Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Global Evidence on Economic Preferences

Armin Falk, Benjamin Enke, Anke Becker, David Huffman en Uwe Sunde

DP 12/2016-051

Global Evidence on Economic Preferences

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

December 21, 2016

Abstract

This paper presents the Global Preference Survey (GPS), the first global survey focused on measuring a set of fundamental economic preferences: risk preference, time preference, positive and negative reciprocity, altruism, and trust. The sample includes 80,000 individuals, drawn as representative samples from 76 countries around the world, representing 90 percent of both the world's population and income. The paper shows that these preferences differ substantially across countries, but heterogeneity within countries is even more pronounced. The preferences vary systematically with plausibly exogenous individual characteristics – gender, cognitive ability, age, and cultural differences as captured by language structure – as well as country-level characteristics like geography. Preference differences are also correlated with differences in a wide range of individual-level outcomes, including savings decisions, labor market choices, and prosocial behaviors, and these relationships are similar across countries. Country-level preference differences are correlated with variation in important aggregate outcomes, ranging from economic development, to frequency of armed conflicts.

We thank the editor, Andrei Schleifer, and three anonymous referees. 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, SITE Experimental Economics, University of Virginia, WZB Berlin, Maastricht University, LISER (Luxemburg), IZA, Bonn, University of Copenhagen, University of Bologna, FU Berlin, LMU (Munich)and UC San Diego. Armin Falk acknowledges financial support from the European Research Council through ERC # 209214. Thomas Dohmen acknowledges financial support from the Netherlands Organization for Scientific Research through NWO Vidi grant 452-10-006, and from Netspar. Dohmen, Falk: University of Bonn, Department of Economics; tdohmen@uni-bonn.de, armin.falk@uni-bonn.de. Becker, Harvard University, Department of Human Evolutionary Biology: ankebecker@fas.harvard.edu. Enke, Harvard University, Department of Economics: enke@fas.harvard.edu. Huffman: University of Pittsburgh, Department of Economics; huffmand@pitt.edu. Sunde: University of Munich, Department of Economics; uwe.sunde@lmu.de.

JEL classification: D01; D03; F00
Keywords: Economic preferences; cultural variation.

Contents

1 Introduction

This paper presents evidence on how fundamental economic preferences are distributed around the world. The analysis uses the Global Preference Survey (GPS), a new global survey designed to measure a set of preferences that play a central role in economic theory. While economic models abstract away from many details of preferences, they explicitly model preferences over certain attributes – timing, risk, and implications for the payoffs of others – that are almost always relevant for the trade-offs involved in economic decisions. Accordingly, the GPS includes measures of risk preference, time preference, and three conceptually distinct types of social preferences: unconditional altruism, positive reciprocity, and negative reciprocity. The GPS also includes a novel measure of trust; although at least partly a belief rather than a preference, trust has also been argued to be fundamental for a wide range of economic transactions (e.g., Arrow, 1972).

The preference measures featured in the GPS were developed using a methodology grounded in economic theory. In the context of economic models, it is possible to specify an ideal choice situation, in which choices reveal preferences. This approach can reveal differences in preferences across individuals, as well as potentially across different cultures and populations, because it holds everything constant: stakes, probabilities, and relevant information conditions. As a way to approximate the ideal choice setting for revealing a given preference, economists have used incentivized choice experiments; there are now specific, widely used experimental measures of different key preferences. These types of measures are costly, however, to implement in a globally representative sample.[1] We therefore use a set of survey measures of preferences, which are lower cost, but were specifically developed to provide a good approximation to incentivized, revealed preference measures.

Specifically, the survey preference measures in the GPS were developed in an initial optimization exercise (for details see Falk et al., 2016). The exercise involved conducting multiple incentivized choice experiments for each preference, and testing the relative abilities of a wide range of different types of survey measures, to predict behavior in these choice experiments. The procedure lead to the selection

[1] For example, the measure should ideally involve large menus of choices, to give tight identification of preferences, but this is costly in terms of time. Also, to allow real choices, experiments should involve real stakes, but this is financially costly on a large scale. Data sets that contain experimental preference measures for several countries typically come from small- or medium-scale experiments and are based on student or other convenience samples (e.g., Vieider et al., 2015b,a; Wang et al., 2016). See also (Rieger et al., forthcoming) for a cross-country survey on risk preference among economics students.

of two survey items for each preference that had the best joint predictive power.[2] Combining each pair into a single measure yields a single measure for each preference is not only validated, in the sense of being related to experiments; it combines the best performing survey formats out of a menu of alternatives.

This paper first uses the GPS to provide descriptives on the nature of global variation in the set of preferences. For each of the preferences, we document substantial variation not just across individuals, but also across entire countries. This cross-country heterogeneity follows pronounced economic, geographic and cultural patterns. The various preference measures are also correlated, giving rise to distinct "preference profiles" of groups of countries: Patience and willingness to take risks are one pair of correlated preferences, and the pro-social traits of positive reciprocity, altruism, and trust form another grouping. Although between-country variation is substantial, within-country heterogeneity is even more pronounced.

Our next step is to investigate the relationship of the preferences to characteristics, which previous literatures have hypothesized might determine preferences. At the individual level we show that the preferences vary systematically with gender, cognitive ability, and age. Some relationships, such as between risk aversion and gender, go in the same direction in almost all cultures, whereas others, such as the age profile for patience, appear more variable across cultures. The preferences are also related to exogenous characteristics at the country level, including geographic and climatic features. Drilling deeper into potential determinants of preference heterogeneity, we show that the preferences vary systematically with cultural differences, as proxied by grammatical structure of language. People who speak languages for which the grammar does not require an explicit coding of the future(see Chen, 2013) are more patient, positively reciprocal, trusting, and altruistic, both across and within countries.

Turning to economic outcomes, we provide evidence that these are systematically related to preference variation. At the individual level, 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. These relationships of preferences with outcomes are qualitatively similar across almost all countries. This provides an additional, out of context check on the ability of the GPS measures to capture the same underlying traits across a wide range of cultures. Important country level outcomes are also re-

[2]The procedure also involved cross-validation and testing out-of-sample predictive power.

lated to cross-country variation in preferences. Greater patience is associated with higher GPD; risk aversion with stricter labor regulations; altruism with volunteering and donation as a fraction of GDP; negative reciprocity with greater frequency of armed conflicts.

The GPS complements other, existing global surveys, by providing a new set of measures. Existing surveys, such as the World Values Survey (WVS), include a wide range of valuable measures, but do not have questions designed to capture the set of preferences measured in the GPS. Some questions in these surveys might be able to serve as proxies for the preferences, e.g., because the question was designed to capture a trait from another field of study, which has a conceptual overlap with the notion of a preference in economic theory.[3] The challenge, however, is identifying good proxies and eliminating poor proxies: traditionally, researchers have had to rely on intuition and subjective judgments. The measures in the GPS have the advantage that they have known relationships to revealed preference benchmarks derived from theory.

The GPS can also enhance the value of existing surveys, by providing a new tool for discriminating between strong and weak preference proxies. To illustrate, at the end of the paper we explore possibilities for finding preference proxies in the WVS. For positive and negative reciprocity, we do not find any plausible candidate proxies. We do find a candidate proxy for risk preference, which turns out to be strongly correlated with the GPS risk preference measure, at the country level, and related to determinants and outcomes in a similar way, as well. This provides an indication that it is a strong proxy, and provides a way to study risk preference in the WVS set of countries. For altruism and time preference, by contrast, the candidate proxies turn out to have relatively weak relationships with the GPS measures, and economic outcomes, raising caveats about their ability to capture the relevant preferences. The exercise enhances the value of some measures in existing surveys, and also underlines that the GPS is contributing new information about the world population that is of particular interest to economists.

The findings in this paper are relevant for a large literature on how fundamental economic preferences are related to individual characteristics, such as gender, age, or cognitive ability. Previous work has typically using non-representative samples or representative samples from within particular countries (See, e.g., Barsky et al., 1997; Frederick, 2005; Dohmen et al., 2008; Croson and Gneezy, 2009; Dohmen et al., 2010, 2011). Some studies have compared how preferences are related to traits across specific societies, e.g., comparing gender differences in preferences in

[3]For example, the trait "value of stimulation" in psychology (Schwartz, 2012) could potentially be related to risk preference.

a matriarchal and a patriarchal society in India (Gneezy et al., 2009). The pervasiveness of such relationships across a broader range of cultures, however, and on a representative basis, has been an open question. Our results help add to knowledge in this area, shedding light on which relationships between preferences and traits are close to universal, and which are relatively culturally specific. Previous work has also hypothesized that geographic and climatic factors might have shaped preference endowments around the world (e.g., Galor and Özak, 2016). We provide new empirical evidence about how preferences vary with geography and climate.

Another contribution is to the literature on economic preferences and individual economic outcomes.[4] Our findings indicate that preference heterogeneity is related to outcomes in a similar way across a wide variety of cultures, something that has not been shown previously in a systematic way. This is both reassuring that the GPS preference measures are capturing something similar irrespective of cultural differences, and about the ability of the fundamentals of economic theory to help explain outcome variation across societies. In addition, the representative cross-country nature of our data permits an investigation of the relationships of preferences to aggregate economic and social outcomes across countries, which to date is uncharted territory.[5]

Evidence on cross-country variation in economic preferences is also relevant for literatures on cultural economics and political economy (Guiso et al., 2006; Fernández, 2011; Alesina and Giuliano, forthcoming; Giuliano and Nunn, 2013). Some of the research on determinants and implications of cultural variation at the macro level has 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., 2015; Galor and Özak, 2016), but has not studied the component of culture related to fundamental economic preferences, mainly due to a lack of representative, cross country data. Another strand of the culture and economics literature has measured some of these preferences, but focusing on small-scale societies or sub-populations existing within countries (Henrich et al., 2001, 2006, 2010; Apicella et al., 2014;

[4]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).

[5]An exception is the literature on the importance of trust, which has been possible due to trust measures in existing global surveys like the WVS (see, e.g., Knack and Keefer, 1997; Guiso et al., 2009; Algan and Cahuc, 2010).

Talhelm et al., 2014).

The GPS data are well suited for many potential research agendas, on the determinants and implications of preference variation. In the conclusion of the paper we discuss several examples of potentially fruitful directions for future research using the GPS. The remainder of the paper proceeds as follows. In Section 2 we give details on the GPS dataset, and present descriptives on global preference variation. Section 3 studies the relationship between the preferences and exogenous characteristics, at the individual and country levels. Section 4 investigates the relationships between preferences and economic outcomes. Section 5 directly compares the preference measures in the GPS to candidate preference proxies in the WVS. Section 6 concludes.

2 Dataset

2.1 General Data Characteristics

The GPS data were collected within the framework of the Gallup World Poll, a survey that includes representative population samples in a large number of countries, and asks about social and economic issues, on an annual basis. We added our survey preference measures to Gallup's 2012 World Poll questionnaire, for 76 countries; the result is the GPS data set. We discuss some noteworthy characteristics of the data 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.

One important feature of the GPS data is that it measures preferences for a nationally representative sample for each country. This means that it is possible to study how preferences vary within the population of a given country, and also to construct country level averages, shedding light on how the preferences of the representative agent vary across countries. The median sample size was 1,000 participants per country.[6] Respondents were selected through probability sampling; ex-post representativeness of the data can be achieved using weights provided by Gallup.[7] In total, we collected preference measures for more than 80,000 participants worldwide.

The 76 countries included in the GPS constitute a geographically and culturally diverse set of nations. They were chosen with the aim of providing a globally repre-

[6] 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.).

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

sentative sample. The collection of countries covers all continents, various cultures, and different levels of development. Specifically, it 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.

Another important feature of the GPS data is a standardized data collection protocol across countries, achieved through several steps. Before the 2012 World Poll, Gallup conducted pre-tests of the preference module in 22 countries of various cultural heritage. This was 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). For all countries, there was a translation of all survey items from the original language, to the local language, and back again in an iterative process; this is Gallup's regular translation scheme, to ensure comparable meaning of the questions across languages. Monetary values used in the survey questions were also calibrated according to median household income for each country, so as to hold monetary stakes constant.[8] Finally, most of the interviews for the World Poll 2012 took place using the same response mode across individuals and countries – face-to-face interviews – although in some exceptional cases telephone interviews were also used. Table XXX in Appendix XXX shows the countries included in the GPS, along with numbers of observations and the survey mode.

2.2 Preference Measures

For each preference, we combine survey items, using weights obtained from the initial optimization exercise. 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., 2016, 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

[8] 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.

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.[9] The quantitative items more closely resemble the choice-based experiment measures, but the qualitative items also have explanatory power.

Time Preference. Our measure of time preference 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 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*

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.

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

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.

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 Preference. 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 (**?**). 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*.[10] This last item captures prosocial punishment and hence a concept akin to norm enforcement. These three items receive weights of about one third each.

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.

[10] In the original survey design exercise, 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., 2016). However, in the pilot in 22 countries, 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.

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). The item was a strong predictor of trusting behavior in incentivized trust games, in the survey design stage. Time constraints determined the choice to have only one measure of trust, and also the fact that there already exists a global measure of trust in the WVS data set.

2.3 Further Variables of Interest

The Gallup World Poll includes a wide range of individual-level background variables 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. The data contain regional identifiers (usually at the state or province level), hence allowing for cross-regional analyses within countries. In the GPS survey module 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.

2.4 Descriptives

The analysis begins with an investigation of the heterogeneity of preferences around the world, as captured by the GPS measures. 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 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

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

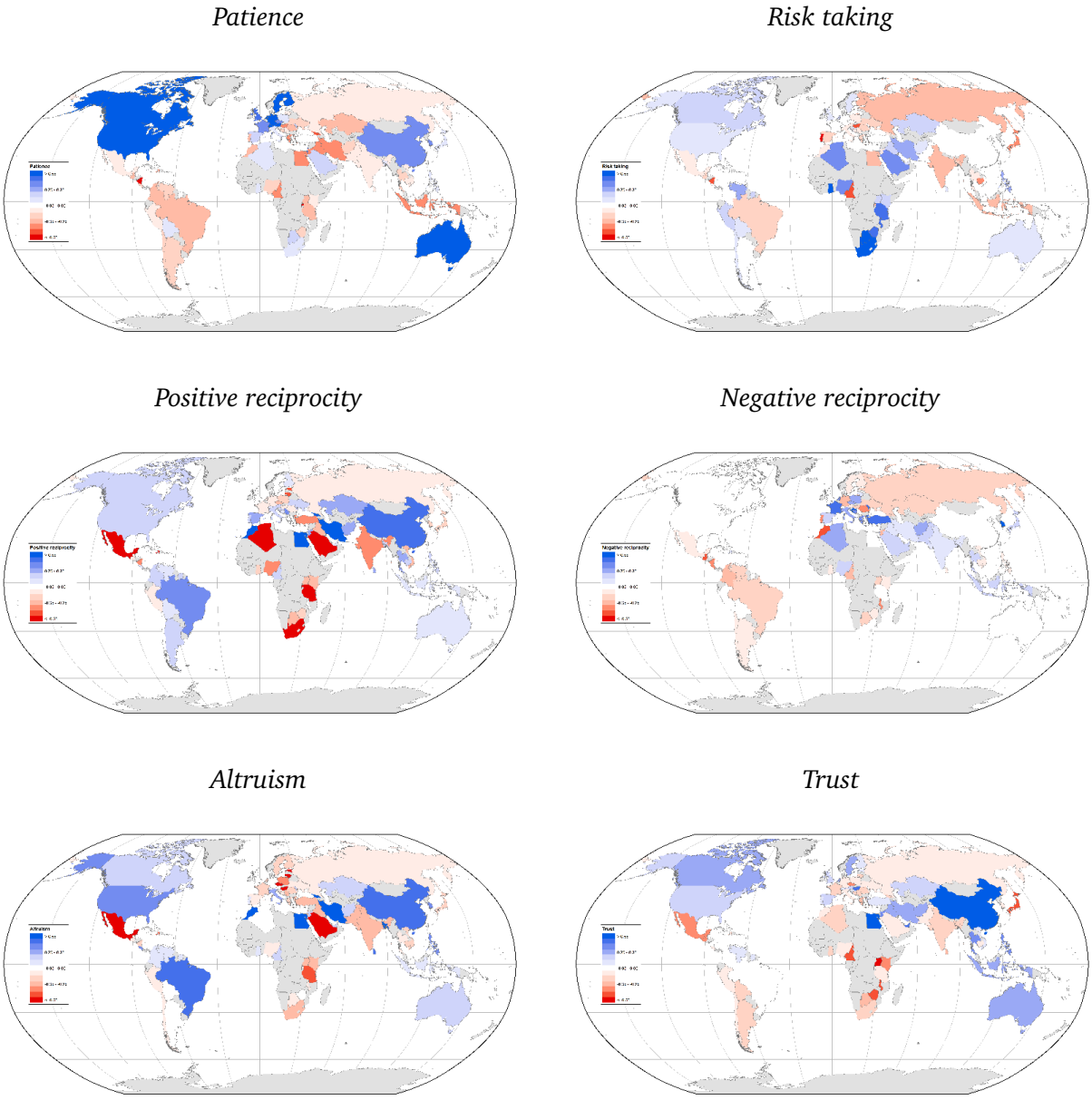

Figure 1: World maps of preferences. In each figure, white denotes the world average. Darker blue indicates higher values of a given trait, while darker red colors indicate lower values, all of which are measured in standard deviations from the world mean. Grey indicates missings.

