NBER WORKING PAPER SERIES

GLOBAL EVIDENCE ON ECONOMIC PREFERENCES

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

Working Paper 23943
http://www.nber.org/papers/w23943

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2017

We thank the editor, Andrei Shleifer, and three anonymous referees. For comments and discussions we are grateful to Johannes Hermle, Benedikt Herrmann, Fabian Kosse, Guido Neidhoefer, and seminar participants at Caltech, CESifo Area Conference on Behavioural Economics, Columbia Business School, EALE conference in Ghent, Bozen-Bolzano, IZA Bonn, LISER, LMU Munich, Maastricht, SITE Experimental Economics, Stockholm, UC San Diego, Bologna, Copenhagen, Konstanz, Toulouse, Virginia, and Berlin. Ammar Mahran provided outstanding research assistance and Markus Antony outstanding administrative support. Falk acknowledges financial support from the European Research Council through ERC # 209214. Dohmen acknowledges financial support from NWO Vidi grant 452-10-006, and from Netspar. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Global Evidence on Economic Preferences
Armin Falk, Anke Becker, Thomas Dohmen, Benjamin Enke, David B. Huffman, and Uwe Sunde
NBER Working Paper No. 23943
October 2017
JEL No. D0,D03,D9

ABSTRACT

This paper studies the global variation in economic preferences. For this purpose, we present the Global Preference Survey (GPS), an experimentally validated survey dataset of time preference, risk preference, positive and negative reciprocity, altruism, and trust from 80,000 individuals in 76 countries. The data reveal substantial heterogeneity in preferences across countries, but even larger within-country heterogeneity. Across individuals, preferences vary with age, gender, and cognitive ability, yet these relationships appear partly country specific. At the country level, the data reveal correlations between preferences and bio-geographic and cultural variables such as agricultural suitability, language structure, and religion. Variation in preferences is also correlated with economic outcomes and behaviors. Within countries and subnational regions, preferences are linked to individual savings decisions, labor market choices, and prosocial behaviors. Across countries, preferences vary with aggregate outcomes ranging from per capita income, to entrepreneurial activities, to the frequency of armed conflicts.

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

Anke Becker
Harvard University
Department of Human Evolutionary Biology
11 Divinity Ave
Cambridge MA 02138
ankebecker@fas.harvard.edu

Thomas Dohmen
Department of Economics
University of Bonn
Adenauerallee 24-42, 53113 Bonn
t.dohmen@uni-bonn.de

Benjamin Enke
Department of Economics
Harvard University
Littauer M008
Cambridge, MA 02138
and NBER
enke@fas.harvard.edu

David B. Huffman
Economics Department
University of Pittsburgh
230 S. Bouquet St.
Pittsburgh, PA 15260
huffmand@pitt.edu

Uwe Sunde
University of Munich
Department of Economics
Schackstr 4
80539 Munich
GERMANY
Uwe.Sunde@econ.lmu.de

1 Introduction

Many theories of human behavior, in economics and neighboring disciplines, assume that a set of preferences drives individual decision-making. This includes preferences about risk, the timing of rewards, and in the social domain, reciprocity, altruism, and trust. Given the importance of preferences in economists' conceptual framework, a substantial empirical literature has focused on understanding the potential determinants and consequences of preference variation. While this literature has produced many insights about individual-level heterogeneity in preferences in certain populations, less is known about the global variation in preferences. This partly reflects the lack of a global data set, representative at the country level, with measures specifically designed to capture economic preferences.

This paper introduces such a dataset, the Global Preference Survey (GPS). The empirical analysis is motivated by a set of questions about the extent and nature of global preference heterogeneity, at different levels of aggregation: Do countries differ in terms of average preferences? Are certain preferences correlated, leading to preference bundles? How large is cross-country variation in preferences relative to within country variation? Regarding the potential determinants of preference heterogeneity, do the GPS preference measures vary with individual characteristics like gender, age, and cognitive ability? To what extent are these differences universal, or more country-specific? Are country-level preference profiles related to differences in geography, culture, language, or religion? Turning to the relationship between preferences and outcomes, how does individual-level heterogeneity in financial, labor market, or prosocial choices vary with preferences around the world? Are differences in aggregate preference profiles correlated with the cross-country variation in outcomes such as economic development, charitable activities, or violent conflict?

This paper explores these questions by making use of the core features of the GPS: (i) coverage of 76 countries that represent approximately 90 percent of the world population; (ii) representative population samples within each country for a total of 80,000 respondents, (iii) measures designed to capture time preference, risk preference, altruism, positive reciprocity, negative reciprocity, and trust, based on an ex ante experimental validation procedure (Falk et al., 2016) as well as pre-tests in culturally heterogeneous countries, (iv) standardized elicitation and translation techniques through the pre-existing infrastructure of a global polling institute, Gallup. *Upon publication, the data will be made publicly available online.* The data on individual preferences are complemented by a comprehensive set of covariates provided by the Gallup World Poll 2012.

The analysis begins by describing the nature of the heterogeneity in preferences, both across and within countries. Many of the world's most patient populations are located in Europe or the English-speaking world, while risk taking is particularly prevalent in large parts of Africa and the Middle East. Prosocial preferences are particularly pronounced throughout Asia and relatively weak in Sub-Saharan Africa. 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 prosocial traits of positive reciprocity, altruism, and trust form another grouping. While the between-country variation in preferences is substantial, within-country variation is larger, suggesting that individual characteristics are even more important for explaining preference differences than national borders.

The analysis next turns to a more systematic, regression-based analysis of potential determinants of preference variation. The results establish that, at the individual level, preferences vary systematically with gender, cognitive ability, and age. For example, women are more impatient, less risk tolerant, and more prosocial than men. Cognitive skills are uniformly positively linked to patience, risk taking, and social preferences, and all preferences are subject to age patterns. At the same time, the relationships between sociodemographics and preferences hide considerable heterogeneity across countries: while some relationships, such as between risk aversion and gender, go in the same direction in almost all countries, others, such as the age profile for patience, appear to depend on the level of development.

Prior research has articulated and tested various hypotheses about how (population-level) preference profiles might be determined by geographic or cultural variables, including a "culture of honor", a "Protestant ethic", a "savings-linguistics" hypothesis, or relationships between agricultural conditions and preferences (e.g., Weber, 1930; Nisbett and Cohen, 1996; Tabellini, 2008; Chen, 2013; Galor and Özak, 2016; Galor et al., 2017; Litina, 2016). The GPS allows investigating the correlations between all preferences and such geographic and cultural variables. Regarding biological and geographic conditions, patience, trust, and negative reciprocity are all significantly positively correlated with absolute latitude and the presence of large domesticable animals, the latter result broadly in line with the culture of honor hypothesis. Trust is significantly decreasing in different measures of agricultural suitability. Turning to cultural variables, patience is strongly and significantly correlated with Protestantism and a linguistic structure called future time reference, in line with previously articulated hypotheses and evidence. Moreover, different measures of individualism are positively correlated with patience.

In a next step, we explore the relationships between preferences and individual-level behaviors and outcomes that economists have emphasized as being potentially driven by risk, time, and social preferences. The data show that patient individuals are more likely to save and have higher educational attainment; more risk tolerant individuals are more likely to be self-employed and to be smokers; and social preferences are 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, which provides an additional, out of context check of the ability of the GPS measures to capture behaviorally relevant heterogeneity across a wide range of cultures.

Finally, the paper studies the correlations between country-level preferences and a selected set of outcome variables that previous literatures have suggested may be related to preferences. In a first step, we focus on the relationship between preferences and economic development. An extensive line of work has studied the relationship between trust and per capita income (e.g.,

Knack and Keefer, 1997; Algan and Cahuc, 2010), and a considerable theoretical literature has emphasized the role of time preference for development. In the GPS data, trust is significantly correlated with development; at the same time, the relationship between patience and income is much stronger, in terms of both quantitative magnitude and statistical significance. For example, when both patience and trust are inserted into a joint regression, trust loses significance. Moving to additional aggregate outcomes, we establish that risk taking is significantly correlated with proxies for entrepreneurial activities, in line with the within-country correlation between risk taking and self-employment. Average social preferences correlate with donations and volunteering across countries, again akin to the corresponding within-country results. Finally, average negative reciprocity in a country is strongly correlated with the frequency of armed conflicts.

The findings of this paper tie into several different literatures in behavioral and experimental economics, cultural economics, and long-run development. Within behavioral and experimental economics, researchers have investigated both potential individual-level determinants and outcomes of preference variation, though typically in smaller and more specialized samples. Work on potential determinants includes Barsky et al. (1997); Frederick (2005); Dohmen et al. (2008); Croson and Gneezy (2009); Dohmen et al. (2010, 2011), while research on outcomes has been conducted by Barsky et al. (1997); Ventura (2003); Kirby and Petry (2004); Eckel et al. (2005); Bonin et al. (2007); Chabris et al. (2008); Guiso and Paiella (2008); Dohmen et al. (2009); Meier and Sprenger (2010); Rustagi et al. (2010); Tanaka et al. (2010); Sutter et al. (2013); Golsteyn et al. (2014); Kosfeld and Rustagi (2015); Åkerlund et al. (2016). This paper speaks to open questions in the literature, e.g., whether certain gender differences in preferences are relatively universal or specific to certain cultures or development levels, or perhaps the product of publication bias (Gneezy et al., 2009; Niederle, 2014).

In cultural economics and political economy, this paper is most closely related to research that has measured variation in preferences across societies by focusing on small selected groups such as small-scale societies or university students (Henrich et al., 2001, 2006, 2010; Apicella et al., 2014; Rieger et al., 2014; Talhelm et al., 2014; Vieider et al., 2015b). The GPS has the potential to open up research agendas on the cultural origins of preference variation, something that has been difficult thus far given the absence of representative cross-country data on preferences. Another related literature has investigated the role of culture or geography in shaping economic behavior, but focusing on variables such as female labor force participation, fertility, individualism, and future-orientation (Giuliano, 2007; Fernández and Fogli, 2009; Gorodnichenko and Roland, 2011; Alesina and Giuliano, 2013; Chen, 2013; Alesina et al., 2015; Galor and Özak, 2016). Finally, the results on the cross-country relationships between preferences and outcomes naturally tie into the literature on comparative development, which makes increasing use of arguments about cultural variation (Ashraf and Galor, 2013; Spolaore and Wacziarg, 2013; Galor and Özak, 2016). While this literature has largely focused on trust, the GPS data may open up the investigation of additional hypotheses.

The remainder of the paper proceeds as follows. Section 2 provides details on the GPS dataset and compare the data with other commonly available data sources. Section 3 presents

descriptives on the global preference variation. Section 4 studies the relationship between preferences and potential determinants of preference variation, both at the individual and at the country level. Section 5 investigates the relationships between preferences and economic outcomes and Section 6 concludes.

2 Dataset

2.1 General Data Characteristics

The GPS data were collected within the framework of the 2012 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. This section discusses some noteworthy characteristics of the data. In addition, Online Appendix A contains an extensive documentation of the data-collection process and details on the construction of the preference measures.

One important feature of the GPS data is that it measures preferences for a nationally representative sample for each of the 76 countries covered. Thus, 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 preferences vary across countries. The median sample size was 1,000 participants per country.[1] Respondents were selected through probability sampling; ex-post representativeness of the data can be achieved using weights provided by Gallup. In total, the sample involves 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 representative sample. The collection of countries spans 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 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 GPS survey items in 22 countries of various cultural heritage. This was in order to ensure the implementability of the preference 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 Online 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

[1] 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.).

stakes constant.[2] 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 18 countries, telephone interviews were also used. Table 11 in Online Appendix A shows the countries included in the GPS, along with numbers of observations and the survey mode.

2.2 Preference Measures

The GPS was 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 payoffs of others – that are typically relevant for the trade-offs involved in economic decisions. Accordingly, the GPS includes measures of time preference, risk 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.[3]

The GPS preference measures are based on twelve survey items, which were selected in an initial survey validation study (see Falk et al., 2016, for details). The validation procedure involved conducting multiple incentivized choice experiments for each preference, and testing the relative abilities of a wide range of different question wordings and formats to predict behavior in these choice experiments. The particular items used to construct the GPS preference measures were selected based on optimal performance out of menus of alternative items (for details see Falk et al., 2016). Experiments provide a valuable benchmark for selecting survey items, because they can approximate the ideal choice situations, specified in economic theory, in which individuals make choices in controlled decision contexts. Experimental measures are very costly, however, to implement in a globally representative sample, whereas survey measures are much less costly.[4] Selecting survey measures that can stand in for incentivized revealed preference measures leverages the strengths of both approaches.

The survey items are summarized in Table 1. For most preferences the optimization procedure resulted in a combination of two survey items, involving one qualitative item, which is more abstract, and one quantitative item, which puts the respondent into a precisely defined

[2] 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 of the quantitative items due to difficulties in assessing the involved monetary amounts.

[3] 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).

[4] 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., Rieger et al., 2014; Vieider et al., 2015b,a).

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	Gift in exchange for help	0.52
	Self-assessment: Willingness to return a favor	0.48
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 Online Appendix A.6 for the wording of the questions and Online Appendix A.9 for a discussion of the weights.

hypothetical choice scenario. The quantitative items more closely resemble the choice-based experiment measures, in that they hold stakes, probabilities, and relevant information conditions constant, helping deliver comparable measures across different individuals and cultures. At the same time, the qualitative items also have explanatory power for behavior in the experiments.

For each preference, the survey items are combined into a single preference measure using the weights that (endogenously) emerged from this experimental validation procedure. In particular, the experimental validation procedure allows an analysis of which linear combination of survey items performs best in predicting the corresponding experimental behavior. These weights are then used to compute the final preference measures, in line with the goal of constructing variables that have experimental counterparts. At the same time, future research using the GPS may wish to focus on selected subsets of our items by, e.g., focusing attention on the quantitative survey formats.

Finally, for ease of interpretation, each preference measure is 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.

Time Preference. The 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 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 Online 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 Table 1 indicates, the quantitative item has a weight of 71% in the time preference measure.

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. Choices were between a fixed lottery, in which the individual could win x or zero, and varying sure payments, y:

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?

Choice of the lottery resulted in an increase of the sure amount being offered in the next question, and vice versa, thereby zooming in around the individual's certainty equivalent. Online Appendix A.6.2 contains an exposition of the entire sequence of survey items. The qualitative item asks for the respondents' self-assessment of their willingness to take risks on an eleven-point scale ("In general, how willing are you to take risks?"). This qualitative subjective self-assessment has previously been shown to be predictive of risk-taking behavior in the field in a representative sample (Dohmen et al., 2011) as well as of incentivized experimental risk-taking across countries in student samples (Vieider et al., 2015b). The qualitative item and the outcome of the quantitative staircase measure were combined through roughly equal weights (Table 1).

Positive Reciprocity. Respondents' propensities to act in a positively reciprocal way were also measured using one quantitative item and one qualitative question. First, respondents 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. Respondents were then asked which out of six presents (worth between 5 and 30 euros, or the respective country-specific equivalents) they would give to the stranger as a "thank you". Second, respondents were asked to provide a self-assessment about how willing they are to return a favor on an 11-point Likert scale. These two items receive roughly equal weights (Table 1).

Negative Reciprocity. Negative reciprocity was elicited through three self-assessments. First, respondents 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*.[5] This last item captures prosocial punishment and hence a concept akin to norm enforcement. These three items receive weights of about one third each (Table 1).

