing the linguistic diversity within a given country, taking into account the structural similarity of languages using a language tree.

Labor protection index. Index capturing the rigidity of employment laws by Botero et al. (2004). Includes data on employment, collective relations, and social security laws and measures legal worker protection.

Homicide rate. Numbers of intentional homicides per 100,000 people. Average 2003–2012, taken from World Bank Development Indicators.

Religion shares. Source: Barro (2003).