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

[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.

preference.

A first observation is that 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.

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.

While individual preferences exhibit geographic variation, preferences might also be correlated amongst each other, giving rise to distinct country-level preference profiles. Table 2 shows Pearson correlations of preferences together with levels of significance.[13] 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.[14] 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 behav-

[13]The results are similar when computing Spearman rank correlations.

[14]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 quantitative risk and patience items is 0.19, while that between the two qualitative risk and patience items is 0.55.

ioral 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.[15] Despite being related to the social domain, negative reciprocity is not at all correlated with prosociality. Instead, it is positively correlated with patience. 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 variation in the world. Table 3 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.[16] 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 D 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 2: Pairwise correlations between preferences at country level

	Patience	Risk taking	Positive reciprocity	Negative 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[15] If the GPS survey item for trust measures mainly the belief-component of trust (as opposed to first-mover behavior in trust games, which is also affected by risk preferences), one would expect a low correlation between trust and risk preference.

[16] Interestingly, 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%.

Table 3: 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.

3 Determinants of Preferences

3.1 Preferences and Individual Level Characteristics

The pronounced within-country heterogeneity in preferences calls for a better understanding of individual-level preference variation. The following analysis investigates whether preference heterogeneity is related to three traits: age, gender and cognitive ability, taking self-reported math skills as a proxy for the latter.[17] These are interesting to study for two main reasons. First, they are associated with important differences in economic outcomes; if preferences vary with these traits, this could be part of the explanation for outcome differences.[18] Second, these traits are plausibly exogenous to preferences. Although the evidence is correlational, the previous literature has proposed various mechanisms, ranging from biological to purely social, through which gender, age, and cognitive ability might determine preferences.[19]

In many cases our findings converge with previous evidence, which is reassuring about the validity of the GPS preference measures. At the same time, the results from the GPS have a much broader scope than previous studies, which have mainly

[17] Subjects report math skills on a scale from 0 to 10. 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 for within-country analyses.

[18] See, 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 research relating preferences to these traits.

[19] See Croson and Gneezy (2009); Dohmen et al. (2011); Benjamin et al. (2013).

used individual countries, or non-representative samples. The GPS thus allows new insights into which relationships might reflect mechanisms that are more universal, and which might be specific to certain societies.

In Table 4 we present OLS regression estimates for how preferences are related to gender, cognitive ability, and age across the GPS sample. We report results with and without country fixed effects. The preference variable are standardized so coefficients are in units of standard deviations. In Appendix E we show that results are robust to adding an additional set of control variables.[20]

Starting with time preference, we see in Table 4 that women are less patient than men, on average across the world, but the difference is quite small. Patience is more pronounced among individuals with higher cognitive ability, and it varies with age, in a hump-shaped pattern: Middle aged individuals are the most patient, compared to the young and the elderly. The small gender difference is in line with previous cross-country surveys of college students.[21] Earlier studies have also found that higher cognitive ability goes with greater patience, but this has been documented in only a small set of countries, e.g., the US, Germany, and Chile.[22] There is little previous evidence, from cross-country or representative data, on how patience varies with age.

Turning to risk preference, Table 4 indicates that women are substantially more risk averse than men, by about a fifth of a standard deviation. Risk aversion is more pronounced for individuals with lower cognitive ability. The elderly are also significantly more risk averse than the young, on average around the world. The gender difference we find for risk aversion is qualitatively in line with the results of many previous studies, for particular countries or non-representative sub-populations.[23] Previous studies have also found a similar relationship between risk aversion and cognitive ability, for a few countries.[24] A similar shaped age profile in risk preference

[20] These include indicators for religious affiliation, physical health, and subjective safety perceptions. 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.

[21] See Wang et al. (2016) for results from a survey with college students across 45 countries.

[22] See, e.g., Frederick (2005), Dohmen et al. (2010), and Benjamin et al. (2013), respectively. Shamosh and Gray (2008) report a positive relationship between time preference and cognitive ability across studies, in a meta-analysis, but the analysis does not identify the set of countries from which the countries are drawn, and subject pools are non-representative.

[23] Vieider et al. (2015a) conduct experiments measuring risk preference in 30 countries, with student subjects, and find that female students are more risk averse than males, on average; the study does not compare gender differences across countries. Meta-analyses have found that, across studies, female subjects are on average more risk averse than males, but the underlying studies used non-representative samples, or were from only a small number of countries Croson and Gneezy (2009); Byrnes et al. (1999).

[24] Frederick (2005), Burks et al. (2009), Dohmen et al. (2010), and Benjamin et al. (2013) find similar results in the US, Germany, and Chile.

Table 4: 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)
1 if female	-0.045***	-0.056***	-0.19***	-0.17***	0.069***	0.049***	-0.14***	-0.13***	0.096***	0.10***	0.067***	0.066***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Subj. math skills	0.045***	0.028***	0.049***	0.046***	0.042***	0.038***	0.043***	0.040***	0.044***	0.044***	0.063***	0.056***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age	1.41***	0.72***	-0.24	-0.083	1.38***	1.02***	0.12	-0.36*	0.052	-0.0060	0.90***	0.37*
	(0.34)	(0.17)	(0.23)	(0.20)	(0.22)	(0.17)	(0.21)	(0.19)	(0.22)	(0.14)	(0.27)	(0.21)
Age squared	-1.64***	-1.45***	-1.18***	-1.20***	-1.39***	-1.17***	-0.88***	-0.45**	-0.11	0.015	-0.49*	0.032
	(0.34)	(0.20)	(0.24)	(0.21)	(0.26)	(0.18)	(0.20)	(0.18)	(0.23)	(0.15)	(0.27)	(0.20)
Constant	-0.46***	-0.37***	0.19**	0.21***	-0.54***	-0.079**	-0.018	0.37***	-0.28***	-0.064**	-0.64***	-0.078**
	(0.08)	(0.04)	(0.09)	(0.04)	(0.08)	(0.04)	(0.07)	(0.05)	(0.07)	(0.03)	(0.07)	(0.04)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	78501	78501	78445	78445	78869	78869	77521	77521	78632	78632	77814	77814
R^2	0.021	0.165	0.088	0.167	0.016	0.128	0.038	0.112	0.016	0.135	0.036	0.111

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

has been documented previously, for individual countries.[25]

Social preferences and trust also vary significantly with individual characteristics. Table 4 shows that positive reciprocity and altruism are more pronounced among women, while negative reciprocity is less strong. Positive reciprocity, altruism,and negative reciprocity are all positively related to cognitive ability. The estimates reveal that positive reciprocity has a hump-shaped relationship to age, negative reciprocity is declining with age, and altruism is not significantly related to age. The few previous cross-country studies relating social preferences to gender and age have mainly focused on students or other non-representative samples, and found varying results.[26] Some previous studies have also found a positive relationship between cognitive ability and altruism, using student subjects.[27]

Finally, the results on trust in Table 4 are broadly in line with evidence from the trust literature, although previous findings are somewhat mixed. We find that women are more trusting, as are the elderly, and higher cognitive ability is associated with higher trust.[28]

We turn next to a country-level analysis, to see whether the aggregate results in Table 4 reflect an underlying uniformity, or instead conceal heterogeneity across societies. For each country separately, we regress a given preference on age, age squared, gender, and cognitive ability. We then summarize the results in three figures. Figure 2 shows the gender coefficients for the different countries, with a separate panel for each preference. Figure 3 presents cognitive ability coefficients in a similar format. Because the relationships between some preferences and age is non-linear and cannot be summarized with a single coefficient, Figure 4 plots age profiles. Showing profiles for 76 countries in one graph is unwieldy, so the figure compares two groupings of countries, OECD members versus non-OECD; this di-

[25] E.g., Dohmen et al. (2011) show that willingness to take risks declines with age in a representative sample of German adults. Mata et al. (2016) show that a WVS measure of "value of stimulation," a trait that is potentially related to risk preference, declines with age, in a sample of 77 countries.

[26] Engel (2011) provides a meta-analysis of studies measuring altruism using dictator games, mainly for student subjects, across 35 countries. The analysis finds no gender difference in altruism, and a positive relationship between age and altruism, in contrast to our findings. Henrich et al. (2001) and Henrich et al. (2010) find no gender or age differences in various social preferences, across selected small-scale societies.

[27] Specifically, Chen et al. (2013) find a similar relationship between altruism and cognitive ability, among college student subjects in the United States.

[28] Alesina and La Ferrara (2002) used GSS data from the U.S., 1974 to 1994, and find a hump-shaped age profile for age as well as lower trust among women. They also find a positive correlation between trust and education. Delhey and Newton (2003) used Euromodule data from Germany, Hungary, Slovenia, South Korea, Spain, and Switzerland, between 1999 and 2001, and find no relationships of trust to gender, age, or education (except women less trusting than men in Switzerland), but this is clearly a much smaller set of countries. Another cross-country analysis, by Whiteley (1999), used 45 countries in the WVS from 1990 to 1993. He finds that the elderly are more trusting, consistent with our findings, while women are less trusting.

vision of countries captures some of the most salient cross-country differences or commonalities. See Appendix E.2 for age profiles at a more disaggregated level.

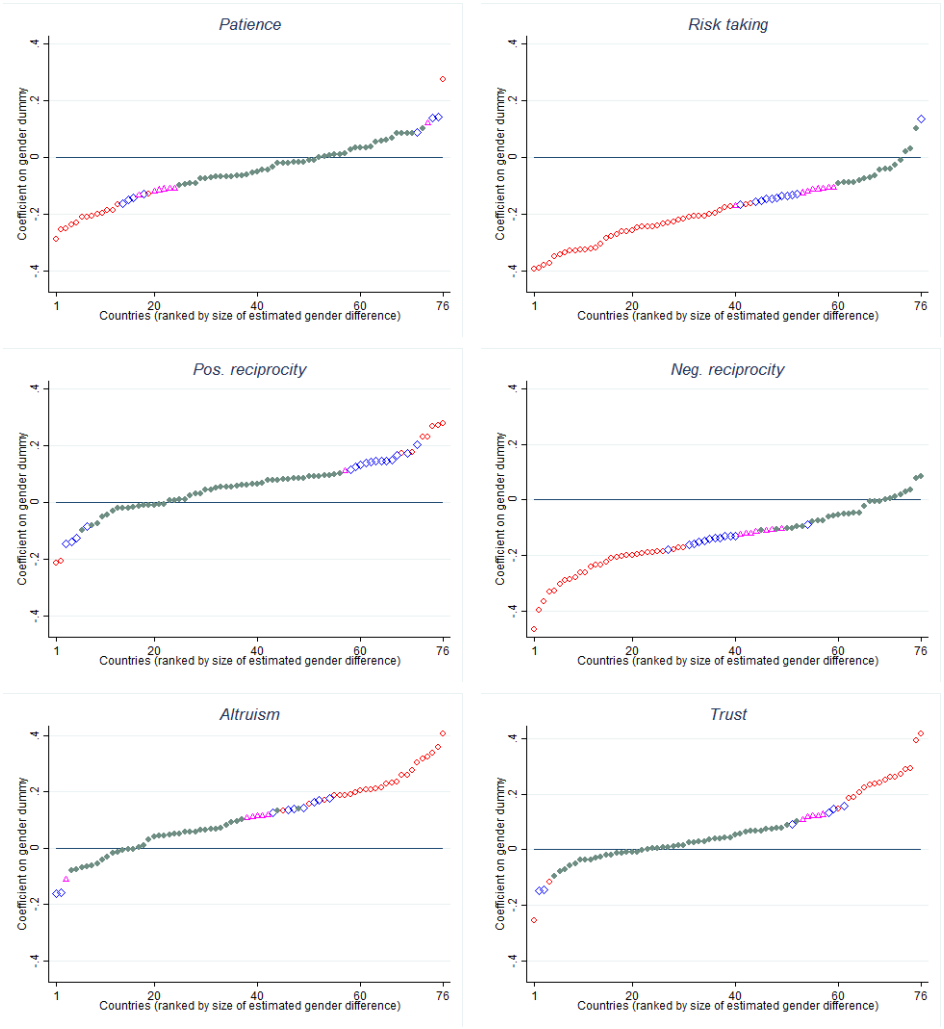

Figure 2: Gender correlations separately by country. Each panel plots the distribution of gender correlations. That is, for each country, we regress the respective preference on gender, age and its square, and subjective math skills, and plot the resulting gender coefficients as well as their significance 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 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.