Altruism. Altruism was measured through a combination of one qualitative and one quantitative item, both of which are related to donations. The qualitative question asked respondents 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 (Table 1).

Trust. The trust measure is based on one item, which asked respondents 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 and the fact that there already exists a global measure of trust in the World Values Survey (WVS) data set determined the choice to have only one item measuring trust.

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, 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)

[5] 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.

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. The GPS survey module 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. The publicly available GPS dataset includes all preference measures as well as information on respondents' age, gender, cognitive skills, and country of residence. In addition, the data contain Gallup's individual-level identifier, so that the preference measures can be matched to Gallup's entire World Poll dataset.

2.4 The GPS as a Complement to Existing Global Surveys

Questions in existing global surveys, although designed for other purposes, could potentially serve as proxies for the set of preferences measured in the GPS. 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. A challenge, however, is distinguishing weak from strong proxies, when relying on intuition. One way that the GPS complements existing surveys is by providing a new source of data to assess the validity of potential preference proxies.

This subsection presents results from such a validation exercise, for measures in the WVS and the data of Hofstede (2001), two widely used global surveys that are designed to measure traits that might be related to preferences: attitudes, beliefs and personality traits.[6] The WVS provides individual-level responses for representative samples from a wide range of countries, while the Hofstede data includes measures for a similar range of countries, but at the country-level, and based on non-representative samples (mainly IBM employees).

Identifying candidate preference proxies in these data sets involved looking for key-words, and types of trade-offs, that seemed plausibly related to a respective preference. This initial identification was necessarily based on intuition. The procedure did not yield any WVS questions or Hofstede cultural dimensions that asked about something that seemed related to positive or negative reciprocity. It did lead to identifying measures that might possibly proxy for the other preferences, with varying degrees of plausibility.

In the WVS, the question that seems most closely related to time preference is an item designed to capture "Long Term Orientation" in terms of childrearing. 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?" This variable is coded as 1 if the individual lists "thrift, saving money and things," regardless of what other qualities the respondent lists. For

[6]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 includes similar measures to the ones that we analyze in the WVS. These surveys have a more limited geographic coverage than the WVS.

risk preference, there is a seemingly plausible proxy in the WVS, 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 Schwartz Values Survey (Schwartz, 2012), and designed to capture a universal "value of stimulation." A WVS survey item that seems to come closest to capturing altruism, albeit from a particular, societal perspective, asks respondents how similar they are to a hypothetical person for whom "[i]t is important [...] to do something for the good of society." The WVS also includes a well-known measure of trust, which we compare 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."

The Hofstede data set contains various cultural dimensions (Hofstede, 2001) that are composed of collections of qualitative survey items. Two cultural dimensions have labels that are evocative of time and risk preference, respectively: The "Long Term Orientation" (LTO) cultural dimension, which is occasionally used in economics, and the "Uncertainty Avoidance" dimension. Both of these measures include individual items that seem distant from either time or risk preference, but the data do not include responses to individual items so it is not possible to use a sub-set of items for preference proxies.[7]

Table 2 reports the country-level correlations. Starting with the WVS, the table shows that the questions for patience and altruism are only weakly correlated with the corresponding GPS preference measures. This might reflect the fact that the LTO item in the WVS is about childrearing rather than individual patience, and the altruism question is about a particular, societal perspective. On the other hand, the WVS value of stimulation measure, and the WVS trust question, are significantly positively correlated with the GPS risk preference and trust measures, respectively.[8] The Hofstede long term orientation and uncertainty avoidance dimensions also turn out to be significantly correlated with the corresponding GPS preference measures, perhaps surprisingly given that some of the individual items underlying the cultural dimensions appear far removed from preferences.

[7]The four items for long term orientation are: (1) value a person places on "doing a service to a friend;" (2) value of "thrift (not spending more than needed);" (3) agreement with "persistent efforts are the surest way to results;" (4) an item asking "How proud are you to be a citizen of your country?". The uncertainty avoidance dimension consists of: (1) "How often do you feel nervous or tense?" ; (2) "All-in-all, how would you describe your health these days?"; (3) agreement with "One can be a good manager without having a precise answer to every question that a subordinate may raise about his or her work;" (4) agreement with "a company's or organization's rules should not be broken – not even when the employee thinks breaking the rule would be in the organization's best interest."

[8]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 19 through 20 in Online Appendix I explore these relationships. 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 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, the WVS time preference proxy has the opposite gender difference to the GPS patience measure, and is negatively related to educational attainment at the individual level, and GDP at the country level, the opposite of what one would expect if it were to capture patience.

Table 2: Relationships between preference proxies in the WVS and Hofstede (2001) and GPS measures

	Spearman's rho	p-Value	Obs.
Correlations with GPS patience			
WVS Long Term Orientation	0.07	0.59	60
Hofstede Long Term Orientation	0.43	<0.01	60
Correlations with GPS risk taking			
WVS Value of Stimulation	0.32	0.03	48
Hofstede Uncertainty Avoidance	-0.34	0.01	60
Ccorrelation with GPS altruism			
WVS Altruism	0.19	0.27	35
Correlation with GPS trust			
WVS Trust	0.49	<0.01	60

In sum, the exercise helps distinguish weaker and stronger preference proxies in existing data sets. It provides evidence for the meaningfulness of the widely used WVS trust measure, and points to a potentially valuable proxy for risk preference for the WVS set of countries. The results also lend some support for using two of the Hofstede dimensions as preference proxies, although the latter are not available on the individual level or for representative samples.

The GPS also complements existing surveys in another way, as it contains data on the whole set of preferences, including positive and negative reciprocity, at the individual, regional, and country level. In contrast to existing datasets, this allows the exploration of the correlation structure among multiple preference dimensions, and investigations of how relationships between preferences and outcomes might be subject to omitted variable concerns because preferences are intra-correlated.

3 Distribution of Preferences Around the Globe

Figures 1 and 2 show how the country averages for each preference compare to the world average.[9] Each preference is normalized to have mean zero and standard deviation one in the individual-level data. Country averages are computed using sampling weights provided by Gallup. In each figure, white denotes countries in which the country-level average is within -0.05 to 0.05 of a standard deviation of the world mean. Darker blue indicates higher values of a given trait, while darker red colors indicate lower values. Grey countries are not included in the GPS. To provide a complementary perspective, Table 3 provides information on the average preferences for various groupings of countries.

[9]Country-level averages are calculated using the sampling weights provided by Gallup. The sampling weights ensure that our measures are representative of the population at the country level. See Online Appendix A.4.3 for more details.

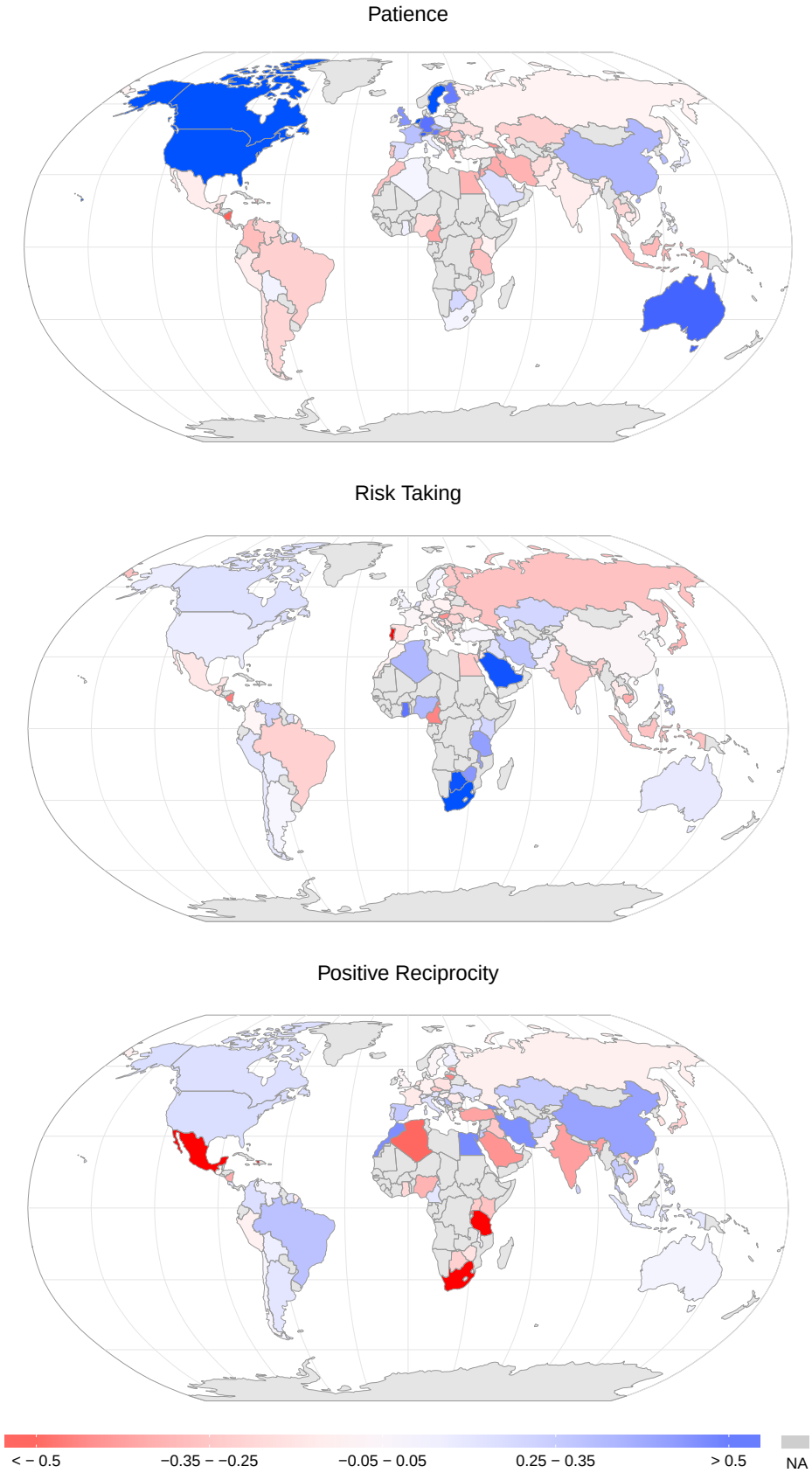

Figure 1: World maps of patience, risk taking, and positive reciprocity.

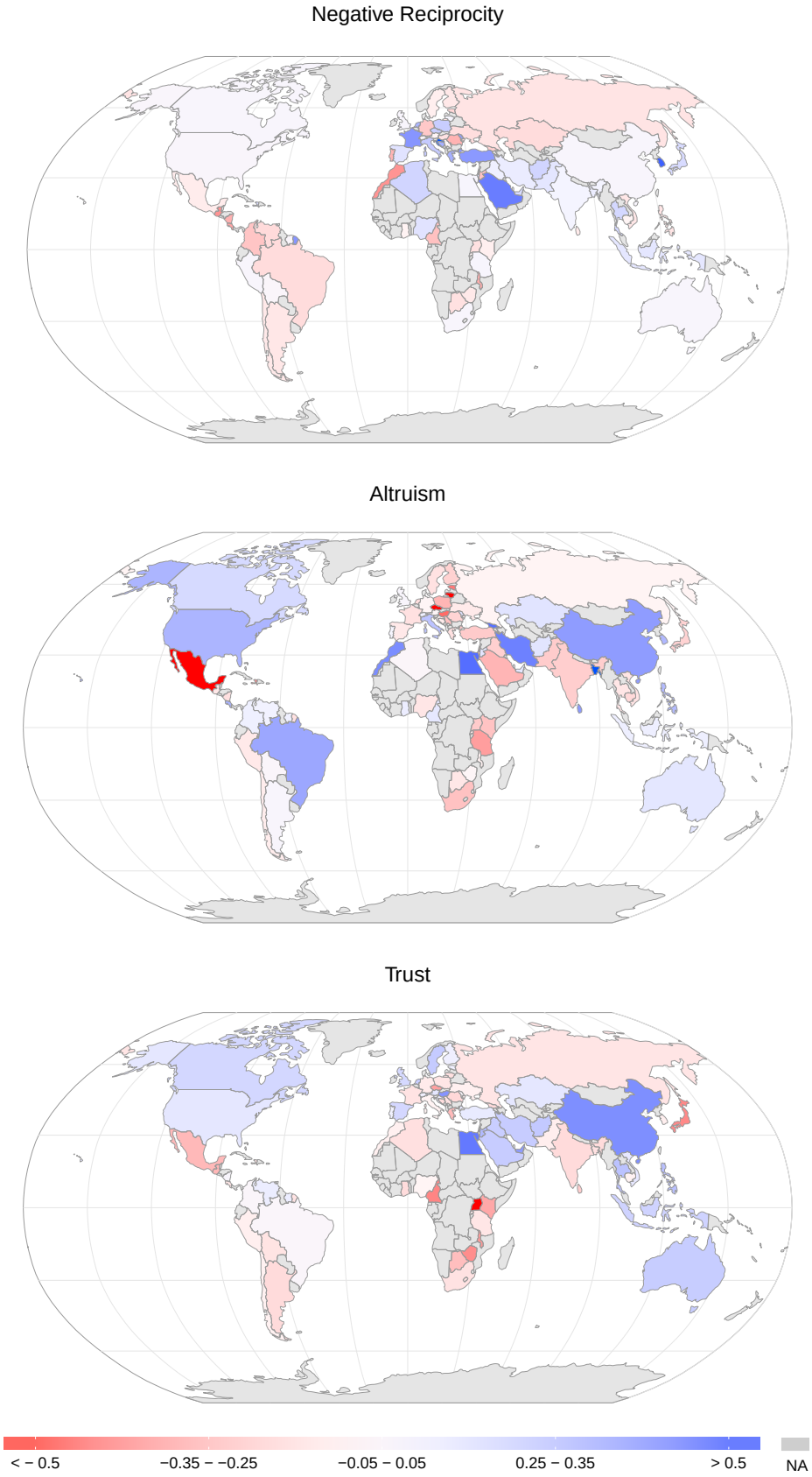

Figure 2: World maps of negative reciprocity, altruism, and trust.

Table 3: Regional averages and variance decomposition

	Patience	Risk taking	Pos. recip.	Neg. recip.	Altruism	Trust	# Obs.
Western Europe	0.49	-0.11	0.06	0.04	-0.04	0.10	11
Eastern Europe	-0.12	-0.12	-0.02	0.10	-0.22	-0.07	16
Neo Europe	0.73	0.15	0.16	0.02	0.26	0.23	3
South and East Asia	-0.01	-0.10	0.07	0.11	0.13	0.04	13
North Africa & ME	-0.14	0.16	0.07	0.08	0.13	0.23	9
Sub-Saharan Africa	-0.16	0.34	-0.34	-0.11	-0.15	-0.33	11
South America	-0.21	-0.03	-0.08	-0.16	-0.05	-0.10	13
% between-country variation	13.5	9.0	12.0	7.0	12.3	8.2	

Notes. Neo-Europe includes the United States, Canada, and Australia. Regional averages of each preference, expressed in terms of standard deviations from the world individual mean. The variance decomposition in the bottom row decomposes the individual-level variation 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. ME = Middle East.

The figures reveal that preferences vary substantially across countries, by at least one standard deviation for each preference.[10] Most differences 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.

In terms of patterns of preference variation, a first observation is that populations of European ancestry tend to be more patient than the world mean. Indeed, all of the ten most patient countries in the world are either located in the Neo-European, English-speaking world, or else in Western Europe, with the Scandinavian countries exhibiting particularly high levels of patience. Trust levels are also particularly high in Neo-Europe, while western European countries are notable for being relatively risk averse.