Beginning with time preference, Figure 2 shows that the slightly larger degree of impatience among women, at the aggregate level, conceals substantial heterogeneity. Only about 68 percent of countries have a coefficient indicating greater impatience for women, and only about 30 percent have a statistically significant difference in that direction. Figure 3 indicates, by contrast, that the relationship between cognitive ability goes in the same direction, and is statistically significant, in almost all countries. This indicates that the relationship is relatively universal, and

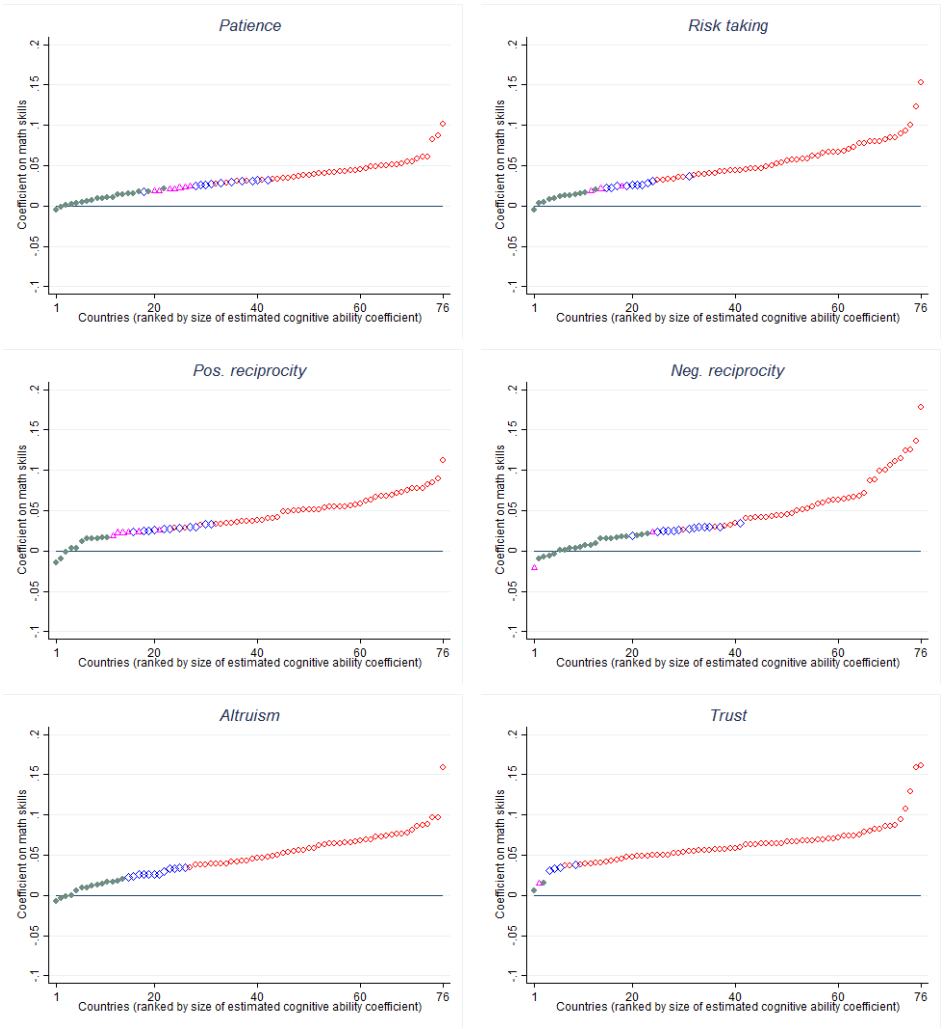

Figure 3: Cognitive ability correlations separately by country. Each panel plots the distribution of cognitive ability correlations. That is, for each country, we regress the respective preference on gender, age and its square, and subjective math skills, and plot the resulting math skill coefficients 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 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.

arguably not the product of specific educational systems or institutions. In Figure 4 we see that the hump-shaped age pattern for patience, observed in the aggregate, is actually only present for OECD-member countries; the profile is different, strictly declining, in non-OECD countries.

Turning to risk preference, Figure 2 reveals that in 95 percent of countries, the gender coefficient is non-zero and in the direction of greater risk aversion among women. Of these, 82 percent are statistically significant at least at the 10-percent level. This shows the relatively universality of the gender difference in risk preference, in qualitative terms, across a wide range of cultures and on a representative basis. Figure 3 shows that in almost all countries, lower cognitive ability is associated with significantly greater risk aversion. The age profiles in Figure 4 imply that risk aversion is increasing with age for both OECD and non-OECD countries. This similarity in age profiles is interesting given the diversity of historical experiences across countries, for different age groups.[29]

For positive reciprocity, some relationships to individual characteristics are more universal than others. While women are more positively reciprocal on average across the world, Figure 2 shows that this is statistically significant for only 26 percent of countries, so the difference is driven by a sub-set of societies. By contrast, In Figure 3, shows that positive reciprocity is associated with higher cognitive ability irrespective of culture. In terms of age profiles, Figure 4 reveals another difference across societies: The profile for positive reciprocity is hump-shaped for OECD countries, but less so for non-OECD countries.

Figure 2 shows that altruism, and negative reciprocity, are related to gender in opposite ways across countries, in line with the aggregate results. In most coun-

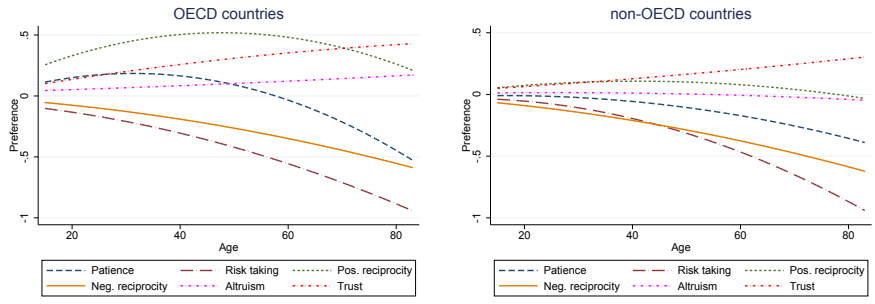

Figure 4: Age profiles by OECD membership. The figures depict the relationship between preferences and age conditional on country fixed effects, gender, and subjective math skills. These are augmented component plus residuals plots, in which the vertical axis represents the component of the preference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

[29]These results are in line with the finding of Mata et al. (2015), that age and gender differences in sensation seeking are qualitatively similar across 77 countries.

tries, altruism is more pronounced among women, whereas negative reciprocity is less pronounced. Altruism and negative reciprocity are both associated with higher cognitive ability in almost every country, as seen in Figure 3. Figure 4 indicates that altruism is weakly increasing with age for OECD countries, and largely flat for non-OECD, whereas negative reciprocity declines with age for both groups of countries
.

Finally, Figures 2, 3, and 4 show that the aggregate results on trust are largely born out in the data on individual countries. One exception is the positive relationship of trust to gender at the aggregate level; at the country level, women are more trusting than men in about 68 percent of countries, but this is statistically significant for only about 33 percent. Previous studies, conducted in different countries, have sometimes found that women are less trusting than men, perhaps reflecting this cultural specificity.[30] In almost all countries, trust is increasing with cognitive ability, and rust increases with age for both OECD and non-OECD countries.

In summary, some relationships between preferences and individual characteristics appear to reflect mechanisms that are relatively universal across a wide range of cultures. There are other relationships, however, like time preference and gender, or positive reciprocity and age, for which the qualitative relationships differ substantially across cultures. These latter findings point to cases where results from one culture might not generalize to other cultures, and where the underlying mechanisms are sensitive to cultural differences. These findings are made possible by the scope and representativeness of the GPS.

3.2 Preferences and Country-Level Characteristics

Section 2.4 showed that preferences vary systematically across countries. To further unpack the nature of this country-level variation, in this section we relate preferences to a set of important country characteristics: distance to the equator, longitude, fraction at risk of malaria, average temperature, average precipitation, and fraction living in the (sub-) tropics. One motivation for investigating the relationship of preferences to these particular features is that they are plausibly exogenous to preferences, and previous literatures have argued that such geographic and climatic characteristics could potentially have played a role in determining preference differences across countries.[31] These particular characteristics of countries are also

[30] See, e.g., Alesina and La Ferrara (2002) or Whiteley (1999).

[31] For example, Galor and Özak (2016) shows that initial differences in geography (suitability for agriculture) lead to the emergence of differences in Future Term Orientation as measured by the corresponding World Values Survey index, which in turn could cause future generations to have more patient time preferences.

of interest as they have been argued to serve as "deep" determinants for economic outcomes at the country level, including national income and development.[32] If geography and climate are correlated with preferences, this could conceivably be part of the reason why these country characteristics are related to aggregate economic outcomes.

Table 5 shows pairwise correlations of average preferences with each of the geographic and climatic variables. We focus on pairwise correlations because the geographic and climatic variables are in some cases highly correlated; in Appendix F we provide alternative results from multivariate regressions, in which we regress each preference on all of the geography and climate variables simultaneously. Beginning with time preference, the table shows that patience increases with distance from the equator, but is unrelated to longitude. Countries with warmer temperatures, and more precipitation, have lower levels of patience, as do countries with a higher risk of malaria and larger fractions of their populations living in the tropics. By contrast, willingness to take risks is not significantly related to geographic and climatic characteristics, except that countries with greater precipitation are more risk averse. Negative reciprocity has relationships to geography and climate that are similar to those for patience, with the exception that negative reciprocity also varies significantly with longitude. The pro-social traits of positive reciprocity are largely unrelated to geography and climate, although greater risk of malaria is associated with weaker positive reciprocity, and more precipitation is associated with greater altruism. Trust is related to geography and climate in a similar way to patience.

3.3 The Cultural Origins of Preference Variation

In this section we drill deeper into the origins of preference differences, exploring a potential role for culture. While there are myriad potential historical and cultural roots of preference differences, we focus on one particular proxy for cultural variation, i.e., grammatical structure of language. Language has been used as proxy for cultural variation in several previous studies.[33] In deriving specific testable hypotheses on the relationship between preferences and linguistic or cultural variation, we follow the work of Chen (2013). As discussed in detail by Chen (2013), some languages have a grammatical structure called strong Future Time Reference (FTR), requiring people to explicitly distinguish between present and future by making use of constructions such as "I *will* go to school tomorrow." Other languages have weak

[32] See Gallup et al. (1999); Diamond (2005); Spolaore and Wacziarg (2013); Olsson and Hibbs Jr (2005); Alsan (2015).

[33] See, e.g., Fearon (2003); Desmet et al. (2009, 2012); Spolaore and Wacziarg (2015).

Table 5: Pairwise correlations between preferences and geograhic and climatic variables

	Distance to equator	Longitude	Fraction at risk of malaria	Average temperature	Average precipitation	Fraction living in (sub-) tropics
Patience	0.487***	0.067	-0.298***	-0.475***	-0.204*	-0.347***
Risk taking	-0.189	-0.007	0.189	0.185	-0.234**	0.023
Positive reciprocity	0.132	0.119	-0.277**	-0.126	0.056	-0.094
Negative reciprocity	0.251**	0.279**	-0.150	-0.127	-0.228**	-0.272**
Altruism	-0.129	0.130	-0.133	0.105	0.192*	0.130
Trust	0.257**	0.165	-0.334***	-0.133	-0.178	-0.256**
Observations	76	76	75	75	75	75

Notes. Pairwise Pearson correlations between average preferences and other geographic and climatic variables at country level. See Appendix K for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

FTR, allowing speakers to talk about the future in present tense. 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.[34]

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 besides just time preference. 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. We investigate the relationship between average preferences and FTR at the country level, but also exploit within-country variation in preferences and FTR.

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.[35] We apply Chen's classification of languages to our dataset, which results in a set of 56 coded languages. In addition, we were able to code an additional 3 languages ourselves using the methodology outlined in Chen (2013).[36] In sum, we have access to 59 classified languages for a total of 75,224 respondents.[37] All results are robust to only making use of the languages coded in Chen (2013).

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

[35] 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 attenuation bias, which works against finding statistically significant effects in our analyses.

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

[37] We could not classify 23 languages, which are mostly spoken by small minorities (5,113 respondents in total). For such small languages, linguistic literatures are not well developed enough to establish an FTR classification. Chen attempted classification of 5 of these languages, but was also unsuccessful.

As a first step in the analysis, 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.[38]

Table 6 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.[39]

In a second step of the analysis, we exploit within-country variation in preferences and FTR. Such analyses are arguably even better suited to identify cultural origins of preferences because they can account for unobserved heterogeneity at the country level. Although we observe some variation in interview languages in many countries in our sample, only in Estonia, Nigeria, and Switzerland (2,925 respondents in total) do we observe within-country variation in FTR. We proceed by regressing individual-level preferences on a dummy for whether a respondent speaks a weak or strong FTR language, conditional on country fixed effects and age, age squared, gender, and our cognitive skills proxy. Columns (1), (3), (5), (7), (9), and (11) of Table 7 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.[40] We do not find significant relationships between FTR

[38] 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.

[39] Table 16 in Appendix G provides an additional robustness check, showing that the country-level results are similar when using OLS rather than WLS.

[40] 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 significant at the 10% level for patience and at the 1% level for positive reciprocity, trust, and altruism. In Table 17, we report the coefficient on FTR

Table 6: 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. See Appendix K for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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, subjective health and institutional quality. Despite this comprehensive set of covariates, and only exploiting within-region variation in FTR and preferences, we obtain almost identical results.[41]

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, our results highlight that the contemporary preference variation may have very deep historical roots,[42] and that the GPS data are well-suited to identify such effects.

4 Preferences and Outcomes

4.1 Preferences and Individual Outcomes

We now turn to investigating the correlations of preferences with individual behaviors and outcomes. Understanding the relationship between our preference measures and individual-level economic and social decisions is important in two respects. First, the resulting correlations provide insights into the potential role of heterogeneity in underlying preference parameters for explaining observed choice behavior, on a global scale. Second, it allows us to evaluate the meaningfulness and behavioral relevance of the GPS preference measures in a culturally and economically highly heterogeneous sample.[43] For each preference or set of preferences, we focus on outcomes that previous conceptual frameworks or models have identified as potentially determining that outcome. We consider the correlations between preferences and individual outcomes on average across the world, but also compare the relationships observed within different countries.

separately for each country in which we observe within-country variation.

[41] 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.

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

[43] 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.

Table 7: 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**
	(0.03)	(0.02)	(0.05)	(0.07)	(0.06)	(0.04)	(0.16)	(0.09)	(0.06)	(0.09)	(0.05)	(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
	(0.18)	(0.18)	(0.21)	(0.21)	(0.18)	(0.17)	(0.19)	(0.20)	(0.14)	(0.16)	(0.21)	(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
	(0.20)	(0.21)	(0.21)	(0.21)	(0.19)	(0.18)	(0.18)	(0.20)	(0.15)	(0.17)	(0.21)	(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)	(0.02)	(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)
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***
	(0.05)	(0.08)	(0.07)	(0.10)	(0.07)	(0.09)	(0.17)	(0.13)	(0.07)	(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
Additional controls	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. Additional controls include: log household income; subjective health status; subjective perception of strength of rule of law; indicators for religion. See Table 14 for the religion categories. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.1.1 Accumulation Decisions

We evaluate the relationship of the GPS patience measure to savings and investments in human capital, because in economic theory, time preference is a crucial determinant of such accumulation decisions.[44] Table 8 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, secondary, and tertiary education).[45]

In Appendix H.3, we show that the significant relationship between our patience variable and accumulation processes is not driven by only a few countries. Plotting the distribution of point estimates and their significance level across countries, we show that the coefficient of patience is positive in more than 90% of countries for both savings and education, and in most cases statistically significant. For instance, the correlation between patience and education is significant at least at the 5% level in 74% of all countries.

4.1.2 Risky Choices

We next investigate the relationship of risk preferences to important behaviors that have often been hypothesized to depend on taste for risk. Specifically, the career choice of being self-employed has been modeled as depending on sufficient willingness to take risks. Likewise, the risky health behavior of smoking has also often been hypothesized to depend on risk preference.[46] As columns (5) and (6) of Table 8 establish, our preference measure is related to 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

[44] See Acemoglu (2008) for an overview of the role of time preferences in determining accumulation and growth.

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

[46] See, e.g., Kihlstrom and Laffont (1979), and Viscusi and Hersch (2001).

Table 8: 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	0.038***	0.027***	0.069***	0.033***						
	(0.01)	(0.01)	(0.01)	(0.00)						
Risk taking					0.027***	0.024***	0.023***	0.019***	0.057***	0.032**
					(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)
Age		0.032		0.97***		1.56***		0.59***		2.56***
		(0.23)		(0.26)		(0.12)		(0.10)		(0.31)
Age squared		-0.21		-1.80***		-1.54***		-0.72***		-2.84***
		(0.23)		(0.25)		(0.12)		(0.11)		(0.30)
1 if female		0.00047		-0.016		-0.053***		-0.019***		-0.57***
		(0.01)		(0.01)		(0.01)		(0.00)		(0.03)
Subj. math skills		0.014***		0.042***		0.0058***		0.0028***		-0.011***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Constant	0.17***	-0.38***	1.32***	0.16**	0.036***	-0.37***	0.084***	0.094***	0.30***	0.067
	(0.00)	(0.09)	(0.00)	(0.07)	(0.00)	(0.04)	(0.00)	(0.03)	(0.00)	(0.11)
Country FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Region FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	15260	14459	79357	68409	72839	62125	57072	50687	15309	14490
R^2	0.073	0.183	0.205	0.359	0.064	0.137	0.108	0.167	0.028	0.229

OLS estimates, standard errors (clustered at country level) in parentheses. Saved last year is a binary indicator, while education level is measured in three categories (roughly elementary, secondary, and tertiary education). Self-employment and planned self-employment are binary, while smoking intensity is measured in three categories (never, occasionally, frequently). For the purposes of this table, age is divided by 100. Additional controls include log household income, and indicators for religious affiliation. See Appendix K for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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 set of additional covariates.