To the East, the former communist Eastern European countries are on average rather risk averse, not very patient, and low on altruism, but the patterns are less clear compared to their Western European counterparts. Countries in East and South Asia are also relatively impatient, except for "Confucian countries" (China, Japan, South Korea). This group of countries is consistently risk averse and relatively negatively reciprocal. On average, altruism is high, but patterns are diverse at the country level.

Middle Eastern and North African populations have in common relatively high levels of risk tolerance and low levels of patience. Social preferences and trust in 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, with most of these in sub-Saharan Africa; in addition, all sub-Saharan

[10] Online Appendix B.1 provide an alternative way to visualize the heterogeneity, with histograms of preferences at the country and individual levels.

Table 4: Pairwise correlations between preferences at country level

	Patience	Risk taking	Pos. reciprocity	Neg. reciprocity	Altruism	Trust
Patience	1					
Risk taking	0.230**	1				
Pos. reciprocity	0.0155	-0.256**	1			
Neg. reciprocity	0.258**	0.193*	-0.154	1		
Altruism	-0.0104	-0.0152	0.711***	-0.132	1	
Trust	0.190	-0.0615	0.363***	0.160	0.273**	1

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

populations are on average lower than the world mean on positive reciprocity, altruism, and trust, and are rather impatient. Finally, in the Southern Americas, most populations appear impatient, and low in terms of negative reciprocity. They have more intermediate values in risk taking and the prosocial traits, i.e., altruism, positive reciprocity, and trust.

In sum, 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 4 shows Pearson correlations of preferences together with levels of significance.

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.

Another grouping of positively correlated traits involves prosociality, i.e., the traits of positive reciprocity, altruism and trust. While trust constitutes a belief rather than a preference, all of these traits share in common that they describe positive behavioral dispositions towards others. The correlation between altruism and positive reciprocity is particularly high, and trust also tends to be higher where people are positively reciprocal. This is intuitive as it is hard to imagine high levels of trust absent positive reciprocity, i.e., trust-rewarding behaviors.

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 Online Appendix C.

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. The last row of 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. Across preferences, about 10% of the total variation is due to between-country variation. Part of the within-country variation might

reflect measurement error, so that the variation in true preferences is overstated. 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 Online Appendix D for details.[11]

4 Determinants and Cultural Correlates of Preferences

4.1 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. Indeed, a large literature in behavioral economics has investigated the relationships between these traits and preference variation, mostly for two reasons. First, they are associated with differences in economic outcomes; if preferences vary with these traits, this could be part of the explanation for outcome differences.[12] 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.[13] Because most previous evidence on preferences has come from individual countries, or non-representative samples, the GPS provides new insights into which relationships might reflect mechanisms that are more universal, and which might be specific to certain societies. For instance, the origins and universality of gender differences across cultures remain an open question (e.g., Gneezy et al., 2009; Niederle, 2014).

For the purposes of our analysis, we make use of the sociodemographic covariates contained in the GPS. As a proxy for cognitive skills, our dataset contains a measure of self-reported math skills that we use to proxy for cognitive skills. While this is an imperfect proxy for cognitive ability, there is evidence that math skills are correlated with cognitive ability in general (Borghans et al., 2016), that subjective assessments of ability are correlated with measured cognitive ability, and that these have predictive power for academic achievement (Marsh, 1990; Marsh et al., 2005; Spinath et al., 2006; Ackerman and Wolman, 2007; Chamorro-Premuzic et al., 2010). Since such relative self-assessments might be interpreted in different ways across countries, we only use self-reported cognitive skills for within-country analyses.

Table 5 presents OLS regression estimates for how preferences are related to gender, cogni-

[11] The between-country variation should be scaled up by the inverse of the fraction of variance that is due to measurement error. A very conservative estimate of the test-retest correlation of a given preference measure that is as low as 0.33 implies that the between-country variation is about three times as high as reported in Table 3. Since test-retest correlations between 0.5 and 0.6 are typically found for single items that constitute our measures, we are confident that the between-country variation does not exceed 50 percent.

[12] See, e.g., Barsky et al. (1997); Donkers et al. (2001); Frederick (2005); Sutter and Kocher (2007); Croson and Gneezy (2009); Dohmen et al. (2010, 2011); Benjamin et al. (2013).

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

tive ability, and age across the GPS sample. We report results with country fixed effects as well as subnational region fixed effects. The preference variables are standardized so coefficients are in units of standard deviations. In Online Appendix E we show that the results are robust to adding a set of additional control variables.

Starting with time preference, Table 5 documents 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. There is limited previous cross-country evidence on time preference, but the small gender difference we find is in line with a cross-country study on college students.[14] 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. There is little previous evidence, from cross-country or representative data, on how patience varies with age.

Turning to risk preference, Table 5 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.[15] Previous studies have also found a similar relationship between risk aversion and cognitive ability, for a few countries. A similar shaped age profile in risk preference has been documented previously, for individual countries.[16]

Social preferences and trust also vary significantly with individual characteristics. Table 5 shows that positive reciprocity and altruism are more pronounced among women, while negative reciprocity is weaker among women. 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.[17] Some previous studies have also found a positive relationship between cognitive ability and altruism, using student subjects (e.g. Chen et al., 2013). Finally, the results

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

[15] 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. In meta-analyses, females tend to be more risk averse in the majority of studies (Byrnes et al., 1999; Croson and Gneezy, 2009), but effect sizes are heterogeneous, and roughly 40% of studies do not find a gender difference (Niederle, 2014). The mixed results across studies could potentially reflect small samples (Niederle, 2014).

[16] E.g., Dohmen et al. (2011) show that willingness to take risks declines with age in a representative sample of German adults. See also Dohmen et al. (forthcoming). Mata et al. (2016) show that the WVS measure of "value of stimulation" declines with age.

[17] 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.

Table 5: Correlates of preferences at individual level

	Dependent variable:											
	Patience		Risk taking		Pos. reciprocity		Neg. reciprocity		Altruism		Trust	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1 if female	-0.056***	-0.060***	-0.17***	-0.18***	0.049***	0.053***	-0.13***	-0.14***	0.10***	0.097***	0.066***	0.059***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Subj. math skills	0.028***	0.026***	0.046***	0.043***	0.038***	0.039***	0.040***	0.035***	0.044***	0.041***	0.056***	0.054***
	(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	0.72***	0.76***	-0.083	-0.080	1.02***	1.10***	-0.36*	-0.39**	-0.0061	-0.0090	0.37*	0.28
	(0.17)	(0.17)	(0.20)	(0.20)	(0.17)	(0.16)	(0.19)	(0.18)	(0.14)	(0.14)	(0.21)	(0.20)
Age squared	-1.45***	-1.50***	-1.20***	-1.19***	-1.17***	-1.27***	-0.45**	-0.40**	0.015	0.010	0.032	0.11
	(0.20)	(0.20)	(0.21)	(0.21)	(0.18)	(0.16)	(0.18)	(0.18)	(0.15)	(0.16)	(0.20)	(0.19)
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
Observations	78501	77585	78445	77543	78869	77949	77521	76718	78632	77721	77814	76922
R^2	0.17	0.21	0.17	0.22	0.13	0.22	0.11	0.18	0.13	0.19	0.11	0.17

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). Age is divided by 100. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

on trust in Table 5 are broadly in line with evidence from the trust literature.

We turn next to a country-level analysis, to see whether the aggregate results in Table 5 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 3 shows the gender coefficients for the different countries, with a separate panel for each preference. Figure 4 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 5 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 division of countries captures some of the most salient cross-country differences or commonalities.

Beginning with time preference, Figure 3 shows that the slightly larger degree of impatience among women, at the aggregate level, conceals substantial heterogeneity. Only 68 percent of countries have a coefficient indicating greater impatience for women, and only 32 percent have a statistically significant difference (p-value < 0.1) in that direction. Figure 4 indicates, by contrast, that the relationship of patience with cognitive ability goes in the same direction, and is statistically significant, in almost all countries. This suggests that the relationship is relatively universal, and arguably not the product of institutions or specific educational and value systems. In Figure 5 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 3 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 reveals the widespread prevalence of the gender difference in risk preference, in qualitative terms, across a wide range of cultures and on a representative basis. Figure 4 shows that in almost all countries, lower cognitive ability is associated with significantly greater risk aversion. The age profiles in Figure 5 imply that risk tolerance is decreasing 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.

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 3 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, Figure 4 shows that positive reciprocity is associated with higher cognitive ability irrespective of culture. In terms of age profiles, Figure 5 reveals another difference across societies: The profile for positive reciprocity is hump-shaped for OECD countries, but less so for non-OECD countries.

Figure 3 shows that altruism and negative reciprocity are related to gender in opposite ways across countries, in line with the aggregate results. In most countries, altruism is more pronounced among women, whereas negative reciprocity is less pronounced. Altruism and neg-

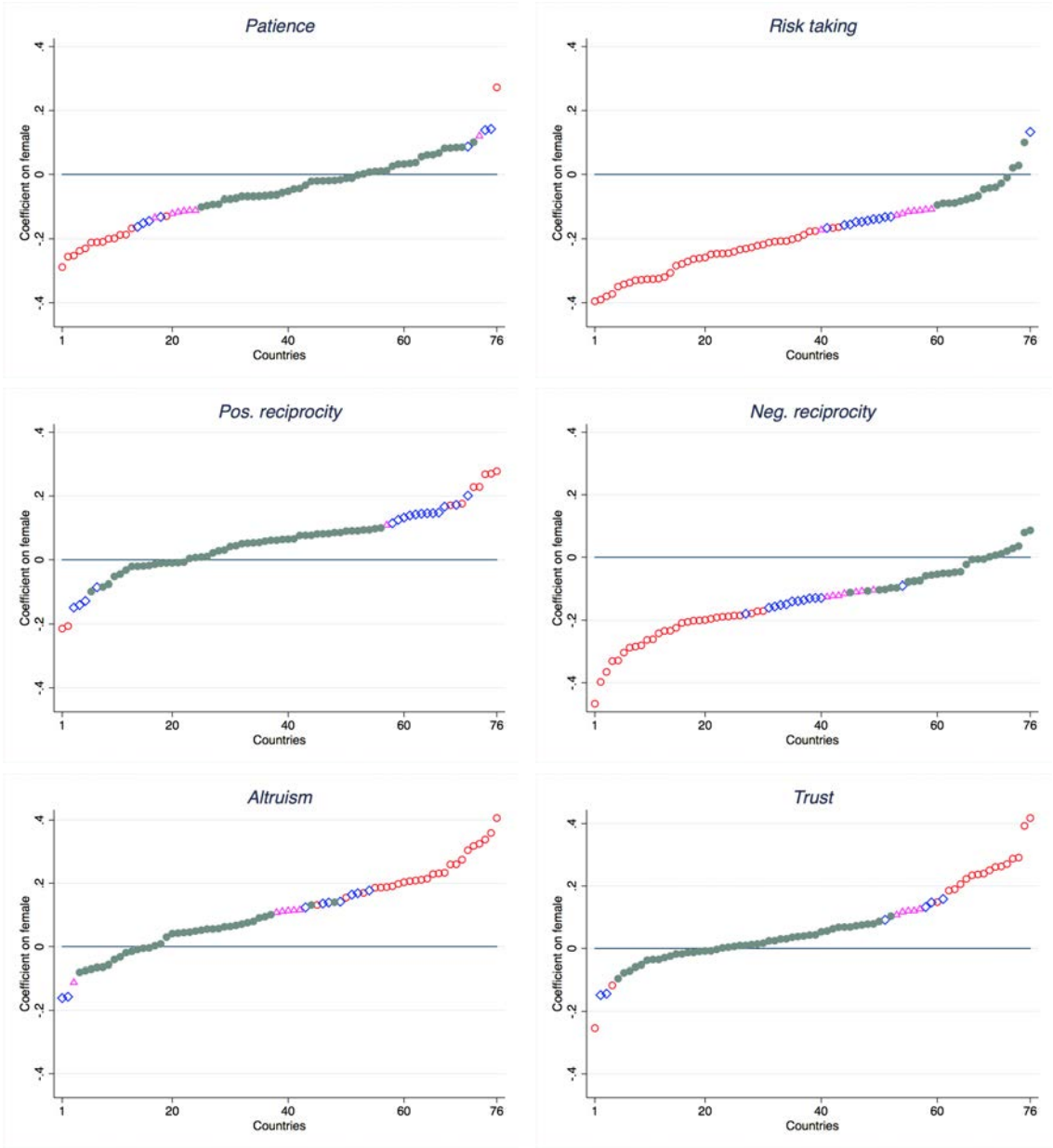

Figure 3: Gender coefficients by country. For each country, we regress the respective preference on gender, age and its square, and subjective math skills, and plot the resulting gender coefficients as well as their significance level. In order to make countries comparable, each preference was standardized (z-scores) within each country before computing the coefficients. Solid green diamonds indicate countries in which the gender coefficient is not statistically different from zero at the 10% level, while red dots / blue diamonds / pink triangles 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.

ative reciprocity are both associated with higher cognitive ability in almost every country, as seen in Figure 4. Figure 5 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 3, 4, and 5 show that the aggregate results on trust are largely born out in

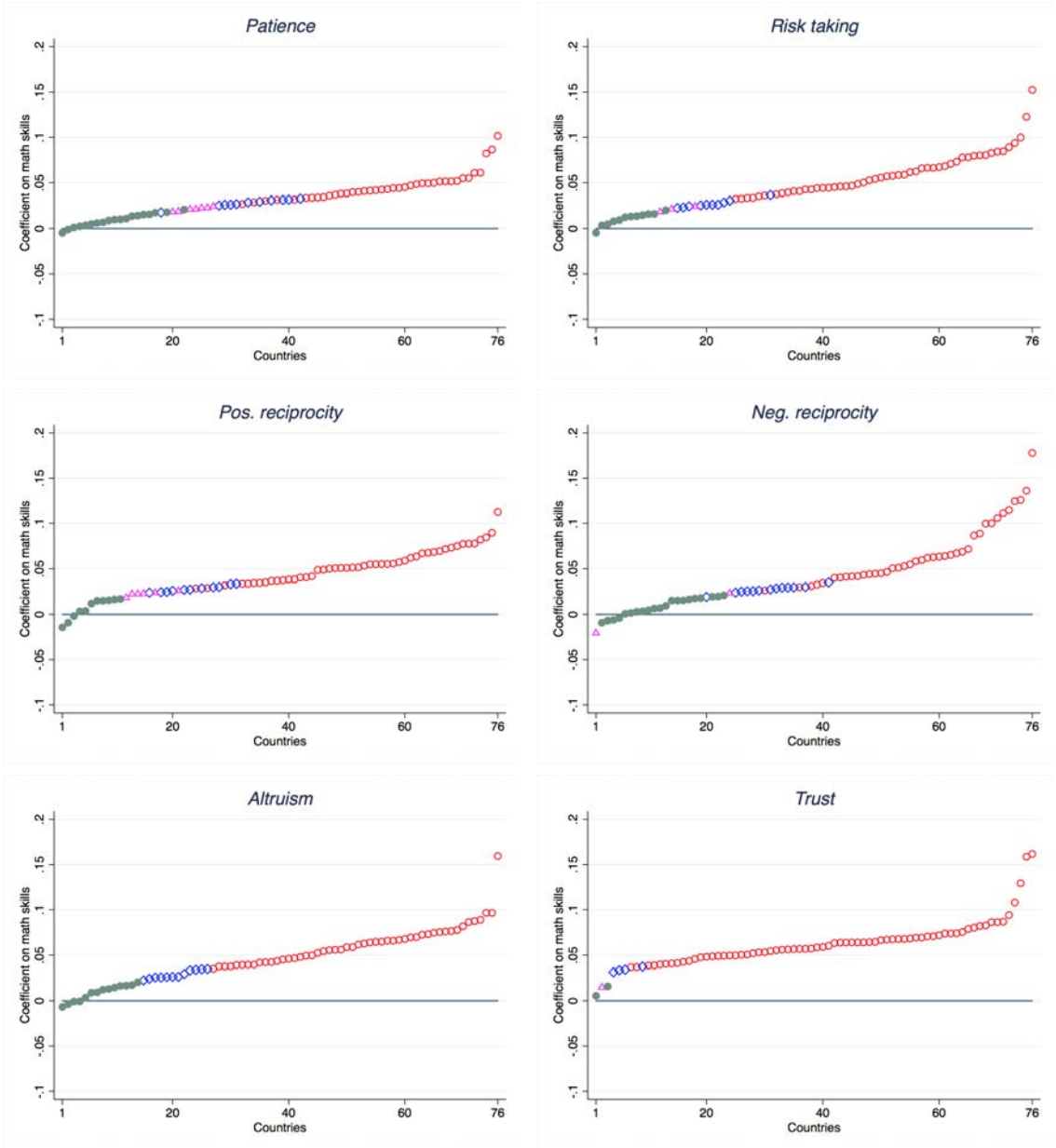

Figure 4: Cognitive ability coefficients by country. For each country, we regress the respective preference on gender, age and its square, and subjective math skills, and plot the resulting coefficients on subjective math skills as well as their significance level. In order to make countries comparable, each preference was standardized (z-scores) within each country before computing the coefficients. Solid green diamonds indicate countries in which the cognitive ability coefficient is not statistically different from zero at the 10% level, while red dots / blue diamonds / pink triangles denote countries in which the effect is significant at the 1% / 5% / 10% level, respectively. Positive coefficients imply that higher cognitive ability is associated with higher values in the respective preference.