Appendix H.3 shows that the correlations between risk preferences and labor market or health decisions are qualitatively similar across countries. 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. These findings are in line with previous studies using data on risk preferences in particular countries to explain self-employment and health behaviors, but shows that the relationship is present in almost all cultures.[47]

4.1.3 Social Interactions

We analyze the relationships of the social preference measures to behaviors and outcomes in the social domain.[48] We focus on behaviors that correspond to unconditional giving, and behaviors that are linked to maintaining social relationships, as these types of outcomes have been hypothesized to depend on altruism, and reciprocity, respectively.[49]

Table 9 summarizes the results. Columns (1)-(8) show that altruism is significantly related to a broad range of giving 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.[50] 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. 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

[47] See, e.g., Dohmen et al. (2011), who analyze a representative sample of German adults and find that willingness to take risks is related to self-employment and to smoking.

[48] 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.

[49] See, e.g., Andreoni (1989) for theoretical work on altruism, and Fehr and Gächter (2002) and Rand et al. (2009) for discussions of how reciprocity may help sustain cooperative relationships.

[50] These baseline probabilities are 31.8%, 21.6%, 48.3%, and 23.7%, respectively (see Table 9 for the order of variables).

Table 9: 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	0.066***	0.059***	0.039***	0.038***	0.055***	0.052***	0.036***	0.032***	0.026***	0.023***	0.021***	0.017***	-0.0012	0.0032
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Positive reciprocity	0.0016	0.0049	0.0065**	0.0031	0.039***	0.034***	0.020***	0.020***	0.0011	-0.0016	0.018***	0.016***	0.012***	0.0085***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Negative reciprocity	-0.0052	-0.0024	0.00058	-0.00058	0.0053	-0.0024	0.0093***	0.0032	0.022***	0.017***	0.010**	0.0016	-0.011***	0.00041
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age		0.64***		0.45***		0.77***		0.26***		1.04***		-0.87***		5.55***
		(0.08)		(0.08)		(0.06)		(0.07)		(0.08)		(0.10)		(0.16)
Age squared		-0.49***		-0.49***		-0.94***		-0.29***		-1.03***		0.73***		-5.38***
		(0.08)		(0.08)		(0.07)		(0.08)		(0.08)		(0.09)		(0.17)
1 if female		0.011*		-0.018***		-0.018***		-0.0018		-0.046***		0.011***		-0.024***
		(0.01)		(0.01)		(0.01)		(0.00)		(0.01)		(0.00)		(0.01)
Subj. math skills		0.0091***		0.0077***		0.0082***		0.0071***		0.0092***		0.0050***		0.0022**
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Constant	0.42***	0.30***	0.19***	0.030	0.63***	0.33***	0.17***	-0.25***	0.20***	0.016	0.55***	0.32***	0.65***	-0.38***
	(0.00)	(0.04)	(0.00)	(0.04)	(0.00)	(0.04)	(0.00)	(0.04)	(0.00)	(0.03)	(0.00)	(0.03)	(0.00)	(0.05)
Country FE	Yes	No	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	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	58229	52686	58213	52677	55991	52473	56253	52812	55944	52421	65986	58479	77881	67420
R^2	0.178	0.241	0.084	0.138	0.083	0.148	0.112	0.179	0.047	0.105	0.093	0.170	0.052	0.237

OLS estimates, standard errors (clustered at country level) in parentheses. For the purposes of this table, age is divided by 100. Additional controls include log household income, and indicators for religious affiliation. See Appendix K for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(11) through (14) examine the relationship between social preferences and respondents' family and friendship relationships. 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.[51]

The overall pattern in Table 9 is that the social preference measures are related to a wide range of behaviors in the social domain. As Appendix H.3 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.

Tables 18 and 19 in the Appendix provide a robustness check by showing that the relationships between outcomes and the corresponding preferences, discussed above, remain similar when controlling for all other preferences simultaneously. For example, regressing savings on patience, but also risk preference, positive reciprocity, negative reciprocity, and altruism, patience is still significantly related to savings (and has a larger point estimate than other preferences).

In sum, all of the GPS preference measures are significantly related to a broad range of economic and social behaviors, in the expected directions based on conceptual frameworks or models. Although the results are correlational, they are consistent with preference heterogeneity being 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 this sense, the correlations provide an important out-of-context validation check for the survey module.

4.2 Preferences and Country-Level Outcomes

In this section we explore the relationship between preferences and outcomes at the country level that might potentially be endogenous to preferences. The results are correlations, so it is important to recognize that the outcomes could also potentially shape preferences, or co-evolve with certain preferences. We follow a similar approach to the analysis of individual level outcomes, analyzing the relationship of each outcome to a particular preference or set of preferences, guided by previous conceptual frameworks or models of how preferences might determine aggregate outcomes.

[51] Also see Dohmen et al. (2009), for similar findings in the adult population of Germany.

4.2.1 Economic development

We begin by investigating whether variation in the level of economic development across countries is related to variation in time preference, as the latter plays a central role in many models of economic development. Patience promotes higher national income in such models, through the channel of accumulation of physical and human capital.[52] Columns (1) to (4) of Table 10 regresses different measures of development on the GPS patience measure and controls. Controls include geographic and climatic variables that have been shown to help explain economic growth, and which were shown to be correlated with patience and other preferences in Section 3.2. Columns (1) and (2) show that average patience in a country is strongly and significantly correlated with GDP per capita. A one standard deviation increase in patience is associated with a 32% increase in log GDP per capita relative to the mean. Although our main focus is investigating possible mechanisms underlying growth, rather than maximizing predictive power, it is noteworthy that patience contributes substantially to explained variation, above and beyond the standard geographic variables. This can be seen comparing the R^2 with preferences included, to R^2 from a regression on controls alone (reported in the bottom row of the table). Adding time preference increases explained variation by 15 percentage points, or about 26%. Columns (3) and (4) show that patience is also significantly related to a broader indicator of development, the Human Development Index.

4.2.2 Riskiness of institutional arrangements and environment

We turn next to studying the social and economic correlates of risk taking. Intuitively, higher risk taking might shape – and be shaped by – the overall riskiness of the economic, institutional, and health environment. For example, from the perspective of a veil of ignorance, highly risk averse societies might implement tighter labor market regulations, such as stronger unemployment protection, as this reduces the risk of unemployment and poverty.[53] Greater aversion to risk might lead societies to mitigate a range of other possibilities for risk in the environment, ranging from physical safety to financial security. Columns (5) and (6) of Table 10 show that an index of labor market regulation is significantly related to the average degree of risk aversion. Thus, more risk averse countries have stricter labor market regulations. The increase in R^2 is also substantial, about 26%. Columns (7) and (8)

[52] For example, in a Ramsey-Cass-Koopmans model, time preference drives savings and accumulation decisions at the micro-level, which translate into higher levels of income at the national level.

[53] See, e.g., Bertola (2004) and Blanchard and Tirole (2008) for models in which risk aversion among workers makes it optimal to have labor market institutions characterized by employment protection and unemployment insurance.

relate risk preference to an index of several factors that make life risky, specifically prevalence of HIV, percent living in poverty, incidence of homicides and traffic death rate. Greater tolerance for risks at the country level is associated with more risky life environment, but the relationship becomes smaller and no longer statistically significant once we add geographic and climatic controls.

4.2.3 Charitable activities and conflict

Finally, we explore the correlations of selected aggregate outcomes with social preferences. Previous literatures have often hypothesized that altruistic preferences play a role in the extent to which a society is characterized by charitable activities. Social preferences, particularly negative reciprocity, have also been hypothesized to play a role in fostering anti-social conflicts.[54] Consistent with a role for altruism in country-level charitable activities, columns (9) and (10) show that countries with a higher average degree of altruism have a larger dollar value of charitable donations and volunteering activities, as a fraction of GDP.[55] Columns (11) and (12) relate the log frequency of armed conflicts to social preferences. Countries with a higher degree of negative reciprocity have experienced significantly more armed conflicts. As a robustness check, Table 22 in Appendix I regresses each of the aggregate outcomes we consider on all preferences simultaneously as well as trust. The results are similar to those obtained from regressing outcomes on only a single preference or subset of preferences.[56]

5 The GPS in comparison to existing global surveys

As the first global survey focused on measuring time preference, risk preference, positive reciprocity, negative reciprocity, and altruism, the GPS complements other, existing global surveys. One way that the GPS adds value is by providing a new

[54] See, e.g., Choi and Bowles (2007) and Herrmann et al. (2008).

[55] Altruism is highly correlated with positive reciprocity at the country level. Indeed, a principal components analysis of social preferences at the country level identifies two principal components, one that loads on altruism and positive reciprocity, and another that loads on negative reciprocity. Regressing volunteering and donation as a fraction of GDP on these two principal components, and controls, the coefficient for the prosocial principal component is still positive, but it is not statistically significant.

[56] One exception concerns willingness to take risks: the relationships to strength of labor regulations, and environmental risk, are still in the same directions, but now the former is not statistically significant with geographic controls, whereas the later becomes statistically significant. Another exception is altruism and volunteering and donation as a fraction of GDP; the relationship is still positive but no longer statistically significant.

Table 10: Country-level outcomes and preferences

	Dependent variables:											
	Ln[GDP p/c ppp]		Human Development index		Labor protection index		Life risk index		Volunteering and donation as fraction of GDP		Ln[Conflicts]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Patience	2.662***	1.914***	0.229***	0.137***								
	(0.261)	(0.257)	(0.0277)	(0.0227)								
Willingness to take risks					-0.238***	-0.173*	2.193***	1.031				
					(0.0827)	(0.0889)	(0.646)	(0.669)				
Positive reciprocity									-0.0175	-1.131	0.937**	0.170
									(0.967)	(0.893)	(0.464)	(0.509)
Negative reciprocity									0.925	0.802	1.607***	0.911**
									(1.058)	(1.040)	(0.412)	(0.398)
Altruism									0.638	2.001**	-0.816**	-0.170
									(0.878)	(0.856)	(0.370)	(0.443)
Constant	8.309***	8.016***	0.703***	0.693***	0.492***	0.280	-0.00263	1.138	1.743***	-0.755	1.825***	1.375
	(0.140)	(0.908)	(0.0137)	(0.0891)	(0.0234)	(0.242)	(0.162)	(1.070)	(0.256)	(2.020)	(0.129)	(1.142)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	76	75	76	75	56	56	67	66	32	32	75	75
R^2	0.397	0.726	0.335	0.744	0.135	0.245	0.203	0.661	0.053	0.404	0.169	0.362
R^2 with controls only		0.576		0.657		0.195		0.623		0.276		0.324

OLS estimates, robust standard errors in parentheses. Labor protection is measured by the LaPorta labor protection index. The life risk index is the principal component from a pca analysis of average life expectancy, log HIV prevalence, and traffic death rate. See Appendix K for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

benchmark for validating candidate proxies for these preferences in existing surveys.

Questions in existing surveys, although designed for other purposes, could turn out to be positively correlated with the measures in the GPS, suggesting that they may be a good proxy for these preferences. This could arise due to happenstance, or because a question was designed to measure a trait studied in another discipline, which has some conceptual overlap with the notion of preferences as defined in economic theory.

In this section we describe such a validation exercise for measures in the World Values Survey (WVS). We focus our analysis on the WVS as it is the main non-commercial, global survey, besides the GPS, designed to measure individual traits and attitudes. It covers an overlapping, but different set of countries from the GPS. Thus, identifying strong preferences proxies in the WVS would allow extending the study of fundamental economic preferences to even more countries.

To select candidate preference proxies, we searched all questionnaires of the WVS. We looked for key-words, and types of contexts and trade-offs, that seemed plausibly related to a respective preference. This initial identification of candidate preferences measures was necessarily based on intuition. We did not find any WVS questions that asked about something that seemed related to positive or negative reciprocity. We were able to identify measures that might possibly proxy for the other preferences, with varying degrees of plausibility.

The question we found in the WVS that seems most closely related to time preference is an item designed to capture "Long Term Orientation." Long Term Orientation (LTO) is defined by Hofstede et al. (2010) as a "cultural value that stands for the fostering of virtues oriented toward future rewards, perseverance and thrift." Specifically, the survey asks: "Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important?" The LTO variable is coded as 1 if the individual lists "thrift, saving money and things," regardless of what other qualities the respondent lists. Importantly, because long term orientation is specifically about fostering patience in children, rather than being a patient individual, it is not clear whether LTO will be a good measure of respondent patience. For example, if respondents who are impatient desire to encourage their children to be more patient than themselves, the measure could actually be inversely related to patience of the respondent on average.

For risk preference, we identified one plausible proxy, which asks the respondent to judge their similarity with a hypothetical person described as follows: "Adventure and taking risks are important to this person; to have an exciting life." This

WVS question was derived from the Schwarz Values Survey (Schwartz, 2012), and designed to capture a universal "value of stimulation." This value is defined in a similar way to the trait of "sensation seeking" in psychology (Zuckerman, 2007). The measure has been used previously as a proxy for tendency to engage in risky behaviors; Mata et al. (2016) state: "We take this item to measure the closely linked constructs of propensity for risk taking and sensation seeking." We hypothesized that the measure could also be related to risk preference, as defined in economic theory.

We also selected the WVS survey item that we judged to come closest to capturing altruism. The question asks respondents how similar they are to a hypothetical person for whom: "It is important for this person to do something for the good of society." This measure might be an imperfect as a proxy for altruism, however, because the broad framing in terms of "good of society" might or might not overlap with a more narrow notion of being kind to individuals. For example, there are circumstances where helping an individual might not benefit society as a whole.

The WVS also includes a well-known measure of trust, which we included for comparison to the GPS trust measure. The WVS measure asks whether the respondent thinks "most people can be trusted" or whether they would rather say that "you can't be too careful."[57] The GPS measure asks whether someone believes that others typically have "good intentions." The GPS trust measure was selected based on ability to predict trusting behavior in incentivized one-shot trust games. Evidence is mixed on whether or not the WVS trust measure correlates with trusting behavior in such games.[58]

Table 11 reports the correlations, at the country level, of the GPS measures and corresponding candidate measures from the WVS. The table shows that the measure of valuing stimulation in the WVS has a substantial correlation with the GPS risk preference measure. The WVS and GPS trust measures are also strongly correlated. By contrast, the candidate measure of altruism in the WVS has only a modest correlation with the GPS altruism measure, and the correlation is not statistically significant. The LTO measure is essentially uncorrelated with the GPS patience measure.

If the WVS candidate proxies capture preferences, one might also expect them to be related to determinants, and economic outcomes, in a similar way to the measures in the GPS. Tables 23 through 25 in Appendix J explore these relationships.

[57] A more recent wave of the WVS also includes a general self-assessment of one's inclination to trust others. Since the overlap between the countries of that specific WVS wave and the countries included in the GPS is only 15, we do not include it in our analysis below.

[58] See, e.g., Glaeser et al. (2000); Fehr et al. (2003).

Table 11: Relationships between preference proxies in the WVS and GPS measures

	Spearman's rho	p-Value	Obs.
WVS Long Term Orientation (correlation with GPS patience)	0.0912	0.4881	60
WVS Value of stimulation (correlation with GPS risk)	0.3042	0.0376	47
WVS Altruism (correlation with GPS altruism)	0.1982	0.2612	34
WVS Trust (correlation with GPS trust)	0.4843	0.0001	60

Spearman correlations. Responses to WVS questions asked in multiple waves are averaged across waves.

For the WVS value of stimulation measure, and the trust measure, we find that the relationships to determinants and outcomes are broadly similar to those obtained with the GPS risk and trust measures. For example, there is a statistically significant gender difference, and age profile, for value of stimulation, similar to those observed for the risk preference measure in the GPS. The value of stimulation is also related to self-employment at the individual level, and risk indexes at the aggregate levels, in the same was as the GPS measure, although the country-level relationships are not statistically significant.

For the candidate altruism and time preference proxies in the WVS, by contrast, the variation with determinants and outcomes is different from the corresponding GPS measures. For example, LTO has the opposite relationship to educational attainment at the individual level, and GDP at the country level, compared to the GPS time preference measure. The candidate altruism proxy from the WVS is not significantly related to volunteering and donation as a fraction of GDP, at the country level, although this could reflect small sample size.[59]

In summary, the validation exercise illustrates the value of the GPS data in two ways. First, it shows how the GPS can make existing surveys even more valuable, by offering a new interpretation of the measures. Second, the exercise shows that for several key preferences – time preference, altruism, negative reciprocity, and positive reciprocity – the WVS does not include strong proxies, at least compared to the GPS benchmark. Thus, the GPS measures are adding new information about the world population, above and beyond the traits and values measured in the WVS and other similar surveys.[60]

[59]The individual level outcomes in the social domain, analyzed in Section 4.1, are not available in the WVS, so we do not have a comparable analysis for the WVS altruism measure and economic outcomes. At the country level, there are only 17 WVS countries with measures of altruism and also measures of volunteering and donation as a fraction of GDP.

[60]There are various regional surveys, including the Barometer surveys of different world regions, and the European Values Survey, which have similar features to the WVS. The former mainly contain various measures of trust, whereas the latter is basically a regional version of the WVS, and thus

6 Conclusion

Many theories of human behavior, in economics and also 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 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 measures that were optimized against a theoretically derived benchmark. 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

includes similar measures to the ones that we analyze in the WVS. These surveys have a more limited geographic coverage than the WVS. Past waves of the Gallup World Poll may also include useful preference proxies.

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 more detailed investigation of the link between aggregate outcomes and preferences at the country level. 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. One example concerns time preference; a large body of dynamic theories of comparative development and growth highlight the crucial role of time preference for aggregate income and growth through accumulation processes. Consistent with such theories, in a follow-up paper, Dohmen et al. (2016), we show that patience is not only predictive of GDP per capita, but also of historical growth rates, and of variation in income across regions within countries. Furthermore, patience is correlated with the accumulation channels specified in growth models, including savings rates, investment in education, and investment in the stock of ideas and knowledge. A second example involves the sub-component of negative reciprocity that captures willingness to punish in response to seeing someone else harmed; such third-party punishment has been hypothesized to be a key contributing factor in the human ability to sustain large scale cooperation (See, e.g., Fehr and Gächter, 2002; Henrich et al., 2006). An initial analysis of the third-party punishment component of the GPS negative reciprocity measure reveals an intriguing finding: a strong positive correlation between willingness to punish third parties and GDP per capita, controlling for other preferences. Third party punishment is also correlated, however, with frequency of armed conflicts. Exploring these dual roles of negative reciprocity is a potentially fascinating future use for the GPS data.

References

Acemoglu, Daron, *Introduction to Modern Economic Growth*, Princeton University Press, 2008.

Alesina, Alberto and Eliana La Ferrara, "Who Trusts Others?," *Journal of Public Economics*, 2002, *85* (2), 207–234.

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

_ , _ , **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*, 2015, *13* (4), 599–630.

Algan, Yann and Pierre Cahuc, "Inherited Trust and Growth," *American Economic Review*, 2010, *100* (5), 2060–2092.

Alsan, Marcella, "The Effect of the TseTse Fly on African Development," *American Economic Review*, 2015, *105* (1), 382–410.

Andreoni, James, "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 1989, pp. 1447–1458.

Apicella, Coren L, Eduardo M Azevedo, Nicholas A Christakis, and James H Fowler, "Evolutionary Origins of the Endowment Effect: Evidence from Hunter-Gatherers," *American Economic Review*, 2014, *104* (6), 1793–1805.

Arrow, Kenneth J, "Gifts and Exchanges," *Philosophy & Public Affairs*, 1972, *1* (4), 343–362.

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.

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 and Empirical Properties of Measures of Risk Preferences," *Working Paper*, 2015.

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.

Bertola, Giuseppe, "A Pure Theory of Job Security and Labour Income Risk," *Review of Economic Studies*, 2004, *71* (1), 43–61.

Blanchard, Olivier J. and Jean Tirole, "The Joint Design of Unemployment Insurance and Employment Protection: A First Pass," *Journal of the European Economic Association*, 2008, *6* (1), 45–77.

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.

Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini, "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment," *Proceedings of the National Academy of Sciences*, 2009, *106* (19), 7745–7750.

Byrnes, James P., David C. Miller, and William D. Schafer, "Gender Differences in Risk Taking: A Meta-Analysis.," *Psychological Bulletin*, 1999, *125* (3), 367.

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, Chia-Ching, I-Ming Chiu, John Smith, and Tetsuji Yamada, "Too Smart to be Selfish? Measures of Cognitive Ability, Social Preferences, and Consistency," *Journal of Economic Behavior and Organization*, 2013, *90*, 112–122.

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.

Choi, Jung-Kyoo and Samuel Bowles, "The Coevolution of Parochial Altruism and War," *Science*, 2007, *318* (5850), 636–640.

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.

Delhey, Jan and Kenneth Newton, "Who Trusts?: The Origins of Social Trust in Seven Societies," *European Societies*, 2003, *5* (2), 93–137.

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.

Diamond, Jared M., *Guns, Germs, and Steel*, National Geographic, 2005.

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*, 2016.

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, pp. 219–260.

Engel, Christoph, "Dictator Games: A Meta Study," *Experimental Economics*, 2011, *14* (4), 583–610.

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," *IZA Discussion Paper 9674*, 2016.

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

Fehr, Ernst and Simon Gächter, "Altruistic Punishment in Humans," *Nature*, 2002, *415* (6868), 137–140.

_ , **Urs Fischbacher, Bernhard von Rosenbladt, Jürgen Schupp, and Gert G. Wagner**, "A Nation-Wide Laboratory: Examining Trust and Trustworthiness by Integrating Behavioral Experiments into Representative Surveys," 2003. Mimeo, Institute for Empirical Research in Economics, University of Zurich.

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. and Jeffrey D. Sachs, "The Economic Burden of Malaria," *American Journal of Tropical Medicine and Hygiene*, 2001, *64* (1-2 Suppl), 85–96.

Gallup, John Luke, Jeffrey D. Sachs, and Andrew D. Mellinger, "Geography and Economic Development," *International Regional Science Review*, 1999, *22* (2), 179–232.

Galor, Oded and Ömer Özak, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, *106* (10), 3064â–3103.

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.

Glaeser, Edward L, David I Laibson, Jose A Scheinkman, and Christine L Soutter, "Measuring Trust," *Quarterly Journal of Economics*, 2000, *115* (3), 811–846.

Gneezy, Uri, Kenneth L. Leonard, and John A. List, "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society," *Econometrica*, 2009, *77* (5), 1637–1664.

Golsteyn, Bart H.H., Hans Grönqvist, and Lena Lindahl, "Adolescent Time Preferences Predict Lifetime Outcomes," *Economic Journal*, 2014, *120* (580), F739–F761.

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, Jean Ensminger, Richard McElreath, Abigail Barr, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins

Gwako, Natalie Henrich et al., "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment," *Science*, 2010, *327* (5972), 1480–1484.

_ , **Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich et al.**, "Costly Punishment across Human Societies," *Science*, 2006, *312* (5781), 1767–1770.

_ , **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.

Kihlstrom, Richard E and Jean-Jacques Laffont, "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion," *Journal of Political Economy*, 1979, pp. 719–748.

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, *112* (4), 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.

Mata, Rui, Anika K: Josef, and Ralph Hertwig, "Propensity for Risk Taking Across the Life Span and Around the Globe," *Psychological Science*, 2016, *27*, 231–243.

Meier, Stephan and Charles Sprenger, "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2010, *2* (1), 193–210.

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.

Olsson, Ola and Douglas A. Hibbs Jr, "Biogeography and Long-Run Economic Development," *European Economic Review*, 2005, *49* (4), 909–938.

Rand, David G, Anna Dreber, Tore Ellingsen, Drew Fudenberg, and Martin A Nowak, "Positive Interactions Promote Public Cooperation," *Science*, 2009, *325* (5945), 1272–1275.

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

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

Salamon, LM, "Global Civil Society: Dimensions of the Nonprofit Sector (Vol. 2). The John Hopkins University Institute for Policy Studies," *Center for Civil Society Studies. Bloomfield, CT: Kumarian Press, Inc*, 2004.

Schwartz, Shalom H, "An Overview of the Schwartz Theory of Basic Values," *Online readings in Psychology and Culture*, 2012, *2* (1), 11.

Shamosh, Noah A and Jeremy R Gray, "Delay Discounting and Intelligence: A Meta-Analysis," *Intelligence*, 2008, *36* (4), 289–305.

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, "How Deep are the Roots of Economic Development?," *Journal of Economic Literature*, 2013, *51* (2), 325–369.

_ **and** _ , "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.

Talhelm, Thomas, X Zhang, Shige Oishi, Chen Shimin, D Duan, X Lan, and S Kitayama, "Large-Scale Psychological Differences within China Explained by Rice versus Wheat Agriculture," *Science*, 2014, *344* (6184), 603–608.

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*, 2015, *13* (3), 421–452.

Vieider, Ferdinand M., 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.

Viscusi, W Kip and Joni Hersch, "Cigarette Smokers as Job Risk Takers," *Review of Economics and Statistics*, 2001, *83* (2), 269–280.

Wang, Mei, Marc Oliver Rieger, and Thorsten Hens, "How Time Preferences Differ: Evidence from 53 Countries," *Journal of Economic Psychology*, 2016, *52*, 115–135.

Whiteley, Paul F, "The Origins of Social Capital," *Social Capital and European Democracy*, 1999, pp. 25–44.

Zuckerman, Marvin, *Sensation Seeking and Risky Behavior*, American Psychological Association, 2007.

APPENDIX

A Construction of the 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 Survey Optimization Exercise

To maximize the behavioral validity of the preference measures, subject to constraints of necessary brevity, all underlying survey items were selected through an initial (constrained) optimization procedure (see Falk et al. (2016) 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 Cross-Cultural Pilot 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 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). 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.

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 Countries Included in the GPS and Selection Criteria

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.[61]

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.[62] 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.

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

[62] 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 5.

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*

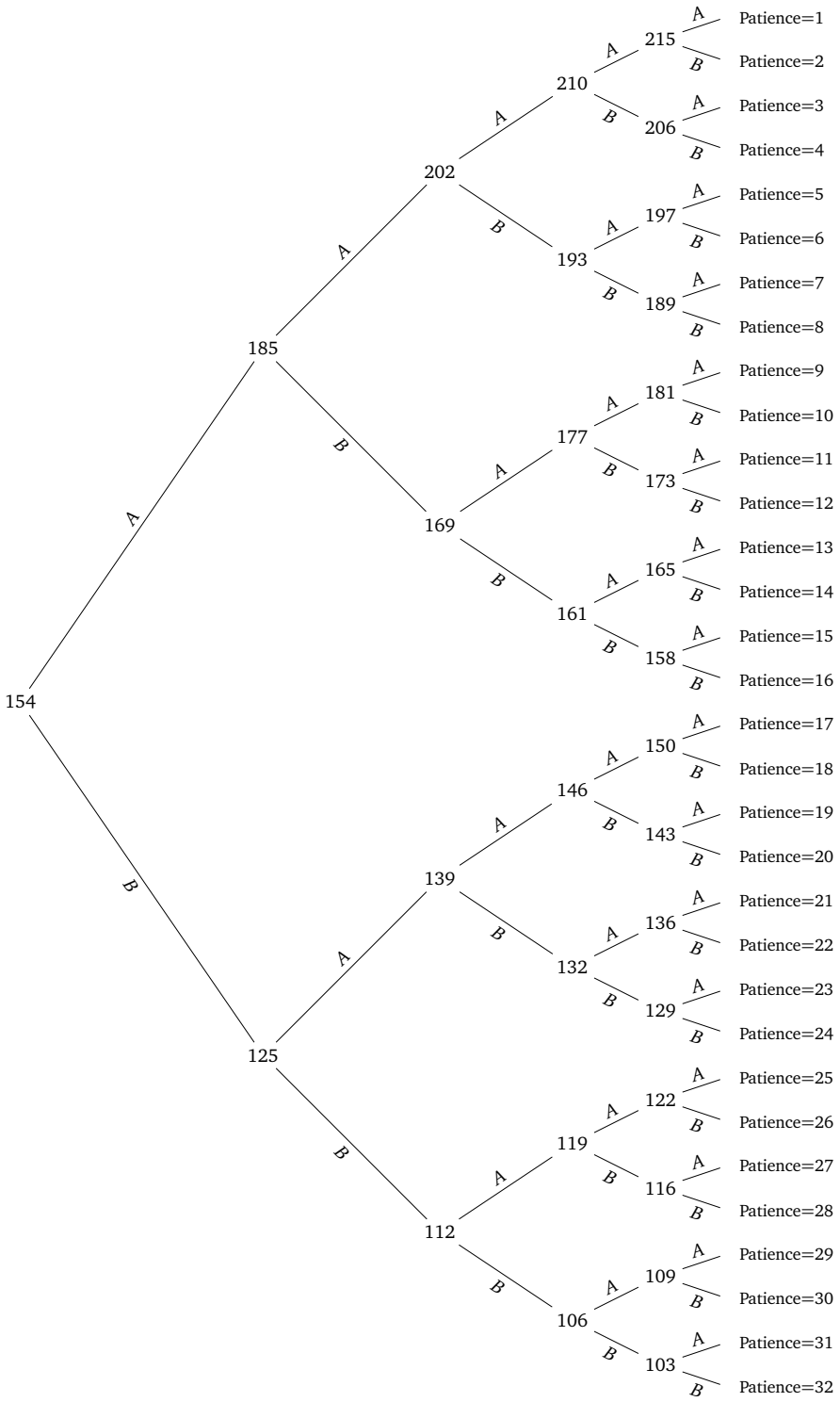

Figure 5: Tree for the staircase time task (numbers = payment in 12 months, A = choice of "100 euros today", B = choice of "x euros in 12 months". The staircase procedure worked as follows. First, each respondent was asked whether they would prefer to receive 100 euros today or 154 euros in 12 months 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?*