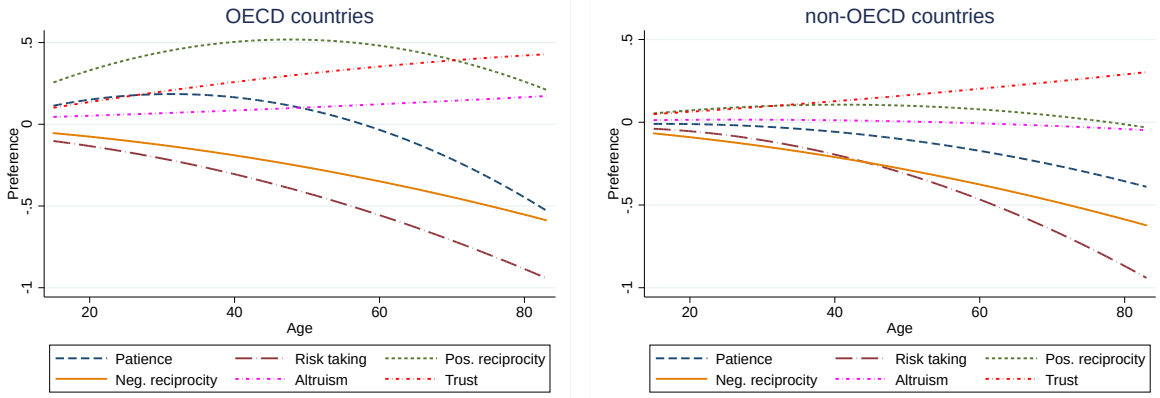

Figure 5: Age profiles by OECD membership. The figures depict the relationship between preferences and age conditional on country fixed effects, gender, and subjective math skills. These are augmented component plus residuals plots, in which the vertical axis represents the component of the preference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

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. In almost all countries, trust is increasing with cognitive ability, and trust 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 countries. There are other relationships, however, such as between time preference and gender, or positive reciprocity and age, for which the qualitative relationships differ substantially across countries. These latter findings point to cases where results from one country might not generalize to others, and where the underlying mechanisms might be sensitive to cultural differences.

4.2 Geographic and Cultural Correlates

To unpack the nature of country-level variation, this section relates preferences to a set of geographic and cultural variables that have been proposed as potential determinants of preferences in the literature. While the results presented here are to be understood as simple raw correlations, they nonetheless speak to previously articulated narratives, hypotheses, or empirical results. For example, various authors have proposed that the evolution of time preference and trust is related to geographic conditions (Galor and Özak, 2016; Litina, 2016), and that negative reciprocity is intimately linked to biological endowments as in the "culture of honor" hypothesis of Nisbett and Cohen (1996). The GPS allows for a comprehensive evaluation of these hypotheses using experimentally validated survey measures of preferences.

The analysis is divided into (bio-) geographical characteristics, which are more likely to be

exogenous to preferences, and cultural variables that are potentially endogenous to preferences, at least in the long run. The first five rows of Table 6 present the Pearson correlations between all preferences and the following geographic conditions: (i) a summary statistic of geographic conditions proposed by Spolaore and Wacziarg (2013), which consists of the first principal component of absolute latitude, agricultural suitability, rate of East-West orientation, and size of landmass, coded such that the first component is positively correlated with per capita income (Olsson and Hibbs, 2005); (ii) a summary statistic of biological conditions, which is the principal component of number of annual or perennial wild grasses and number of domesticable big animals, again coded such that the first component is positively correlated with per capita income (Olsson and Hibbs, 2005); (iii) distance from the equator, (iv) agricultural suitability, adjusted for post-Columbian migration flows using the migration matrix of Putterman and Weil (2010) and (v) a recently developed index of crop suitability, also ancestry-adjusted (Galor and Özak, 2016).

First focusing on the geography summary statistic of the variables in Olsson and Hibbs (2005), we see that most preferences are significantly related to geographic conditions. For example, patience, negative reciprocity and trust are all positively correlated with those geographic variables that have previously been argued to be conducive for economic development (Diamond, 2005; Spolaore and Wacziarg, 2013).

While the Olsson-Hibbs summary statistic has the advantage of capturing various dimensions of geography, its interpretation as principal component is not fully transparent, so we also report separate correlations between preferences and distance to the equator as well as agricultural suitability indices. The results show that patience, negative reciprocity and trust all increase in distance from the equator. On the other hand, the results show that agricultural suitability and crop suitability are only very weakly correlated with the GPS patience variable.[18] However, both agricultural and crop suitability are significantly negatively correlated with trust. All other preferences are largely uncorrelated with agricultural suitability.

Turning to the Olsson-Hibbs summary statistic of biological endowments, patience, negative reciprocity and trust all positively covary with biological conditions. Further unpacking these relationships, negative reciprocity is positively correlated with the number of large domesticable animals, broadly in line with the culture of honor hypothesis (Nisbett and Cohen, 1996; Grosjean, 2014). On the other hand, those bio-geographic conditions that are conducive to development are negatively correlated with risk taking.

Next, we study correlations between preferences and cultural variables that have been proposed as potential drivers of preference variation. We consider linguistic structures, religion, individualism, and family ties. First, various recent papers have argued that language might shape people's preferences and behaviors (Tabellini, 2008; Chen, 2013; Sutter et al., 2014; Galor et al., 2017). In particular, a linguistic feature called weak future time reference (FTR) has attracted attention because it correlates with future-oriented decisions. This linguistic variable

[18] Galor and Özak (2016) find a correlation between Hofstede's Long Term Orientation variable and crop suitability for agriculture.

Table 6: Pairwise correlations between preferences and geographic and cultural variables

	Patience	Risk taking	Pos. recip.	Neg. recip.	Altruism	Trust	# Obs.
(Bio-) geography							
Geographic conditions (O-H)	0.45***	-0.31**	0.20	0.39***	-0.10	0.39***	51
Absolute latitude	0.48***	-0.19	0.13	0.25**	-0.13	0.26**	76
Agricultural suitability (aa)	-0.02	-0.14	0.03	0.03	-0.22*	-0.47***	73
Crop suitability (aa)	0.11	-0.18	-0.11	0.08	-0.22*	-0.37***	73
Biological conditions (O-H)	0.37***	-0.37***	0.30**	0.30**	-0.00	0.44***	51
Culture							
Weak future time reference	0.32***	-0.13	0.13	-0.04	0.07	0.21*	68
Pronoun drop not allowed	0.57***	0.08	0.04	0.06	0.01	0.18	67
Share Protestants	0.45***	0.10	-0.20*	-0.17	-0.14	-0.01	76
Individualism	0.65***	-0.01	-0.03	0.14	-0.15	0.15	60
Family ties	-0.57***	0.34**	0.11	-0.02	0.27*	0.09	49

Notes. Pairwise Pearson correlations between average preferences and other geographic and climatic variables at country level. See Online Appendix J for additional information about the variables. In analyses with language variables, the sample only includes countries for which we could classify the interview language of at least 50% of our respondents. Geographic and biological conditions are the first principal components of the geography and biological variables in Olsson and Hibbs (2005), also see Spolaore and Wacziarg (2013). (aa) = ancestry-adjusted. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

assumes a value of one if a given language allows to speak about the future in present tense, and zero otherwise. Table 6 shows that in our data patience is also highly and significantly correlated with weak FTR.[19] Moreover, weak FTR is postively correlated with trust. Second, we study the linguistic feature of pronoun drop, which was originally used by Tabellini (2008). This variable assumes a value of one if a language does not allow to drop pronouns, which is hypothesized to invoke a stronger emphasis on individual needs as opposed to those of other people. However, perhaps in contrast with this notion, we find that pronoun drop is uncorrelated with all of the social preferences. Instead, it is strongly correlated with patience.[20]

A prominent hypothesis in the social sciences is Weber's (1930) argument of a "Protestant ethic", which, among other aspects, is believed to have made people more patient. We investigate this argument on a correlational basis by relating our patience measure to the share of Protestants in a given country (Barro, 2003). Consistent with Weber's hypothesis, we find that Protestantism is strongly correlated with patience. This correlation is robust to restricting

[19]For this analysis, we made use of the classification by Chen (2013), with minor additions and changes. First, we set Persian to missing after corresponding with him (he originally classified Persian as strong FTR, which is open to discussion). Second, we managed to classify Moroccan Arabic (strong), Fula (strong), and Khmer (weak).

[20]Online Appendix F investigates the relationship between preferences and FTR and pronoun drop at the individual level within countries by exploiting within-country variation in interview language. In these analyses, weak FTR is again significantly associated with patience and trust, and additionally with altruism and positive reciprocity. The linguistic feature of no pronoun drop is also again correlated with patience, and in addition, with risk taking and altruism.

attention to Europe or predominantly Christian countries.

Finally, we turn our attention to variables that measure aspects of social and family structure. Hofstede (2001) proposes a measure of individualism that has subsequently been used in economics (Gorodnichenko and Roland, 2011). Likewise, Alesina and Giuliano (2013) extract a measure of family ties from the WVS which measures the importance of the family relative to other aspects in life. Intuitively, it seems possible that family ties and individualism are related to social preferences. However, we find only weak evidence for such relationships. While family ties are correlated with altruism, all other correlations are not significant. Instead, individualism and family ties are also both correlated with patience.

In sum, this section has brought to light that the distribution of preferences across countries is not random, but rather follows geographic and cultural patterns. In particular, patience is strongly related to many cultural and geographic conditions, and future research might tap further into the potential of the GPS to illuminate potential causal channels.

5 Preferences and Outcomes

This section investigates the relationship of economic outcomes to preferences. The focus is on outcomes that previous literatures have hypothesized might depend on a particular preference or set of preferences.

5.1 Preferences and Individual Outcomes

5.1.1 Accumulation Decisions

Economic theory suggests that patience is instrumental for savings and investments in human capital. We evaluate the relationship of the our patience measure to these outcome variables in the GPS. Columns (1) and (2) of Table 7 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 in specifications with country and subnational region 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 15% 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).[21] In Online Appendix G.2, we show that the significant relationship between our patience variable and accumulation processes is not driven by only a few countries. Rather, the coefficient of patience is positive in more than 90% of countries for both savings and education, and in most cases statistically significant.

[21] All results are robust to using (ordered) probit estimations.

Table 7: Patience and accumulation decisions, risk preferences and risky choices

	Dependent variable:									
	Accumulation decisions				Risky choices					
	Saved last year		Education		Own business		Plan start business		Smoking int.	
	(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)
Country FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Region FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	15260	14459	79357	68409	72839	62125	57072	50687	15309	14490
R^2	0.07	0.18	0.21	0.36	0.06	0.14	0.11	0.17	0.03	0.23

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). Additional controls include age, age squared, gender, subjective math skills, log household income, and indicators for religious affiliation. See Online Appendix J for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.1.2 Risky Choices

We next investigate the relationship of risk preferences to behaviors that have been hypothesized to depend on a taste for risk. Specifically, the career choice of being self-employed as well as the risky health behavior of smoking have been modeled as depending on sufficient willingness to take risks (e.g., Kihlstrom and Laffont, 1979; Viscusi and Hersch, 2001). As columns (5) and (6) of Table 7 document, our preference measure is related to actual self-employment. The same pattern holds when considering individuals' intention to start their own business, conditional on not being self-employed (columns (7)-(8)). Columns (9) and (10) relate risk preferences to the respondent's smoking intensity, measured on a three-point scale (never, occasionally, and frequently). We find that more risk-tolerant people are more likely to smoke, both with country and subnational region fixed effects, and conditional on a set of additional covariates. Online Appendix G.2 again shows that the correlations between risk preferences and labor market or health decisions are qualitatively similar across countries.

5.1.3 Social Interactions

Next, we analyze the relationships of the social preference measures to behaviors and outcomes in the social domain. 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.[22]

Table 8 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,

[22]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.

Table 8: 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***	0.038***	0.036***	0.056***	0.052***	0.036***	0.032***	0.025***	0.022***	0.020***	0.017***	-0.0025	0.0031
	(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	0.0010	0.0046	0.0060*	0.0022	0.038***	0.034***	0.019***	0.020***	0.00095	-0.0015	0.018***	0.016***	0.012***	0.0083***
	(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)
Trust	0.0088**	0.0054	0.0077**	0.010***	-0.0077**	-0.0066*	0.0017	0.0020	0.0033	0.0023	0.0011	0.0020	0.010***	0.0024
	(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	-0.0059*	-0.0027	-0.00031	-0.0014	0.0062	-0.0022	0.0088**	0.0030	0.021***	0.017***	0.0096**	0.00064	-0.012***	0.00038
	(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)
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
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	57456	52124	57439	52116	55233	51919	55519	52276	55191	51873	65117	57835	76888	66693
R^2	0.18	0.24	0.08	0.14	0.08	0.15	0.11	0.18	0.05	0.10	0.09	0.17	0.05	0.24

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

or sending money or goods to other people in need. Across the different behavioral categories, the point estimate is very consistent and implies that an increase in altruism by one standard deviation is correlated with an increase in the probability of engaging in prosocial activities of 3.5–6.5 percentage points, which corresponds to an increase of roughly 15–20% compared to the respective baseline probabilities.[23] Positive reciprocity is a significant correlate of helping people in need (columns (5) through (8)), perhaps a manifestation of generalized reciprocity in the sense that reciprocal people who have been helped before are also willing to help others. In contrast, the negative reciprocity variable is virtually uncorrelated with all of the prosocial activities in the first eight columns. As columns (9) and (10) show, however, negative reciprocity is a significant predictor of whether people are willing to voice their opinion to a public official. Columns (11) through (14) examine the relationship between social preferences and respondents' family and friendship 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.

The overall pattern in Table 8 is that the social preference measures are related to a wide range of behaviors in the social domain. As Online Appendix G.2 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 16 and 17 in the Online 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 all preferences, 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. 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 heterogeneous sample. In this sense, the correlations provide an important out-of-context validation check for the survey module.

5.2 Preferences and Country-Level Outcomes

This section explores the correlation between preferences and outcomes at the country level, again focusing on outcomes that previous literatures have hypothesized might be endogenous to particular preferences.

[23] These baseline probabilities are 31.8%, 21.6%, 48.3%, and 23.7%, respectively (see Table 8, columns (1) - (8), for the order of variables).

Table 9: Economic development and preferences

	Dependent variable: Log [GDP p/c]									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Patience	2.63***	1.73***							2.67***	1.92***
	(0.26)	(0.27)							(0.29)	(0.31)
Trust			1.58**	0.56					0.73	0.31
			(0.68)	(0.48)					(0.56)	(0.45)
Risk taking					-0.53	0.59*			-1.34***	-0.53
					(0.56)	(0.33)			(0.50)	(0.39)
Negative reciprocity							1.30**	0.51	0.54	0.092
							(0.51)	(0.50)	(0.52)	(0.45)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	76	73	76	73	76	73	76	73	76	73
R^2	0.39	0.70	0.08	0.59	0.01	0.59	0.05	0.59	0.48	0.71

OLS estimates, robust standard errors in parentheses. Controls include distance to equator, average temperature, average precipitation, the share of the population living in (sub-)tropical zones, terrain ruggedness, average distance to nearest waterway, and an island dummy. See Online Appendix J for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.2.1 Patience, Trust and Economic Development

We begin by investigating whether variation in per capita income across countries is related to variation in those preferences that previous literatures or models have highlighted as potential drivers of development. This includes time preference, as many models of economic development such as standard Ramsey-Cass-Koopmans models involve a key role for time preference. Another literature, on social capital, has emphasized that trust may play an important role in development (Knack and Keefer, 1997; La Porta et al., 1997; Algan and Cahuc, 2013). Research in anthropology and behavioral economics has lead to the hypothesis that sanctioning of inefficient behaviors, driven by negative reciprocity, may help sustain large scale cooperation and hence generate efficient outcomes (Fehr and Gächter, 2002; Boyd et al., 2003; Henrich et al., 2006). Finally, willingness to take risks has been found previously to be correlated with income at the individual level (Barsky et al., 1997; Dohmen et al., 2011).

Columns (1) and (2) of Table 9 provide evidence that patience is strongly correlated with per capita income, in specifications with and without geographic controls. In a statistical sense, patience "explains" 40% of the variation in income. Columns (3) and (4) establish that the GPS trust measure is also significantly correlated with per capita income, yet this correlation is no longer statistically significant once controls are accounted for. Columns (5) and (6) show that willingness to take risks is uncorrelated with per capita income, but weakly positively correlated once controls are accounted for (this correlation is not very robust across specifications). Columns (7) and (8) document that between negative reciprocity and per capita income are significantly correlated, yet this correlation loses significance once controls are accounted for.

Finally, columns (9) and (10) show a "horse race" between the set of preferences that have been linked conceptually to development. The results show that patience is the only variable

that is robustly correlated with per capita income. The insight that patience "outperforms" trust in the GDP regressions is robust to using the standard WVS trust question as opposed to the GPS trust variable. These results appear noteworthy given the strong emphasis in the previous literature on the importance of trust, and provide a first piece of evidence for the potential of the GPS in furthering understanding of the relationship between development and preference or belief variables.[24]

Although the main focus of the analysis is investigating correlations, rather than maximizing predictive power, it is noteworthy that patience contributes substantially to explained variation, above and beyond standard geographic variables. This can be seen comparing the R^2 with preferences included, to the R^2 from a regression on controls alone (0.58). Adding all preferences increases explained variation by 13 percentage points and adding patience alone increases the R^2 by 12 percentage points.

5.2.2 Risk Taking and Risky Entrepreneurial Activities

Turning to risk preference, previous literatures have hypothesized that willingness to take risks may drive entrepreneurship, and have also shown evidence of a link between risk preference and self-employment at the individual level. Columns (1)–(6) of Table 10 investigate the relationship between the GPS risk taking variable and different proxies for risky entrepreneurial activities at the country level.[25] Specifically, as dependent variables, the analysis uses the number of patent applications per capita, the number of scientific articles published in a given country per capita, and total factor productivity as a measure of the stock of ideas and knowledge.

The results reveal that risk taking is uncorrelated with patent applications, but risk taking is significantly correlated with the number of scientific articles per capita and Total Factor Productivity (TFP), once the confounding effects of the geographic and climatic covariates are taken into account. The increase in R^2 from adding risk taking to the set of covariates is small, but notable.[26]

5.2.3 Social Preferences, Charitable Activities and Conflict

Finally, the analysis explores the country-level correlations between the social preferences and outcomes that are conceptually linked to the respective preferences. A first dependent variable is the dollar value of charitable donations and volunteering activities, as a fraction of GDP (Salamon, 2004). Given the many studies showing that social preferences and charitable giving are correlated at the individual level, it is natural to explore whether cross-country variation

[24]The relationship of patience to GDP remains strong and significant with positive reciprocity and altruism in the regression as additional controls.

[25]Cross-country data on self-employment are not very meaningful for our purposes because self-employment may refer to very different business concepts in developed and developing economies.

[26]Online Appendix H shows that results are similar with social preferences included in the regression as controls. Adding patience, however, causes risk taking to no longer be statistically significant, whereas patience is significantly positively related to entrepreneurial activities.

Table 10: Country-level outcomes and preferences

	Dependent variable:									
	Entrepreneurship						Social outcomes			
	Patent applic. p/c		Scientific articles p/c		TFP		Volunt. & donat.		Armed conflicts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Risk taking	-0.031	0.28	-0.013	0.094**	0.11	0.22**				
	(0.98)	(0.43)	(0.05)	(0.05)	(0.10)	(0.09)				
Prosociality							0.85	1.23**		
							(0.57)	(0.48)		
Negative reciprocity									1.59***	1.20***
									(0.41)	(0.41)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	64	61	69	67	60	59	32	32	76	73
R^2	0.00	0.66	0.00	0.44	0.02	0.49	0.06	0.42	0.13	0.32

OLS estimates, robust standard errors in parentheses. The dependent variables in columns (1)–(2) and (3)–(4) are the logs of the number of patent applications p/c and the number of scientific articles p/c, respectively. In columns (7)–(8), the dependent variable is volunteering and donation as a fraction of GDP. Frequency of conflicts is measured by the log of conflicts according to PRIO, in the Quality of Government dataset. Prosociality is the first principal component of altruism, positive reciprocity, and trust. Controls include distance to equator, average temperature, average precipitation, the share of the population living in (sub-)tropical zones, terrain ruggedness, average distance to nearest waterway, and an island dummy. See Online Appendix J for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in charitable activity might be related to variation in average pro-sociality at the population level. Since altruism, positive reciprocity and trust are highly correlated at the country level, we collapse these variables into a single "prosociality" variable by computing the first principal component at the individual level and aggregating this score at the country level. Columns (7) and (8) show that this score is significantly correlated with donations and volunteering once the baseline set of controls is taken into account.

Second, motivated by research at the individual level on how punishment can trigger conflict in the form of vengeful counter-punishment (Herrmann et al., 2008; Nikiforakis, 2008), the analysis correlates average negative reciprocity of a country's population with the log of the frequency of armed conflicts. The conflict variable is based on the Peace Research Institute Oslo (PRIO) dataset. Columns (9) and (10) show that countries with a higher degree of negative reciprocity have experienced significantly more armed conflicts, conditional on controls. Here, the raw correlation is particularly pronounced ($\rho = 0.37$), and the inclusion of negative reciprocity leads to an increase in R^2 of seven percentage points relative to the set of controls.[27]

6 Conclusion

The evidence in this paper shows that (i) preferences exhibit large heterogeneity across and within countries, (ii) this variation is at least partly systematic and linked to both individual-level characteristics and aggregate cultural or biogeographic endowments, and (iii) the survey measures of preferences appear to capture heterogeneity that is relevant for explaining out-

[27]Online Appendix H shows that prosociality and negative reciprocity are still significantly related to charitable activities, and frequency of conflicts, when time and risk preference are included as controls.

comes. These findings are only a first step towards tapping the potential of the GPS. The data are well suited for many potential research agendas, on the determinants and implications of preference variation. One example is deepening understanding of the observed correlation structure for preferences across countries, investigating which mechanisms could potentially be involved in the co-evolution of different preference combinations. Another direction is exploring in more detail the nature of individual differences in preferences, for example, whether gender differences in preferences are related to measures of the degree of female empowerment across societies. Differences in how preferences relate to individual economic outcomes across countries could potentially be understood from the perspective of how preferences interact with institutional differences. Finally, the relationship between country-level preference profiles and aggregate economic outcomes is essentially unchartered territory. In this respect, the paper has provided evidence of some novel raw correlations, e.g., between per capita income and time preference, or negative reciprocity and conflicts, which call for a more detailed analysis of the underlying causal pathways.

References

Ackerman, Phillip L. and Stacey D. Wolman, "Determinants and Validity of Self-Estimates of Abilities and Self-Concept Measures," *Journal of Experimental Psychology: Applied*, 2007, *13* (2), 57–78.

Åkerlund, David, Bart H.H. Golsteyn, Hans Grönqvist, and Lena Lindahl, "Time Discounting and Criminal Behavior," *Proceedings of the National Academy of Science (PNAS)*, 2016, *113* (22), 6160–6165.

Alesina, Alberto and Paola Giuliano, "Family Ties," *Handbook of Economic Growth*, 2013, *2*, 177.

_ , 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.

_ and _ , "Trust and growth," *Annu. Rev. Econ.*, 2013, *5* (1), 521–549.

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.

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

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

Beauchamp, Jonathan, David Cesarini, and Magnus Johannesson, "The Psychometric 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.

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, Bart H.H. Golsteyn, James J. Heckman, and John Eric Humphries, "What Grades and Achievement Tests Measure," *Proceedings of the National Academy of Sciences*, 2016, *113* (47), 13354–13359.

Boyd, Robert, Herbert Gintis, Samuel Bowles, and Peter J. Richerson, "The Evolution of Altruistic Punishment," *Proceedings of the National Academy of Sciences*, 2003, *100* (6), 3531–3535.

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

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.

Chamorro-Premuzic, Tomas, Nicole Harlaar, Corina U. Greven, and Robert Plomin, "More Than Just IQ: A Longitudinal Examination of Self-Perceived Abilities As Predictors of Academic Performance in a Large Sample of UK Twins," *Intelligence*, 2010, *38* (4), 385–392.

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.

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.

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.

— , **Falk Armin, Bart Golsteyn, David Huffman, and Uwe Sunde**, "Risk Attitudes across the Life Course," *Economic Journal*, forthcoming.

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.

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

Fernández, Raquel 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.

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

— , — , **and Assaf Sarid**, "Geographical Origins and Economic Consequences of Language Structures," *Working Paper*, 2017.

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

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.

Grosjean, Pauline, "A History of Violence: The Culture of Honor and Homicide in the US South," *Journal of the European Economic Association*, 2014, *12* (5), 1285–1316.

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

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: Papers & Proceedings*, 2001, *91* (2), 73–78.

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

Hofstede, Geert, *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations*, 2 ed., London: Sage, 2001.

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.

Litina, Anastasia, "Natural Land Productivity, Cooperation and Comparative Development," *Journal of Economic Growth*, 2016, *21* (4), 351–408.

Marsh, Herbert W., "Causal Ordering of Academic Self-Concept," *Journal of Educational Psychology*, 1990, *82* (4), 646–656.

_ , Ulrich Trautwein, Oliver Lüdtke, Olaf Köller, and Jürgen Baumert, "Academic Self-Concept, Interest, Grades, and Standardized Test Scores: Reciprocal Effects Models of Causal Ordering," *Child Development*, 2005, *76* (2), 397–416.

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.

Niederle, Muriel, "Gender," 2014. Mimeo, Stanford University Economics Department.

Nikiforakis, Nikos, "Punishment and Counter-Punishment in Public Good Games: Can We Really Govern Ourselves?," *Journal of Public Economics*, 2008, *92* (1), 91–112.

Nisbett, Richard E. and Dov Cohen, *Culture of Honor: The Psychology of Violence in the South*, Westview Press, 1996.

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.

Nunn, Nathan and Diego Puga, "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics*, 2012, *94* (1), 20–36.

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

Porta, Rafael La, Florencio Lopez-De-Silane, Andrei Shleifer, and Robert W. Vishny, "Trust in Large Organizations," *American Economic Review*, 1997, *87* (2), 333–38.

Putterman, Louis and David N. Weil, "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality," *Quarterly Journal of Economics*, 2010, *125* (4), 1627–1682.

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*, 2014, *61* (3), 637–648.

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, Lester M., "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.

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.

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.

Tabellini, Guido, "Institutions and Culture," *Journal of the European Economic Association*, 2008, *6* (2-3), 255–294.

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.

Weber, Max, *The Protestant Ethic and the "Spirit" of Capitalism*, New York: Charles Scribner's Sons., 1930.

ONLINE APPENDIX

A Construction and Content 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 four steps. First, an experimental validation procedure was conducted to select the survey items. Second, the survey items were translated and quantitative amounts were adjusted to ensure comparability across countries. Third, we implemented a pre-test of the selected survey items in a variety of countries to ensure implementability in a culturally diverse sample. Fourth, 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 Sampling and Survey Implementation

A.4.1 Background

The collection of our preference data was embedded into the regular World Poll 2012.[28] The international polling company Gallup has conducted an annual World Poll since 2005, in which it surveys representative population samples in almost every country – partly on a rotating basis – around the world on, e.g., economic, social, political, and environmental issues. The GPS was conducted in a subset of countries that were surveyed by Gallup in 2012.

A.4.2 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

[28] See http://www.gallup.com/services/170945/world-poll.aspx.

Sub-Saharan Africa: Botswana, Cameroon, Ghana, Kenya, Malawi, Nigeria, Rwanda, South Africa, Tanzania, Uganda, Zimbabwe

A.4.3 Sampling within Countries

In general, samples are probability based and nationally representative of the resident population aged 15 and older. The coverage area is the entire country including rural areas, and the sampling frame represents the entire civilian, non-institutionalized adult population of the country. Exceptions are noted in Table 11 and include areas where the safety of the interviewing staff is threatened and scarcely populated areas.[29]

Selecting Households and Respondents
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, which are stratified by population size and/or geography. 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. Next, households are selected using a random route procedure. Unless an outright refusal 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 initial sampled household, he or she uses a simple substitution method.

In countries where telephone interviewing is employed, Gallup uses a random digit dialing method or a nationally representative list of phone numbers. In select countries where cellphone penetration is high, Gallup uses a dual sampling frame. In face-to-face and telephone methodologies, random respondent selection within household is achieved by using either the latest birthday or Kish grid method.[30] Gallup makes at least three attempts to reach a person in each household.