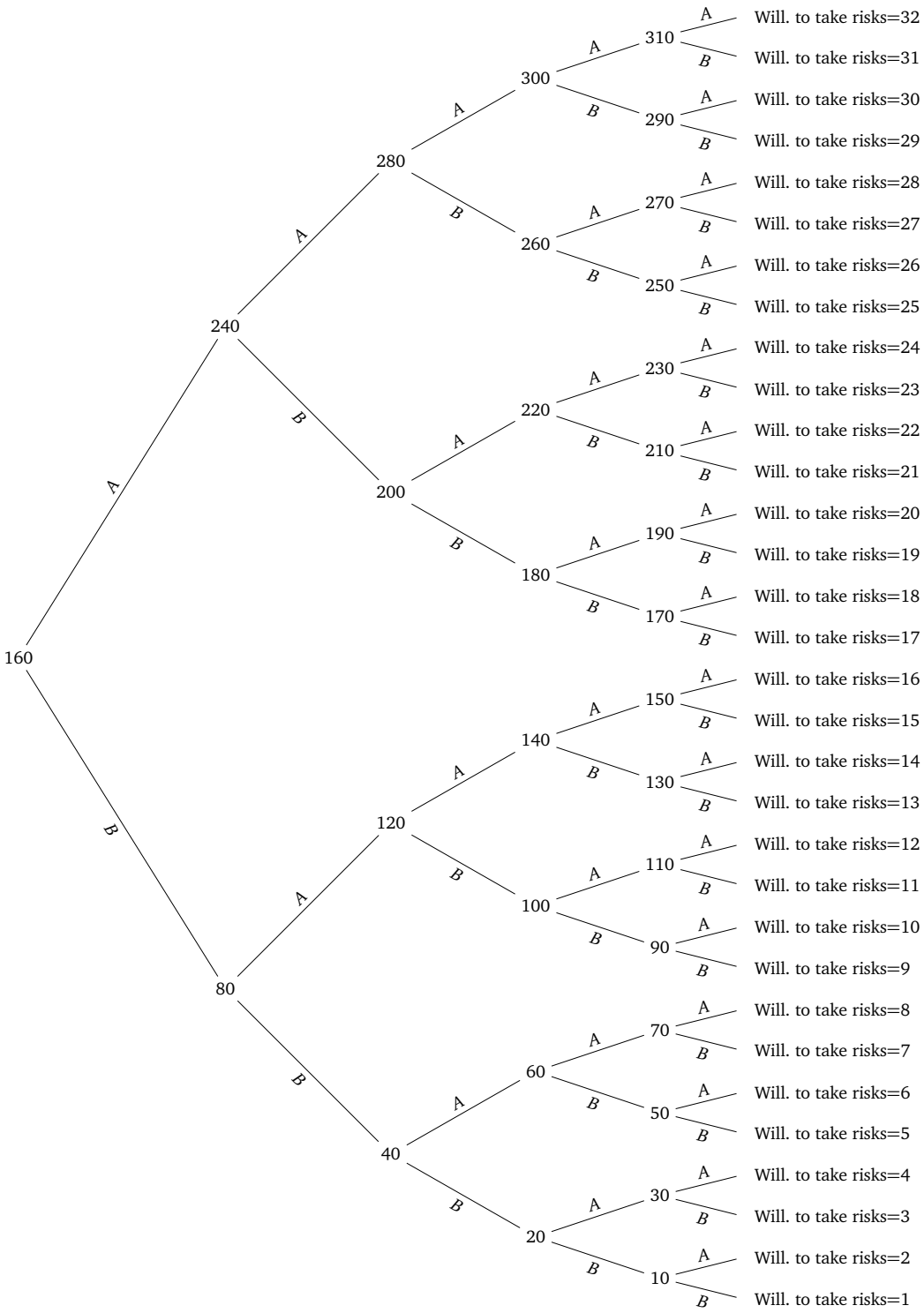

Figure 6: Tree for the staircase risk task (numbers = sure payment, A = choice of sure payment, B = choice of lottery). The staircase procedure worked as follows. First, each respondent was asked whether they would prefer to receive 160 euros for sure or whether they preferred a 50:50 chance of receiving 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 Missing Values

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. (2016). 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 above. 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

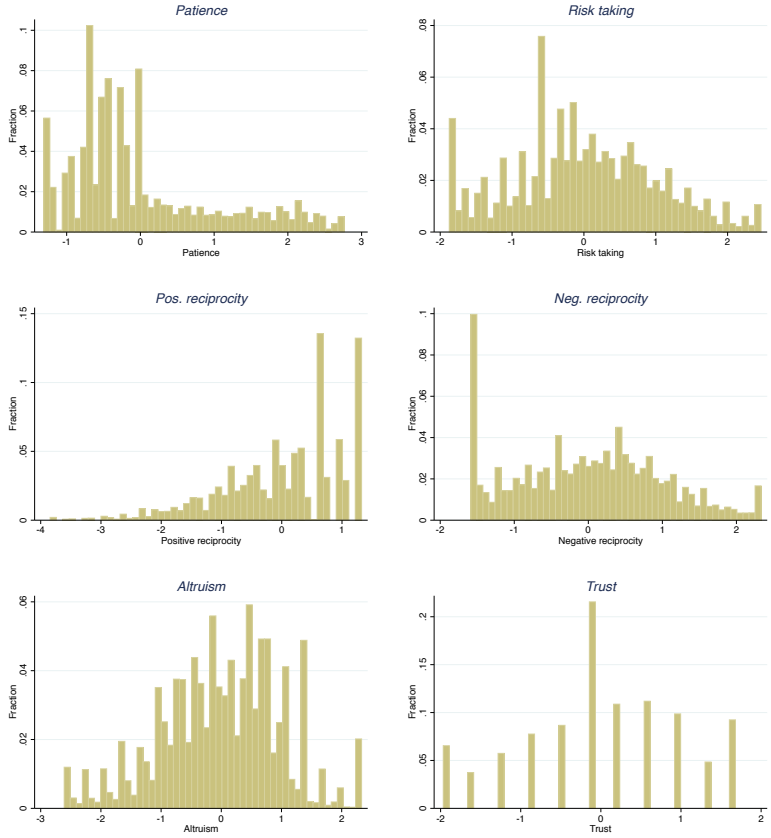

Figure 7: Distribution of preferences at individual level. The figure plots the distribution of standardized preference measures at the individual level. All data are standardized at the level of the individual in the full sample.

A.8.2 Country Level

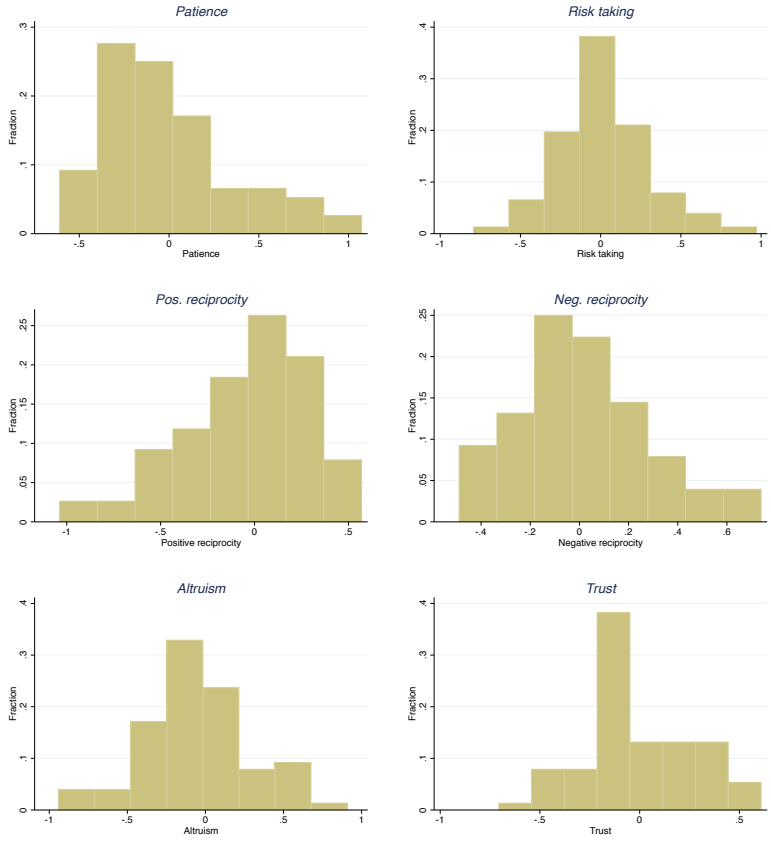

Figure 8: Distribution of preferences at country level. The figure plots the distribution of country averages of standardized preferences. All data are standardized at the level of the individual using the full sample.

B Correlations Among Preferences at the Individual Level

Table 12 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 12: 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 13 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 13: 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

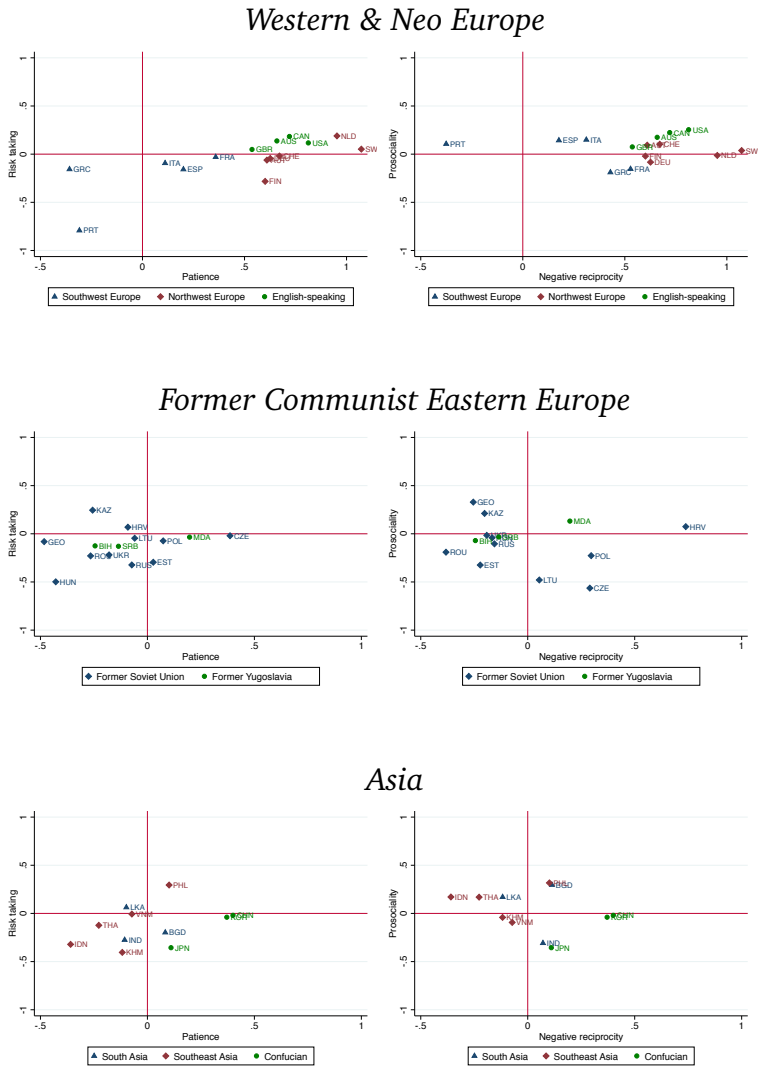

Figure 9a: Risk, time, and social preferences by world region (1/2). Each subpanel (row) plots risk taking, patience, negative reciprocity, and prosociality of all countries within a given world region. The prosociality score is computed as the average of altruism, positive reciprocity, and trust.

D 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 er-

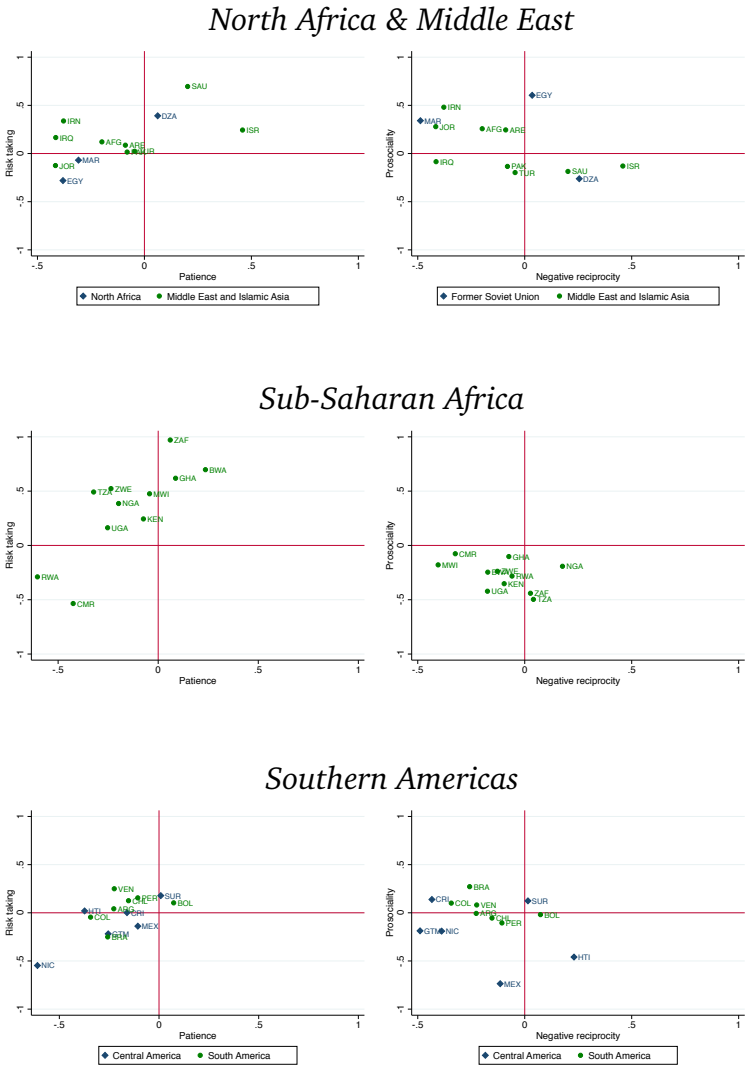

Figure 9b: Risk, time, and social preferences by world region (2/2). Each subpanel (row) plots risk taking, patience, negative reciprocity, and prosociality of all countries within a given world region. The prosociality score is computed as the average of altruism, positive reciprocity, and trust.

ror 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: $Var(\delta)/Var(P) = 1\text{-}\ Var(\epsilon)/Var(P)$. 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 3, 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., 2015). 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.

E Additional Results on Individual-Level Determinants

E.1 Robustness Check for Individual-Level Determinants

Table 14: 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)
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)
Age	0.0072***	0.0069***	-0.00083	0.0046**	0.010***	0.0079***	-0.0036*	-0.0018	-0.000060	0.00024	0.0037*	0.00071
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age squared	-0.00014***	-0.00013***	-0.00012***	-0.00017***	-0.00012***	-0.000090***	-0.000045**	-0.000069***	0.0000015	-0.000013	0.0000032	0.000024
	(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		0.00058***		0.00044*		0.00021		-0.00066**		0.00037*		0.0022***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Subj. health index		0.00097***		0.0012***		0.0012***		-0.00056**		0.0010***		0.00052**
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
1 if christ		-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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E.2 Age Profiles Separately by World Region

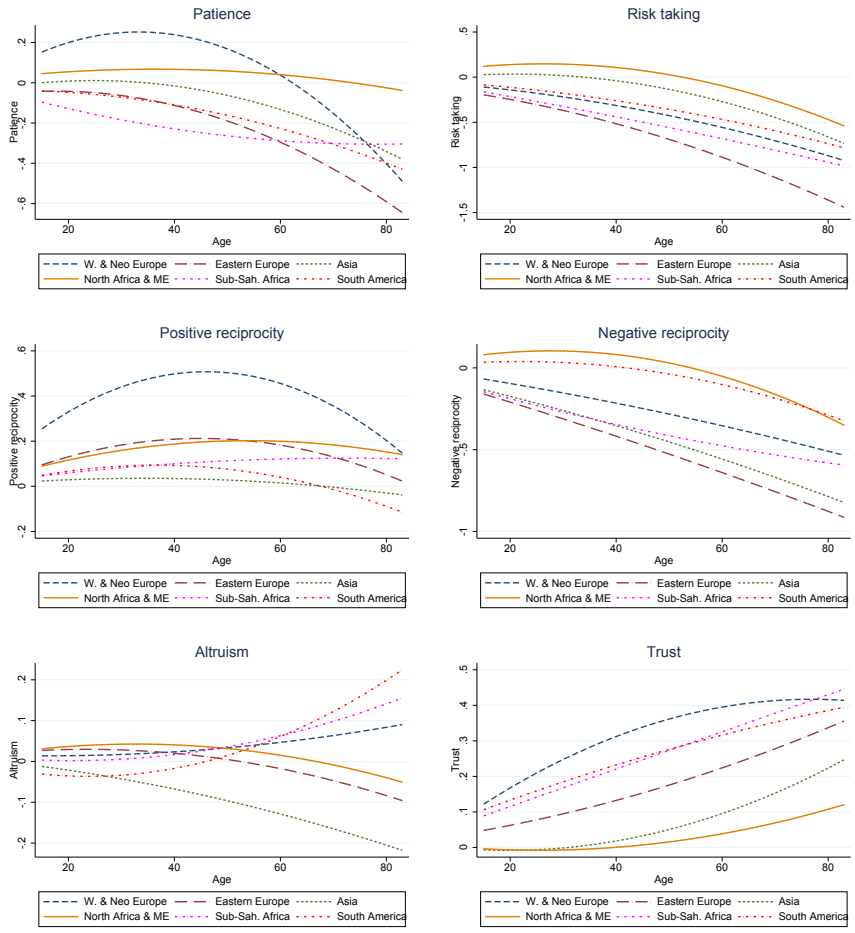

Figure 10: Age profiles separately by continent.