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

[29]This paragraph is taken from www.gallup.com/178667/gallup-world-poll-work.aspx

[30]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.

samples and that the correct person is randomly selected in each household.

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.

Overview: Countries, Respondents and Interview Mode

Table 11: GPS-countries: Sample size, interview mode, interview language, and sample exclusions

Country	# Obs.	Interview Mode	Interview Language	Exclusions (Samples are nationally representative unless noted otherwise)
Afghanistan	1,000	Face-to-Face	Dari, Pashto	Gender-matched sampling was used during the final stage of selection.
Algeria	1,022	Face-to-Face	Arabic	Sparsely populated areas in the far South were excluded, representing appr. 10% of the population.
Argentina	1,000	Face-to-Face	Spanish	
Australia	1,002	Landline and Cellular Telephone	English	
Austria	1,001	Landline and Cellular Telephone	German	
Bangladesh	999	Face-to-Face	Bengali	Three hill districts in Chittagong (Rangmati, Khagrachori, and Bandarban) were excluded for security reasons, representing appr. 1% of the population.
Bolivia	998	Face-to-Face	Bolivia	
Brazil	1,003	Face-to-Face	Portuguese	
Cambodia	1,000	Face-to-Face	Khmer	
Cameroon	1,000	Face-to-Face	English, French, Fulfulde	The sample has a larger-than-expected proportion of respondents who report completing secondary education when compared with the data used for post-stratification weighting.
Canada	1,001	Landline and Cellular Telephone	English, French	Yukon, Northwest Territories, and Nunavut

				were excluded from the sample.
Chile	1,003	Face-to-Face	Spanish	
China	2,574	Face-to-Face, Landline Telephone	Chinese	Xinjiang and Tibet were excluded from the sample representing less than 2% of the population.
Colombia	1,000	Face-to-Face	Spanish	
Costa Rica	1,000	Face-to-Face	Spanish	
Croatia	992	Face-to-Face	Croatian	
Czech Republic	1,005	Face-to-Face	Czech	
Egypt	1,020	Face-to-Face	Arabic	
Estonia	1,004	Face-to-Face	Estonian, Russian	
Finland	1,000	Landline and Cellular Telephone	Finnish	
France	1,001	Landline and Cellular Telephone	French	
Georgia	1,000	Face-to-Face	Georgian, Russian	South Ossetia and Abkhazia were not included for the safety of interviewers, representing approx. 7% of the population.
Germany	997	Landline and Cellular Telephone	German	
Ghana	1,000	Face-to-Face	English, Ewe, Twi, Dagbani	
Greece	1,000	Face-to-Face	Greek	
Guatemala	1,000	Face-to-Face	Spanish	
Haiti	504	Face-to-Face	Creole	
Hungary	1,004	Face-to-Face	Hungarian	
India	2,539	Face-to-Face	Hindi, Tamil, Kannada, Telugu, Marathi, Gujarati, Bengali, Malayalam, Odia, Punjabi, Assamese	Excluded population living in Northeast states and on remote islands, representing less than 10% of the population.

Indonesia	1,000	Face-to-Face	Bahasa Indonesia	
Iran	2,507	Landline and Cellular Telephone	Farsi	
Iraq	1,000	Face-to-Face	Arabic, Kurdish	
Israel	999	Face-to-Face	Hebrew, Arabic	The sample does not include the area of East Jerusalem.
Italy	1,004	Landline and Cellular Telephone	Italian	
Japan	1,000	Landline Telephone	Japanese	Excluded 12 municipalities near the nuclear power plant Fukushima, representing less than 1% of the population of Japan.
Jordan	1,000	Face-to-Face	Arabic	Excluded population living in Madaba, Mafraq, Ajloun, Ma'an, Tafiliah, and Aqaba governorates, representing approx. 14% of the population.
Kazakhstan	999	Face-to-Face	Kazakh, Russian	
Kenya	1,000	Face-to-Face	English, Swahili	
Lithuania	999	Face-to-Face	Lithuanian	
Malawi	1,000	Face-to-Face	Chichewa, English, Tumbuka	
Mexico	1,000	Face-to-Face	Spanish	
Moldova	1,000	Face-to-Face	Romanian, Russian	Transnistria (Prednestrovie) was excluded for safety of interviewers, representing approx. 13% of the population.
Morocco	1,000	Face-to-Face	Moroccan Arabic, French, Berber	Excludes the Southern provinces, representing approx. 3% of the population.
Netherlands	1,000	Landline and Cellular Telephone	Dutch	
Nicaragua	1,000	Face-to-Face	Spanish	
Nigeria	1,000	Face-to-Face	English, Yoruba, Hausa,	

			Igbo, Pidgin English	
Pakistan	1,004	Face-to-Face	Urdu	Did not include Azad and Jammu Kashmir (AJK), representing approx. 5% of the population. Gender-matched sampling was used during the final stage of selection.
Peru	1,000	Face-to-Face	Spanish	
Philippines	1,000	Face-to-Face	Filipino, Iluko, Hiligaynon, Cebuano, Bicol, Waray, Maguindanaon	
Poland	999	Face-to-Face	Polish	
Portugal	998	Landline and Cellular Telephone	Portuguese	
Romania	994	Face-to-Face	Romanian	
Russia	1,498	Face-to-Face	Russian	North Ossetia, Kabardino-Balkariya, and remote small settlements in far-Eastern Siberia were excluded, representing approx. 5% or less of the population.
Rwanda	1,000	Face-to-Face	Kinyarwanda, French, English	
Saudi Arabia	1,035	Face-to-Face	Arabic	Includes Saudis and Arab expatriates; non-Arabs were excluded (representing approx. 20% of the adult population). Gender-matched sampling was used during the final stage of selection.
Serbia	1,023	Face-to-Face	Serbian	
South Africa	1,000	Face-to-Face	Afrikaans, English, Sotho, Zulu, Xhosa	

South Korea	1,000	Landline and Cellular Telephone	Korean	
Spain	1,000	Landline and Cellular Telephone	Spanish	
Sri Lanka	1,000	Face-to-Face	Sinhala, Tamil	
Suriname	504	Face-to-Face	Dutch	
Sweden	1,000	Landline and Cellular Telephone	Swedish	
Switzerland	1,000	Landline and Cellular Telephone	German, French, Italian	
Tanzania	1,000	Face-to-Face	Swahili, English	The Tanga region was excluded, representing approx. 5% of the population.
Thailand	1,000	Face-to-Face	Thai	
Turkey	1,000	Face-to-Face	Turkish	
Uganda	1,000	Face-to-Face	English, Luganda, Ateso, Runyankole	
Ukraine	1,000	Face-to-Face	Russian, Ukrainian	
United Arab Emirates	1,000	Face-to-Face	Arabic	Includes only Emiratis and Arab expatriates; non-Arabs were excluded (representing more than half of the adult population).
United Kingdom	1,030	Landline and Cellular Telephone	English	
United States	1,072	Landline and Cellular Telephone	English, Spanish	
Venezuela	999	Face-to-Face	Spanish	
Vietnam	1,000	Face-to-Face	Vietnamese	
Zimbabwe	1,000	Face-to-Face	English, Ndebele, Shona	

A.5 Survey Items of the GPS

A.5.1 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.5.2 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 6.

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 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 S7.

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.*

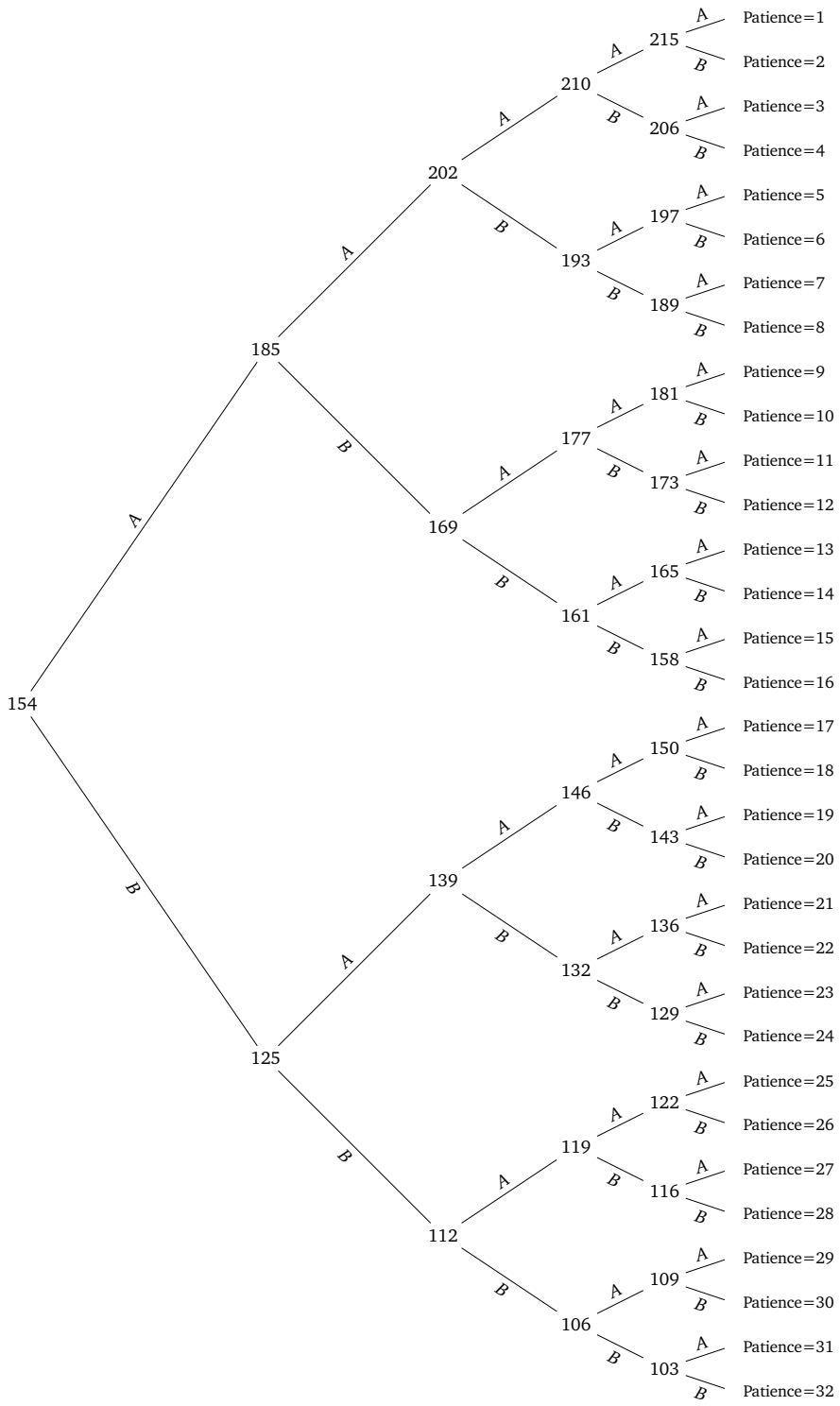

Figure 6: Tree for the staircase time task (numbers = payment in 12 months, A = choice of "100 euros today", B = choice of "x euros in 12 months". The staircase procedure worked as follows. First, each respondent was asked whether they would prefer to receive 100 euros today or 154 euros in 12 months 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.

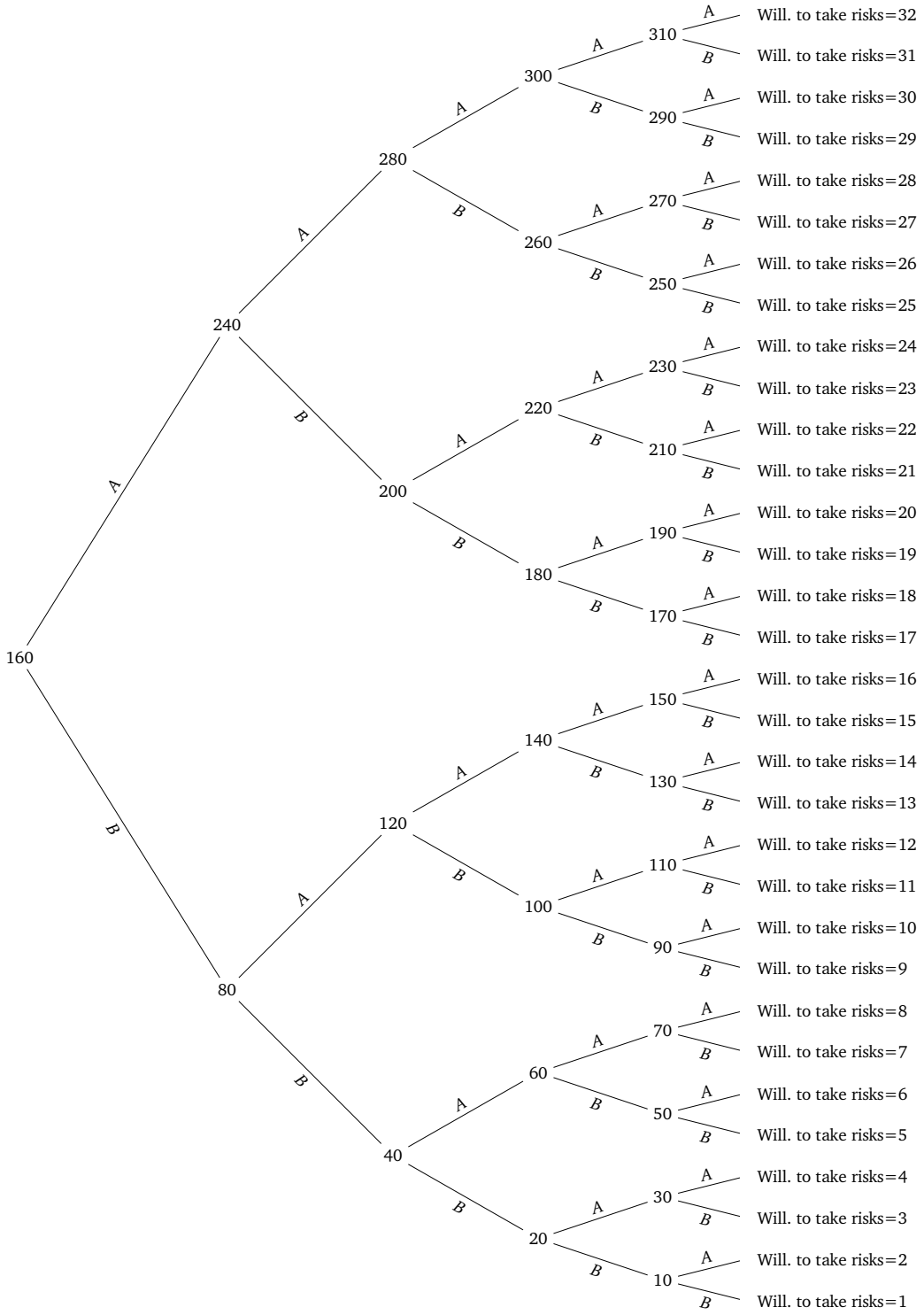

Figure 7: Tree for the staircase risk task (numbers = sure payment, A = choice of lottery, B = choice of sure payment). The staircase procedure worked as follows. First, each respondent was asked whether they would prefer to receive 160 euros for sure or whether they preferred a 50:50 chance of receiving 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.6.4 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?*

A.6.6 Trust

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

A.7 Correction for Implementation Errors

The GPS survey items were implemented with errors in a few countries. While these errors are minor, in this section we describe them in detail and explain how we recode the raw data to take implementation errors into account. To illustrate the majority of implementation errors and how we corrected them, consider Figures 8 and 9, which are the abstract versions of Figures 7 and 6, respectively. A typical error is that the payouts at a given node were not implemented correctly. In these cases, we still have unconfounded information about the preferences of respondents, i.e., behavior up to the erroneous node. For example, suppose that an error exists at node 7 in Figure 8. We then know that the willingness to take risks variable must assume a value between 1 and 4. We impute the midpoint of this interval, 2.5, for such respondents.

A.7.1 Staircase Risk

1. Indonesia; interview language Bahasa: At node 12, respondents should have faced a safe payment of IDR 36,000 (consult the uploaded questionnaire to verify this), but actually faced a safe payment of IDR 26,000. We hence code all 20 respondents who arrived at a willingness to take risk of 9 or 10 as 9.5 (the midpoint of the interval).

2. Pakistan, interview language Urdu: At node 26, respondents should have faced a safe payment of PKR 1040, but actually faced a safe payment of PKR 1140. We hence code all 101 respondents who arrived at willingness to take risk of 25-28 as 26.5.

3. Ukraine, interview language Ukrainian: At node 16, respondents should have faced a safe payment of UAH 130, but actually faced a safe payment of UAH 140. We hence code all 12 respondents who arrived at a willingnes to take risks of 13-14 as 13.5.

4. Vietnam, interview language Vietnamese: At node 6, all respondents should have faced a safe payment of VND 100,000, but actually faced a safe payment of VND 140,000. We hence code all 32 respondents who arrived at a willingness to take risks of 5-6 as 5.5. In addition, at node 31, all respondents should have faced a safe payment of VND 620,000, but actually faced a safe payment of VND 580,000. We hence code all 118 respondents who arrived at a willingness to take risks of 31-32 as 31.5.

5. Malawi, interview language Chichewa: At node 5, respondents should have faced a safe payment of MWK 175, but actually faced a safe payment of MWK 150. We hence code all 21 respondents who arrived at a willingness to take risks of 7-8 as 7.5.

6. Iran, interview language Farsi: All stakes were multiplied by a factor of 10. We cannot correct for this in the coding procedure. (2,507 respondents)

7. Uganda, interview language Ruanyankole: In all questions, the risky payoff was multiplied by a factor of 10. We cannot correct for this in the coding procedure. (132 respondents)

A.7.2 Staircase Patience

1. Vietnam; interview language Vietnamese: At node 14, respondents should have faced a future payment of VND 234,000, but actually faced a safe payment of VND 217,000. We hence code all 36 respondents who arrived at a patience of 17-20 as 18.5. In addition, at node 18, respondents should have faced a future payment of VND 323,000, but actually faced a future payment of VND 246,000. We hence code all 676 respondents who arrived at a patience of 1-8 as 4.5.

2. Iran, interview language Farsi: All stakes were multiplied by a factor of 10. We cannot correct for this in the coding procedure. (2,507 respondents)

A.7.3 Donation Variable

1. Iraq, interview language Kurdish: Respondents should have been asked how much of IQD 300,000 they would like to donate, but were actually asked how much of IQD 30,000 they would like to donate. Given that our “donation” variable is simply the fraction of

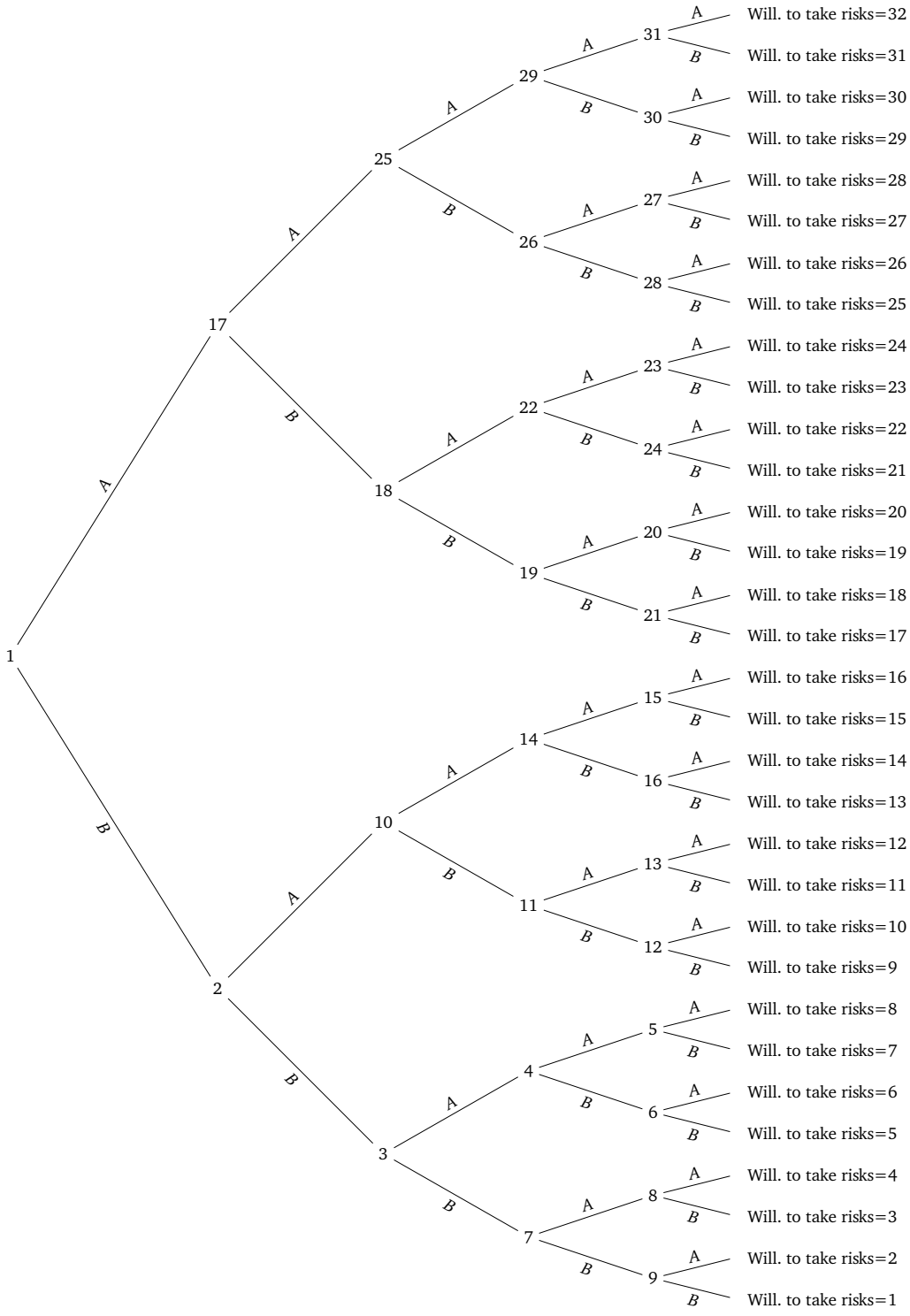

Figure 8: Tree for the staircase risk task (A = choice of lottery, B = choice of sure payment). Node labeling is for expository purposes only.

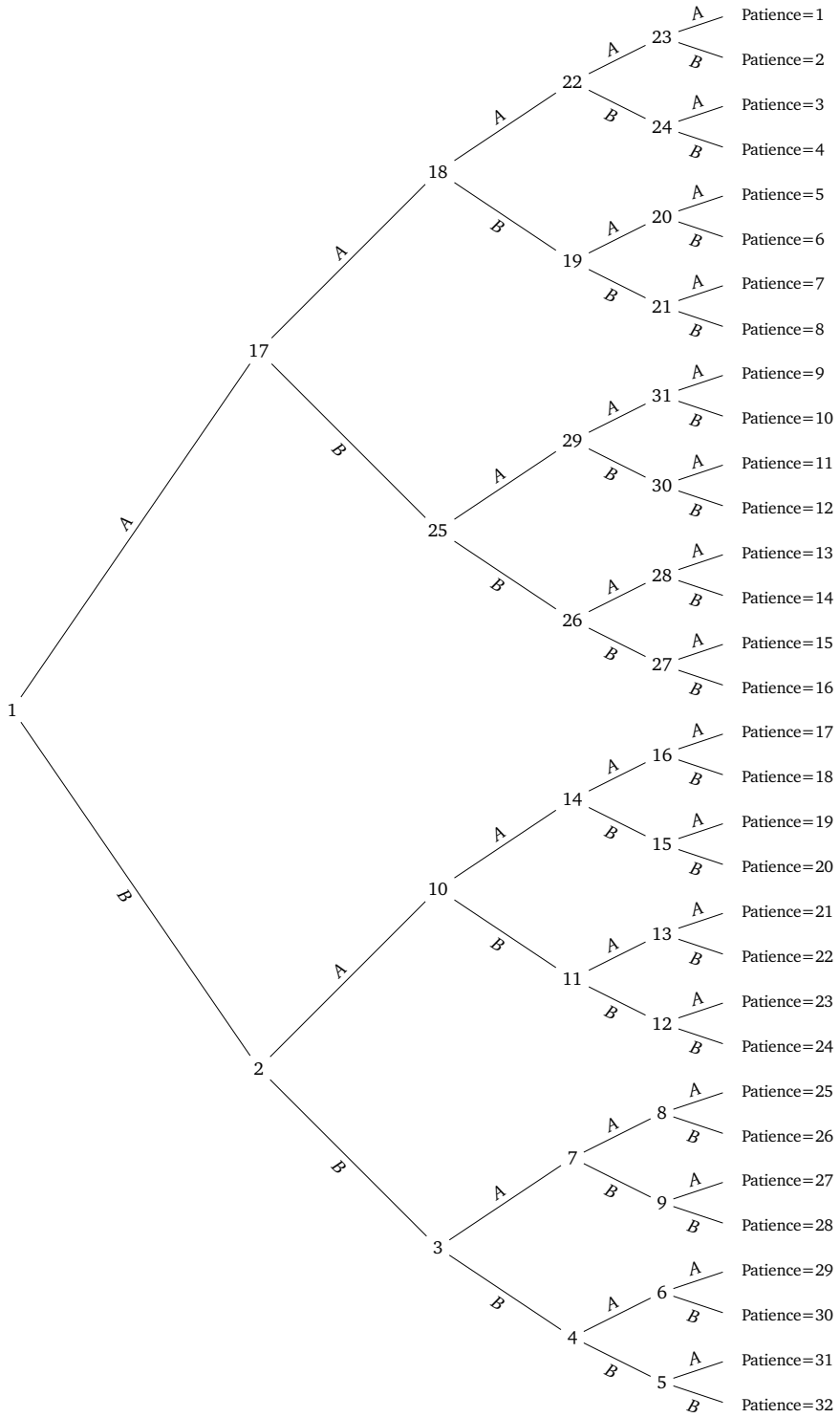

Figure 9: Tree for the staircase time task (A = choice of "100 euros today", B = choice of "x euros in 12 months". Node labeling is for expository purposes only.

the monetary endowment that respondents were willing to donate, we divide the actual donation amount of all Kurdish-speaking Iraqis by 30,000 rather than 300,000.

A.8 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.9 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:

$$
\begin{aligned}
\text{Patience} &= 0.7115185 \times \text{ Staircase patience } + 0.2884815 \times \text{ Will. to give up sth. today} \\
\text{Risk taking} &= 0.4729985 \times \text{ Staircase risk } + 0.5270015 \times \text{ Will. to take risks} \\
\text{Pos. reciprocity} &= 0.4847038 \times \text{ Will. to return favor } + 0.5152962 \times \text{ Size of gift} \\
\text{Neg. reciprocity} &= 0.5261938/2 \times \text{ Will. to punish if oneself treated unfairly} \\
&\quad + 0.5261938/2 \times \text{ Will. to punish if other treated unfairly} \\
&\quad + 0.3738062 \times \text{ Will. to take revenge} \\
\text{Altruism} &= 0.5350048 \times \text{ Will. to give to good causes } + 0.4649952 \times \text{ Hypoth. donation} \\
\text{Trust:} &\text{The survey included only one corresponding item.}
\end{aligned}
$$

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.

A.10 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.

B Additional Descriptive Results

B.1 Histograms by Preference

B.1.1 Individual Level

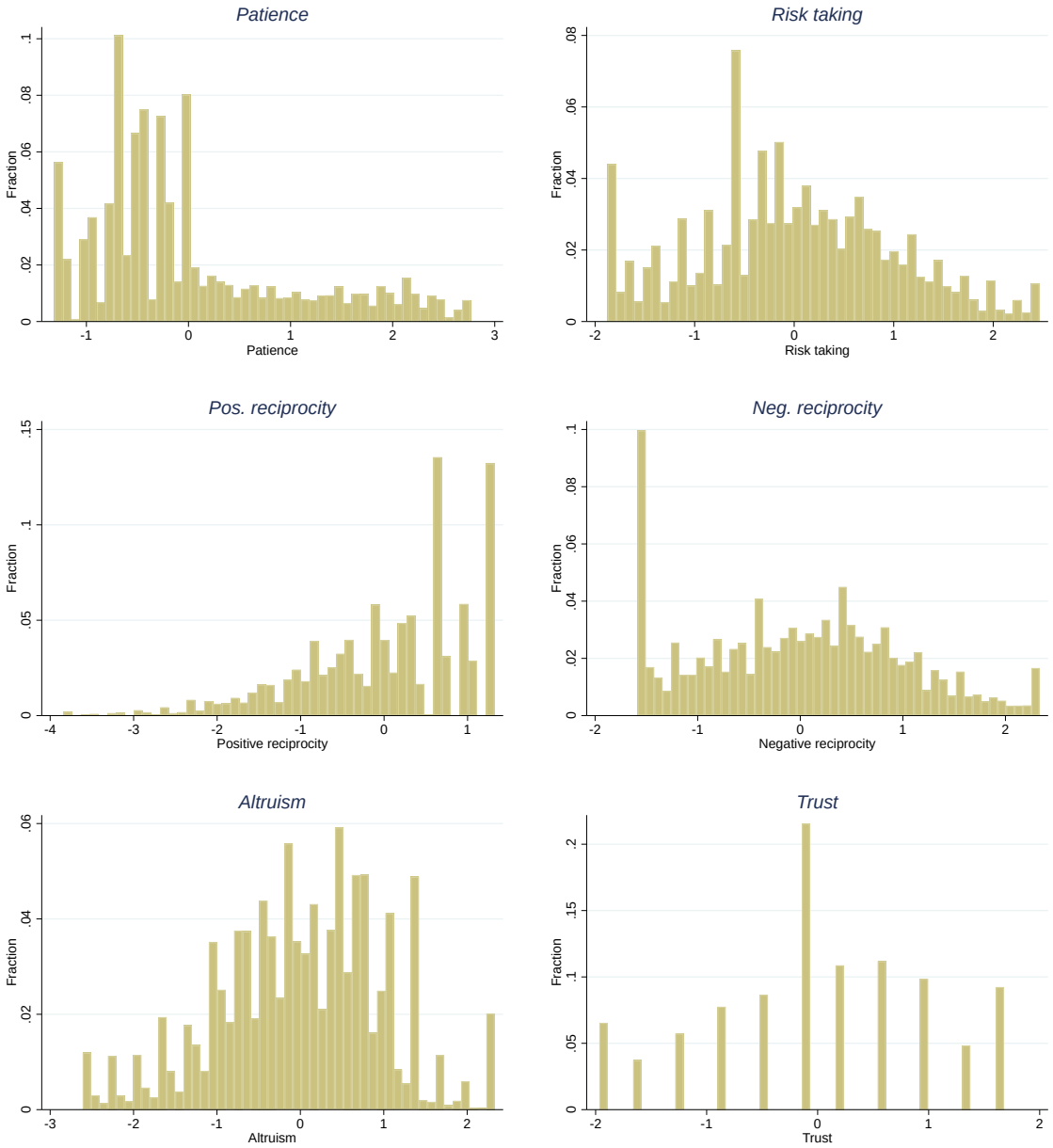

Figure 10: Distribution of preferences at individual level. The figure plots the distribution of standardized preference measures at the individual level. All data are standardized at the level of the individual in the full sample.