F Additional Results on Country-Level Determinants

Table 15: Preferences and Country-Level Determinants: Multivariate Regression Analysis

	Dependent variable:					
	Patience	Risk taking	Pos. reciprocity	Neg. reciprocity	Altruism	Trust
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to equator	0.010	-0.005	-0.008	0.004	-0.011*	0.002
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)
Longitude	0.000	-0.000	0.002**	0.001*	0.002**	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
% at risk of malaria	-0.000	0.007	-0.495**	0.027	-0.627***	-0.344**
	(0.196)	(0.166)	(0.193)	(0.157)	(0.190)	(0.156)
Average temperature	-0.009	0.002	-0.011	0.008	-0.010	0.005
	(0.009)	(0.008)	(0.009)	(0.007)	(0.009)	(0.007)
Average precipitation	0.000	-0.003***	0.001	0.000	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
% living in (sub-)tropical zones	0.154	0.140	0.059	-0.200	0.272	0.124
	(0.285)	(0.241)	(0.279)	(0.227)	(0.276)	(0.227)
Constant	-0.243	0.372	0.378	-0.218	0.477	-0.104
	(0.355)	(0.301)	(0.349)	(0.284)	(0.344)	(0.283)
Observations	75	75	75	75	75	75
R-squared	0.268	0.212	0.162	0.161	0.201	0.178

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). See Appendix K for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G Additional Details on Relationship Between Preferences and FTR

G.1 Robustness Check for Country-Level FTR Regressions: OLS

While the main text reported WLS estimates, Table 16 reports OLS estimates.

Table 16: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G.2 Additional Within-Country Results

Table 17: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

H Additional Results on Individual-Level Outcomes

H.1 Robustness Checks: All Preferences Simultaneously

Table 18: Patience and accumulation decisions, risk preferences and risky choices: All preferences

	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	0.026***	0.021***	0.046***	0.026***	0.0034	0.0022	0.0062***	0.0052**	-0.0038	-0.0069
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Risk taking	0.030***	0.014**	0.068***	0.021***	0.026***	0.023***	0.019***	0.016***	0.051**	0.030**
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)
Positive reciprocity	0.013	0.0097	0.035***	0.024***	0.012***	0.0083***	0.0065**	0.0073***	0.0026	-0.0024
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Negative reciprocity	0.021**	0.011	0.015***	-0.0060	-0.0019	-0.0019	0.0038	0.0019	0.038***	0.033***
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Altruism	0.017**	0.018***	0.028***	0.025***	0.0042*	0.0068***	0.011***	0.011***	-0.020**	-0.0055
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Trust	-0.0052	-0.0024	-0.016***	-0.016***	-0.0021	-0.0018	0.00058	0.0020	-0.030***	-0.011
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Age		0.00018		0.0093***		0.016***		0.0060***		0.026***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Age squared		-0.000018		-0.00017***		-0.00016***		-0.000073***		-0.00028***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
1 if female		0.0040		-0.016		-0.054***		-0.021***		-0.57***
		(0.01)		(0.01)		(0.01)		(0.00)		(0.03)
Subj. math skills		0.012***		0.040***		0.0052***		0.0019***		-0.012***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Constant	0.16***	-0.38***	1.29***	0.17**	0.034***	-0.36***	0.083***	0.10***	0.29***	0.050
	(0.00)	(0.09)	(0.00)	(0.07)	(0.00)	(0.04)	(0.00)	(0.03)	(0.01)	(0.11)
Country FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Region FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	14436	13835	76442	66366	70149	60304	54821	49143	14452	13847
R^2	0.082	0.187	0.225	0.365	0.065	0.138	0.111	0.170	0.033	0.233

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 K). Self-employment and planned self-employment are binary, while smoking intensity is measured in three categories (never, occasionally, frequently). Additional controls include log household income, and indicators for religious affiliation. See Appendix K for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 19: Patience and accumulation decisions, risk preferences and risky choices: All preferences

	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***	0.058***	0.035***	0.034***	0.053***	0.050***	0.034***	0.031***	0.023***	0.020***	0.018***	0.016***	-0.0014	0.0029
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Positive reciprocity	0.000096	0.0038	0.0046	0.0013	0.037***	0.034***	0.018***	0.019***	-0.00044	-0.0024	0.017***	0.016***	0.012***	0.0078***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Negative reciprocity	-0.0087***	-0.0049*	-0.0047	-0.0041	-0.00066	-0.0052	0.0046	0.00076	0.016***	0.013***	0.0043	-0.0015	-0.010***	-0.00054
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Trust	0.0085**	0.0052	0.0076**	0.010***	-0.0082**	-0.0069**	0.0012	0.0016	0.0030	0.0021	0.00064	0.0020	0.010***	0.0022
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Patience	0.012**	0.0089**	0.016***	0.012***	0.0086**	0.0074**	0.015***	0.011***	0.010***	0.0077**	0.0049**	0.0012	0.0047	0.0070***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Risk taking	0.0065*	0.0096***	0.013***	0.013***	0.027***	0.018***	0.013***	0.0092***	0.021***	0.019***	0.021***	0.012***	-0.015***	0.00018
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age		0.0064***		0.0043***		0.0076***		0.0025***		0.010***		-0.0087***		0.056***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Age squared		-0.000047***		-0.000045***		-0.000090***		-0.000027***		-0.000098***		0.000074***		-0.00054***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
1 if female		0.012**		-0.015***		-0.013**		0.00026		-0.044***		0.014***		-0.024***
		(0.01)		(0.01)		(0.01)		(0.00)		(0.01)		(0.00)		(0.01)
Subj. math skills		0.0084***		0.0067***		0.0079***		0.0065***		0.0083***		0.0044***		0.0021*
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Constant	0.43***	0.31***	0.19***	0.048	0.63***	0.33***	0.18***	-0.24***	0.20***	0.032	0.55***	0.32***	0.64***	-0.38***
	(0.00)	(0.04)	(0.00)	(0.04)	(0.00)	(0.04)	(0.00)	(0.04)	(0.00)	(0.03)	(0.00)	(0.03)	(0.00)	(0.05)
Country FE	Yes	No	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	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	57091	51839	57080	51835	54887	51639	55184	52007	54846	51598	64690	57486	76360	66287
R^2	0.179	0.242	0.087	0.140	0.087	0.150	0.113	0.180	0.050	0.107	0.094	0.170	0.053	0.237

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 K). Self-employment and planned self-employment are binary, while smoking intensity is measured in three categories (never, occasionally, frequently). Additional controls include log household income, and indicators for religious affiliation. See Appendix K for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

H.2 Robustness Checks: Alternative Estimation Methods

This appendix reports robustness checks on the relationship between preferences and behaviors at the individual level. Specifically, while Section 4.1 of the main text reported the results of OLS estimations, we now re-estimate all specifications using probit or ordered probit regressions. As Tables 20 and 21 show, the results are unchanged.

Table 20: Patience and accumulation decisions, risk preferences and risky choices: Probit estimates

	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)
WP11292 Profession Outside Country										
Patience	0.12***	0.091***	0.13***	0.072***						
	(0.02)	(0.02)	(0.01)	(0.01)						
Risk taking					0.14***	0.12***	0.15***	0.13***	0.11***	0.073***
					(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Age		0.000053		0.022***		0.096***		0.062***		0.066***
		(0.01)		(0.01)		(0.01)		(0.00)		(0.01)
Age squared		-0.000065		-0.00041***		-0.00098***		-0.00081***		-0.00074***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
1 if female		-0.00068		-0.036		-0.27***		-0.14***		-1.35***
		(0.04)		(0.03)		(0.05)		(0.03)		(0.13)
Subj. math skills		0.051***		0.095***		0.032***		0.018***		-0.024***
		(0.01)		(0.00)		(0.00)		(0.00)		(0.01)
Constant	-0.94***	-2.99***			-1.77***	-4.33***	-1.39***	-1.81***		
	(0.00)	(0.31)			(0.01)	(0.18)	(0.01)	(0.15)		
cut1										
Constant			0.50***	3.55***					0.91***	1.72***
			(0.01)	(0.16)					(0.02)	(0.26)
cut2										
Constant			2.26***	5.60***					1.22***	2.12***
			(0.04)	(0.18)					(0.02)	(0.31)
Country FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Region FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	15260	14011	79357	68409	72839	59646	57072	45016	15309	14490
Pseudo R^2	0.065	0.161	0.116	0.226	0.079	0.160	0.144	0.196	0.020	0.197

(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 K). Self-employment and planned self-employment are binary, while smoking intensity is measured in three categories (never, occasionally, frequently). Additional controls include log household income, and indicators for religious affiliation. See Appendix K for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 21: Social preferences and social interactions: Probit estimates

	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)
WP11292 Profession Outside Country														
Altruism	0.22***	0.20***	0.14***	0.15***	0.15***	0.15***	0.13***	0.13***	0.093***	0.087***	0.084***	0.074***	-0.0030	0.010
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Positive reciprocity	0.0086	0.022*	0.024**	0.012	0.10***	0.097***	0.074***	0.082***	0.0027	-0.0065	0.071***	0.069***	0.032***	0.027***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Negative reciprocity	-0.015	-0.0075	0.0048	0.0016	0.015	-0.0068	0.036***	0.015	0.078***	0.064***	0.040***	0.0040	-0.030***	0.0033
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age		0.023***		0.017***		0.022***		0.0098***		0.041***		-0.039***		0.17***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.01)
Age squared		-0.00018***		-0.00019***		-0.00027***		-0.00011***		-0.00041***		0.00032***		-0.0016***
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
1 if female		0.038*		-0.069***		-0.052***		-0.0048		-0.17***		0.050***		-0.087***
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)
Subj. math skills		0.032***		0.029***		0.023***		0.027***		0.034***		0.022***		0.0065*
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Constant	-0.20***	-0.91***	-0.89***	-1.54***	0.34***	-0.51***	-0.96***	-2.70***	-0.85***	-1.62***	0.11***	-0.54***	0.38***	-2.63***
	(0.01)	(0.12)	(0.01)	(0.13)	(0.01)	(0.12)	(0.01)	(0.13)	(0.01)	(0.12)	(0.00)	(0.13)	(0.00)	(0.16)
Country FE	Yes	No	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	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	58229	51598	58213	51422	55991	52130	56253	50531	55944	51408	65986	56577	77881	67375
Pseudo R^2	0.147	0.196	0.082	0.125	0.062	0.112	0.103	0.153	0.046	0.098	0.093	0.163	0.039	0.192

Probit estimates, standard errors (clustered at country level) in parentheses. For the purposes of this table, age is divided by 100. Additional controls include log household income, and indicators for religiou affiliation. See Appendix K for details on all dependent variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

H.3 Distributions 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 4.1 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 11 shows that the positive correlation between patience and savings holds in virtually all countries in our sample.

While Figure 11 reports the results for patience and risktaking, Figure 12 visualizes the relationships between altruism and behaviors. Finally, Figure 13 presents the correlations between positive and negative reciprocity and the behaviors discussed in Section 4.1 of the main text.

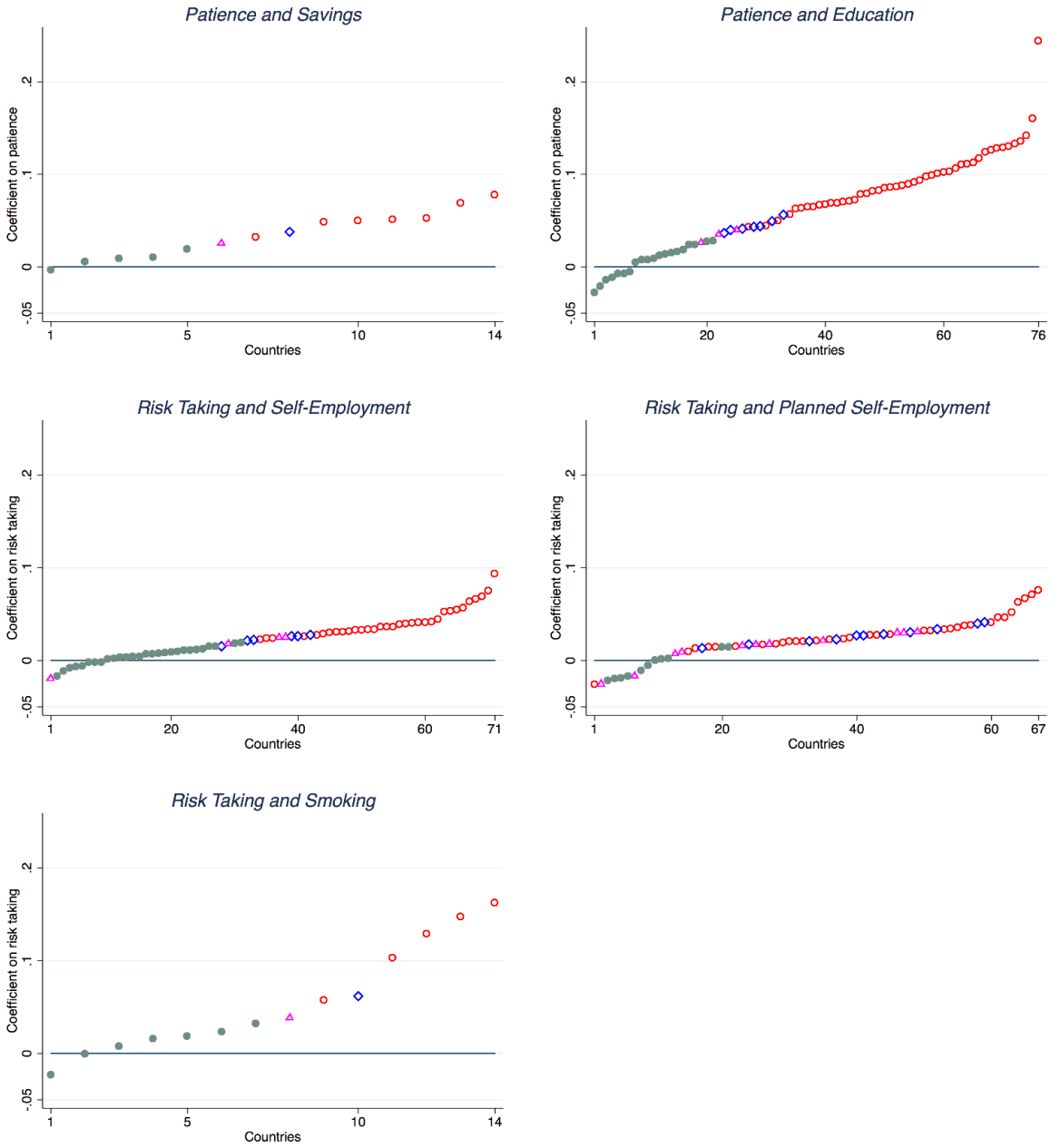

Figure 11: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable, 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.

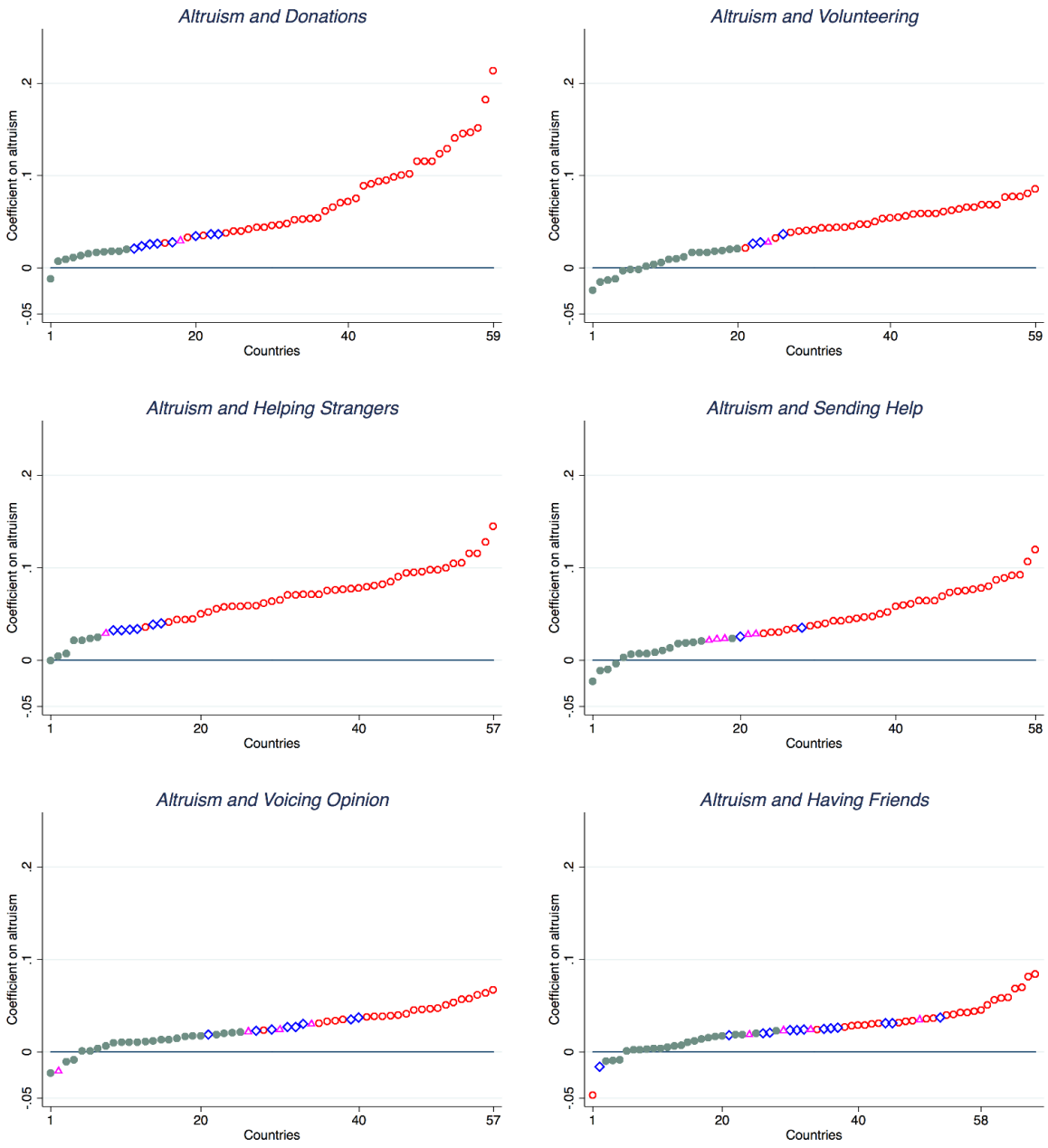

Figure 12: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable, 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.

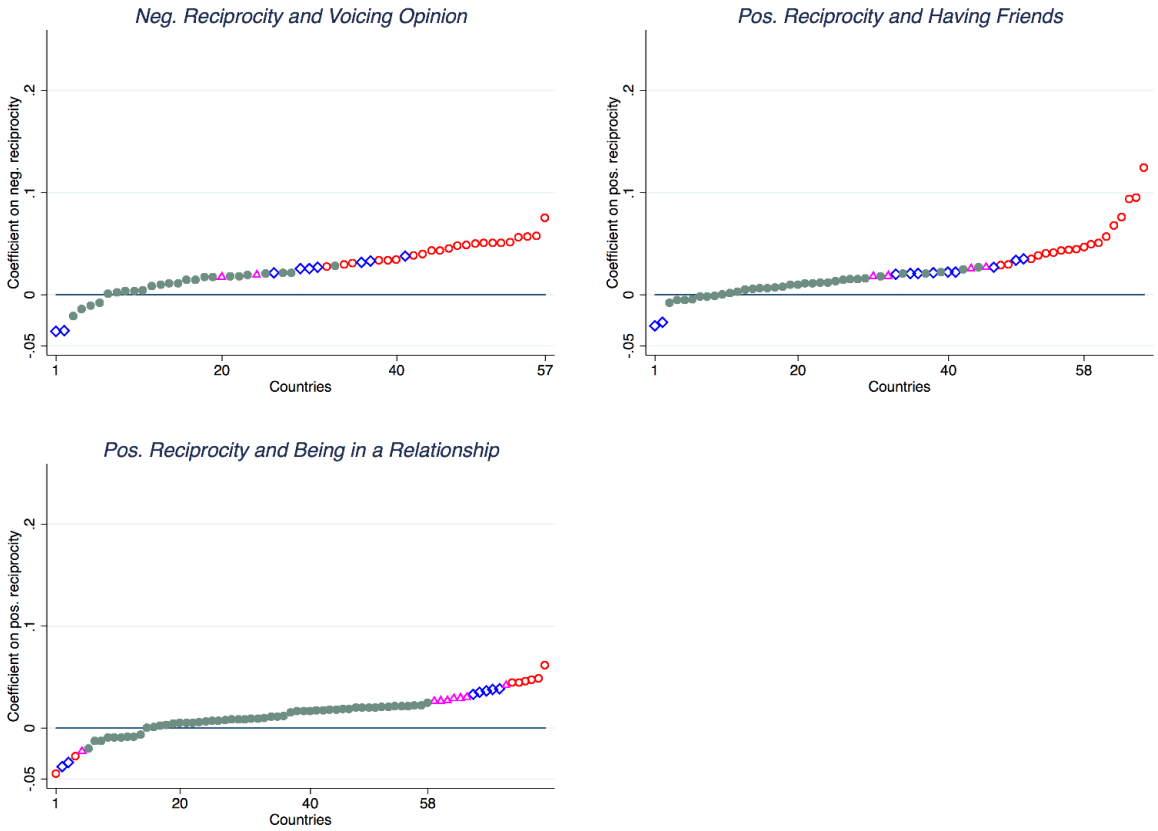

Figure 13: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable, 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.

I Additional Results on Country-Level Outcomes

Table 22: Country-level outcomes: All preferences and trust

	Dependent variables:											
	Ln[GDP p/c ppp]		Human Development index		Labor protection index		Life risk index		Volunteering and donation as fraction of GDP		Ln[Conflicts]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Patience	2.670***	2.130***	0.228***	0.155***	0.00900	-0.109	-1.914***	-1.401***	1.837***	1.809**	0.749**	0.285
	(0.302)	(0.287)	(0.0299)	(0.0287)	(0.0782)	(0.0891)	(0.266)	(0.342)	(0.533)	(0.741)	(0.299)	(0.367)
Risk taking	-1.192**	-0.605	-0.128***	-0.0441	-0.234**	-0.121	2.537***	1.749***	0.730	0.231	-0.902**	-0.660
	(0.526)	(0.378)	(0.0475)	(0.0304)	(0.104)	(0.113)	(0.482)	(0.609)	(0.719)	(0.820)	(0.407)	(0.547)
Positive reciprocity	0.453	-0.0628	0.101*	0.0486	-0.0475	-0.0767	-0.996*	-0.452	-0.553	-0.974	0.270	-0.130
	(0.578)	(0.472)	(0.0569)	(0.0404)	(0.114)	(0.123)	(0.538)	(0.609)	(1.147)	(1.138)	(0.560)	(0.522)
Negative reciprocity	0.417	0.0869	0.0548	0.0151	-0.0180	-0.0178	-0.675	-0.326	-0.326	0.320	1.378***	0.938**
	(0.563)	(0.467)	(0.0486)	(0.0385)	(0.0853)	(0.0999)	(0.442)	(0.417)	(0.767)	(0.759)	(0.379)	(0.403)
Altruism	-0.879	-0.581	-0.0977*	-0.0855*	-0.0371	-0.00177	0.362	0.219	0.138	0.930	-0.482	-0.0277
	(0.582)	(0.525)	(0.0552)	(0.0458)	(0.109)	(0.121)	(0.492)	(0.523)	(0.626)	(0.690)	(0.403)	(0.466)
Trust	0.869	0.439	0.0872*	0.0414	0.109	0.0928	-0.601	-0.165	1.133	1.074	0.544	0.465
	(0.568)	(0.439)	(0.0520)	(0.0387)	(0.123)	(0.129)	(0.625)	(0.469)	(0.740)	(0.788)	(0.485)	(0.553)
Constant	8.320***	8.660***	0.705***	0.744***	0.492***	0.271	-0.0184	0.413	1.365***	-0.698	1.839***	1.790
	(0.133)	(0.956)	(0.0123)	(0.0923)	(0.0250)	(0.242)	(0.116)	(0.896)	(0.223)	(1.615)	(0.125)	(1.186)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	76	75	76	75	56	56	67	66	32	32	75	75
R-squared	0.508	0.752	0.512	0.773	0.162	0.289	0.586	0.751	0.503	0.633	0.264	0.388
R^2 with controls only		0.576		0.657		0.195		0.623		0.276		0.324

OLS estimates, robust standard errors in parentheses. Labor protection is measured by the LaPorta labor protection index. The life risk index is the principal component from a pca analysis of average life expectancy, log HIV prevalence, and traffic death rate. See Appendix K for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J Results on Candidate Preference Proxies in the WVS

Table 23: WVS preference proxies and individual-level determinants

	Dependent variable:							
	WVS Long Term Orientation		WVS Value of stimulation		WVS altruism		WVS trust	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.023**	0.028***	-0.25***	-0.23***	-0.0091	-0.011	-0.021***	-0.021***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Age	0.35	-0.11	-2.71***	-2.28***	-0.47	-0.15	0.71***	0.27
	(0.26)	(0.17)	(0.31)	(0.27)	(0.38)	(0.25)	(0.26)	(0.17)
Age squared	0.059	0.42***	1.14***	0.98***	0.21	0.33	-0.35*	-0.19
	(0.22)	(0.14)	(0.27)	(0.23)	(0.34)	(0.25)	(0.20)	(0.14)
Constant	-0.21***	0.40***	1.14***	1.05***	0.17**	-0.010	-0.26***	-0.11**
	(0.06)	(0.06)	(0.07)	(0.07)	(0.08)	(0.06)	(0.07)	(0.05)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
N	323270	323270	154729	154729	80881	80881	308162	308162
R^2	0.005	0.071	0.082	0.156	0.002	0.130	0.003	0.098

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

Table 24: WVS preference proxies and individual-level outcomes

	Dependent variable:					
	Saved last year		Education		Self employed	
	(1)	(2)	(3)	(4)	(5)	(6)
WVS Long Term Orientation	0.0015	0.0051**	-0.19***	-0.13***		
	(0.00)	(0.00)	(0.02)	(0.02)		
WVS Value of Stimulation					0.0043***	0.0045***
					(0.00)	(0.00)
Age		-0.40***		-0.57		1.42***
		(0.11)		(0.73)		(0.18)
Age squared		0.29***		-2.08***		-1.25***
		(0.08)		(0.58)		(0.15)
Female		-0.018***		-0.19***		-0.072***
		(0.00)		(0.03)		(0.01)
Constant	0.33***	0.13***	3.67***	3.97***	0.094***	-0.20***
	(0.00)	(0.04)	(0.01)	(0.23)	(0.00)	(0.04)
Country FE	Yes	No	Yes	No	Yes	No
Region FE	No	Yes	No	Yes	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes
Observations	339318	254654	296130	230555	151767	131698
R^2	0.076	0.171	0.107	0.275	0.120	0.182

OLS estimates, robust standard errors in parentheses. Saved last year is a binary indicator, while education level is measured in eight categories. Self-employment is binary. For the purposes of this table, age is divided by 100. Additional controls include log of categorical income variable, and indicators for religious affiliation. See Appendix K for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 25: WVS preference proxies and country-level outcomes

	Ln[GDP p/c ppp]		Human Development Index		Dependent variable: Labor Protection Index		Life Risk Index		Volunteering and donation as fraction of GDP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WVS Long Term Orientation	-0.95	-1.27***	-0.040	-0.11***						
	(0.65)	(0.43)	(0.06)	(0.03)						
WVS Value of Stimulation					-0.067	0.076	2.38***	0.82		
					(0.08)	(0.10)	(0.57)	(0.62)		
WVS Altruism									-1.41	-1.44
									(1.03)	(2.51)
Constant	8.46***	6.64***	0.72***	0.62***	0.48***	0.14	-0.077	1.43	1.44	-2.632
	(0.15)	(1.25)	(0.01)	(0.10)	(0.03)	(0.25)	(0.15)	(0.92)	(0.30)	(5.13)
Observations	96	88	95	87	57	53	62	59	17	17
R^2	0.026	0.515	0.005	0.701	0.011	0.235	0.267	0.644	0.143	0.567

OLS estimates, robust standard errors in parentheses. Labor protection is measured by the LaPorta labor protection index. The life risk index is the principal component from a pca analysis of average life expectancy, log HIV prevalence, and traffic death rate. See Appendix K for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

K Description and Data Sources of Outcome Variables

K.1 Individual-Level Variables

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

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.

Have friends. 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.

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

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).

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.

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.

Saved last year. Binary variable capturing whether the respondent saved any money in the previous year. 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.

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

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?".

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.*

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.

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

K.2 Country-Level Variables

Conflicts The data on conflicts in ACLED are computed as the total number of conflicts since 2001. The number of conflicts according to PRIO are taken from the Quality of Government dataset, http://www.qogdata.pol.gu.se/codebook/codebook_basic_30aug13.pdf.

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

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

HIV prevalence Primary data taken from Unicef. Missings replaced by data from most recent available period of the World Bank Development Indicators.

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

Human Development Index. Average over the period 2000 – 2010. Source: United Nations Development Programme.

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.

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

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).

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 and Sachs (2001) 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. ercentage 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.

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).

Percent living in poverty Fraction of the population living on less than $2 per day, average 2003-2012. Taken from World Bank Development Indicators.

Traffic death rate Traffic deaths per 1,000 population, estimate of the World Health Organization.

Volunteering and donation as fraction of GDP Dollar value of volunteering and giving as a share of GDP by country, including gifts to religious worship organizations where available, average over the period 1995-2002. Source: Salamon (2004).