B.1.2 Country Level

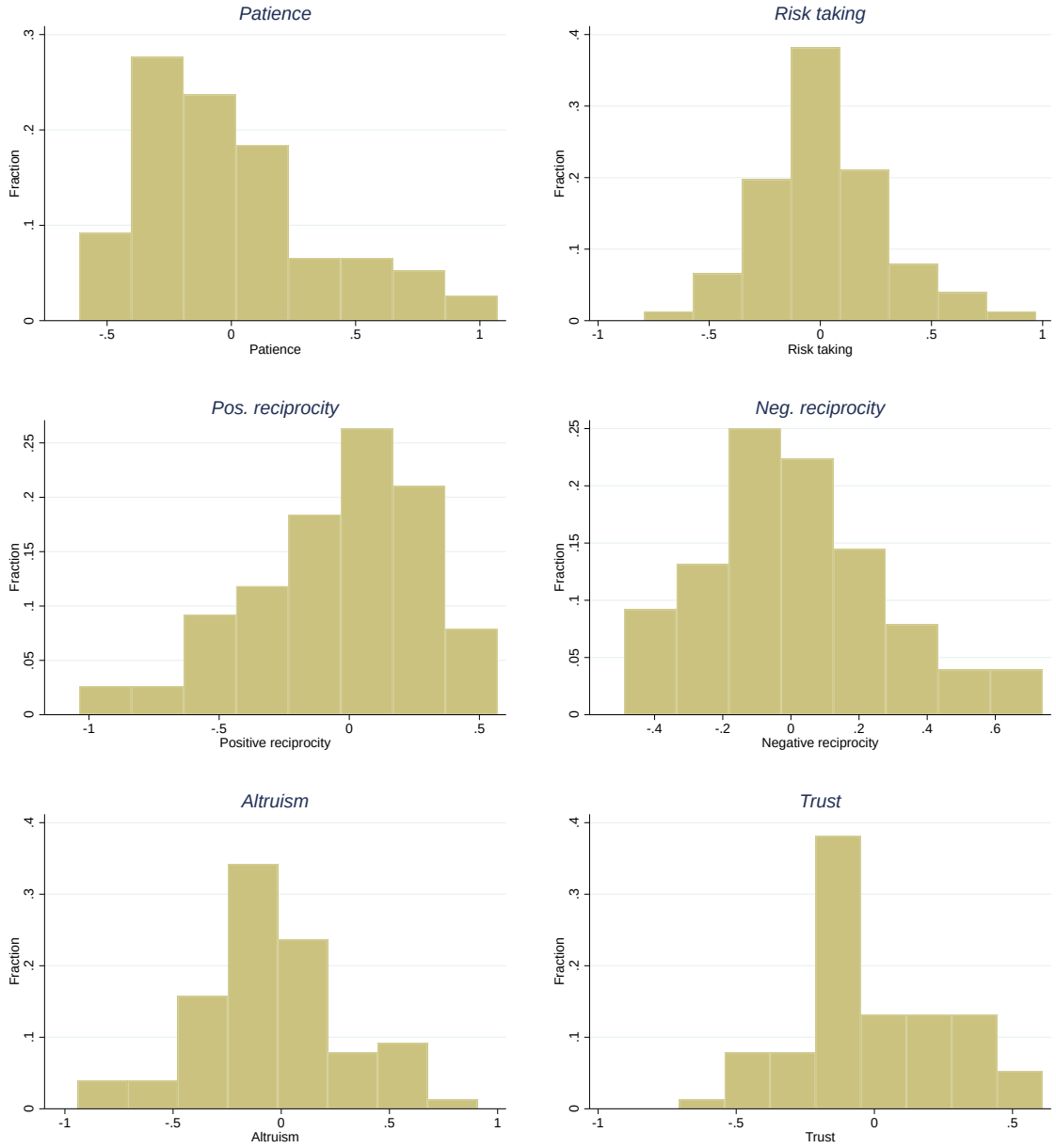

Figure 11: Distribution of preferences at country level. The figure plots the distribution of country averages of standardized preferences. All data are standardized at the level of the individual using the full sample.

C 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	Positive reciprocity	Negative reciprocity	Altruism	Trust
Patience	1					
Risk taking	0.210***	1				
Positive reciprocity	0.084***	0.068***	1			
Negative 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	Positive reciprocity	Negative reciprocity	Altruism	Trust
Patience						
Risk taking	71					
Positive reciprocity	40	30				
Negative 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.

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 error and thus removes this source of variation. This section provides evidence that measurement error is unlikely to be large enough to drive the result.

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

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

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

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

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

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

$$M = P + e,$$

such that we can rewrite

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

The total variance of the preference is hence

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

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

The regression model still allows identifying $Var(\delta)$, but the share of preference variation that is truly explained by the between-country variation is no longer given by the R^2, $Var(\delta)/Var(M)$, but rather by $Var(\delta)/Var(P)$. To assess whether between-country or within-country effects explain a larger share of total variation, one needs to compare $Var(\delta)/Var(P)$ to $Var(\epsilon)/Var(P)$. Since $Var(P) = Var(M) - Var(e)$, $Var(e)$ needs to be determined.

The variance of measurement error, $Var(e)$, is not directly observable, but estimates of test-

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

Take, as an example, the estimate for risk-taking in Table 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)
Age	0.72***	0.74***	-0.083	0.47**	1.02***	0.90***	-0.36*	-0.19	-0.0061	0.041	0.37*	-0.0022
	(0.17)	(0.18)	(0.20)	(0.20)	(0.17)	(0.17)	(0.19)	(0.19)	(0.14)	(0.15)	(0.21)	(0.15)
Age squared	-1.45***	-1.40***	-1.20***	-1.74***	-1.17***	-1.06***	-0.45**	-0.66***	0.015	-0.17	0.032	0.30*
	(0.20)	(0.21)	(0.21)	(0.20)	(0.18)	(0.17)	(0.18)	(0.19)	(0.15)	(0.16)	(0.20)	(0.16)
1 if female	-0.056***	-0.037***	-0.17***	-0.16***	0.049***	0.060***	-0.13***	-0.13***	0.10***	0.094***	0.066***	0.048***
	(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***	0.023***	0.046***	0.040***	0.038***	0.037***	0.040***	0.037***	0.044***	0.039***	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)
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	78501	57616	78445	57588	78869	57867	77521	56973	78632	57675	77814	57110
R^2	0.17	0.21	0.17	0.24	0.13	0.23	0.11	0.19	0.13	0.20	0.11	0.17

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). Additional controls include age, age squared, gender, subjective math skills, log household income, indicators for religious affiliation, a subjective institutional quality index, and a subjective health index. See Appendix J for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

F Language and Preference Variation within Countries

Table 15: Individual-level preferences and language

	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	0.20**	0.16***	0.084	0.063	0.55***	0.31***	0.23**	0.24***	0.43***	0.43***	0.41***	0.44***
	(0.09)	(0.04)	(0.05)	(0.07)	(0.12)	(0.07)	(0.10)	(0.06)	(0.12)	(0.12)	(0.08)	(0.07)
Pronoun drop not allowed	0.17*	0.12***	0.083**	0.11***	0.35***	0.21***	0.071	0.13***	0.24***	0.22***	0.016	-0.016
	(0.09)	(0.03)	(0.03)	(0.04)	(0.08)	(0.05)	(0.07)	(0.04)	(0.05)	(0.06)	(0.04)	(0.03)
Age	0.78***	0.85***	-0.18	-0.064	1.13***	1.20***	-0.52**	-0.54***	-0.082	-0.12	0.24	0.15
	(0.28)	(0.31)	(0.28)	(0.28)	(0.31)	(0.32)	(0.19)	(0.18)	(0.18)	(0.17)	(0.15)	(0.14)
Age squared	-1.56***	-1.65***	-1.13***	-1.21***	-1.27***	-1.38***	-0.36	-0.33	0.062	0.045	0.17	0.21
	(0.31)	(0.34)	(0.29)	(0.30)	(0.33)	(0.34)	(0.22)	(0.21)	(0.16)	(0.16)	(0.16)	(0.15)
1 if female	-0.067***	-0.064***	-0.19***	-0.19***	0.072***	0.071***	-0.13***	-0.14***	0.13***	0.13***	0.10***	0.085***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Subj. math skills	0.027***	0.027***	0.043***	0.040***	0.039***	0.038***	0.034***	0.028***	0.043***	0.039***	0.058***	0.053***
	(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]	0.056***	0.059***	0.069***	0.063***	0.029***	0.043***	0.017*	0.020*	0.034***	0.039***	-0.013**	-0.0049
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	-0.53***	-0.62***	-0.048	0.0034	-1.23***	-1.51***	0.19*	-0.41***	-0.48***	-0.48***	-0.52***	-1.03***
	(0.10)	(0.12)	(0.08)	(0.12)	(0.07)	(0.08)	(0.10)	(0.13)	(0.09)	(0.11)	(0.06)	(0.08)
Country FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Subnational region FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Religion FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	53054	46327	53055	46356	53299	46549	52309	45768	53119	46398	52548	45916
R^2	0.188	0.236	0.184	0.238	0.113	0.202	0.121	0.188	0.124	0.178	0.119	0.165

Notes. OLS estimates, standard errors (clustered at interview language level) in parentheses. The analyses exploit variation in interview language (and associated language structures) within countries or subnational regions. Coefficients are in terms of units of standard deviations of the respective preference (relative to the individual world mean). Age is divided by 100. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G Additional Results on Individual-Level Outcomes

G.1 Robustness Checks: All Preferences Simultaneously

Including all preferences simultaneously is not our preferred approach because it introduces problems of multicollinearity. Still, to check robustness, Tables 16 and 17 pesent the results of the individual-level outcomes regressions with all preferences as explanatory variables.

Table 16: 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**	-0.0037	-0.0070
	(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***	0.0041*	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	0.00057	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)
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 J). Self-employment and planned self-employment are binary, while smoking intensity is measured in three categories (never, occasionally, frequently). Additional controls include age, age squared, gender, subjective math skills, log household income, and indicators for religious affiliation. See Appendix J for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 17: 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***	-0.0015	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	0.000092	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	-0.00065	-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**	0.0075**	0.015***	0.011***	0.010***	0.0077**	0.0049*	0.0011	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***	0.00019
	(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)
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. Saved last year is a binary indicator, while education level is measured in three categories (roughly elementary, secondary, and tertiary education, see Appendix J). Self-employment and planned self-employment are binary, while smoking intensity is measured in three categories (never, occasionally, frequently). Additional controls include age, age squared, gender, subjective math skills, log household income, and indicators for religious affiliation. See Appendix J for additional information about the variables.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G.2 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 5.1 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 12 shows that the positive correlation between patience and savings holds in virtually all countries in our sample.

While Figure 12 reports the results for patience and risktaking, Figure 13 visualizes the relationships between altruism and behaviors. Finally, Figure 14 presents the correlations between positive and negative reciprocity and the behaviors discussed in Section 5.1 of the main text.

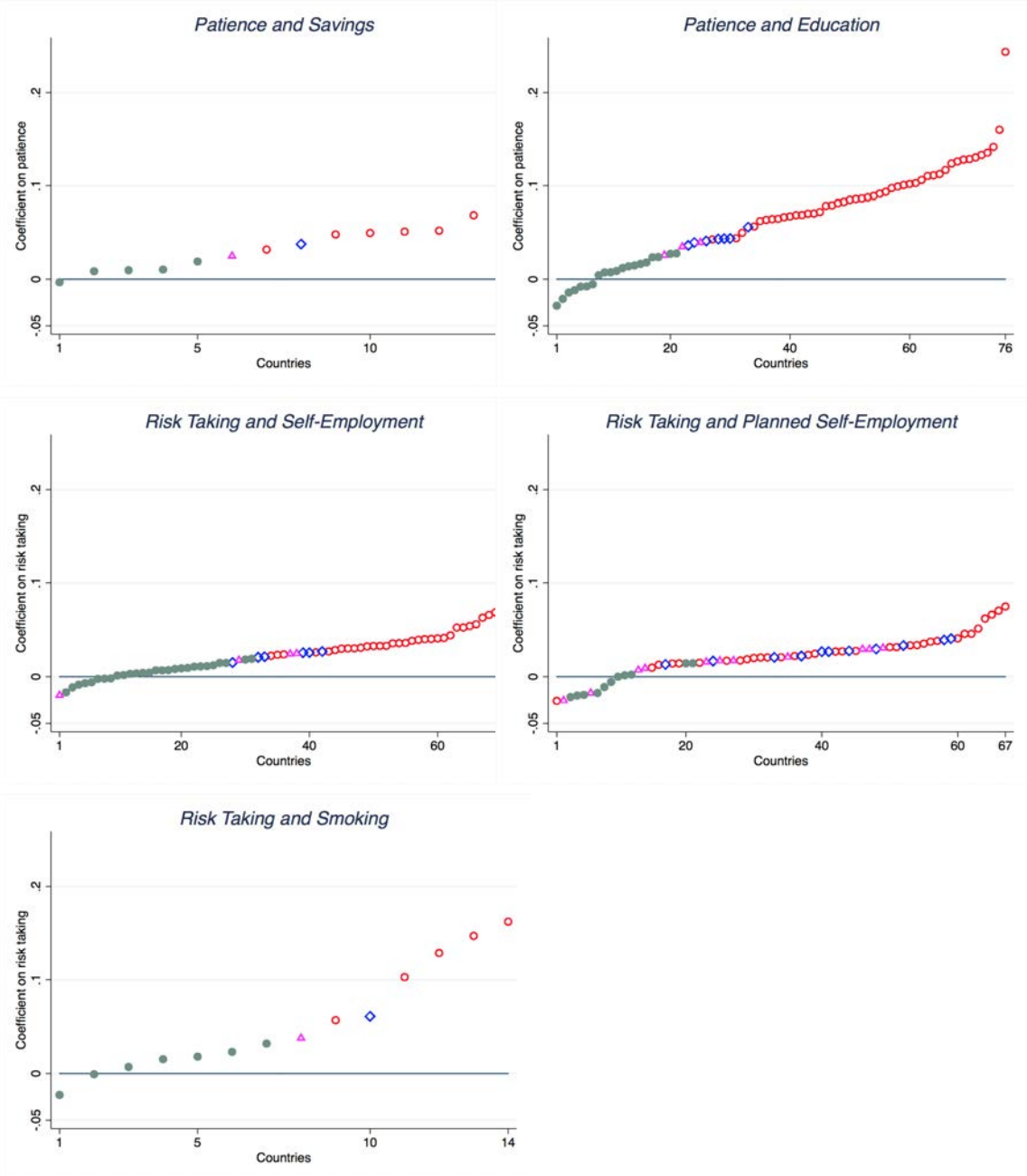

Figure 12: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable,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

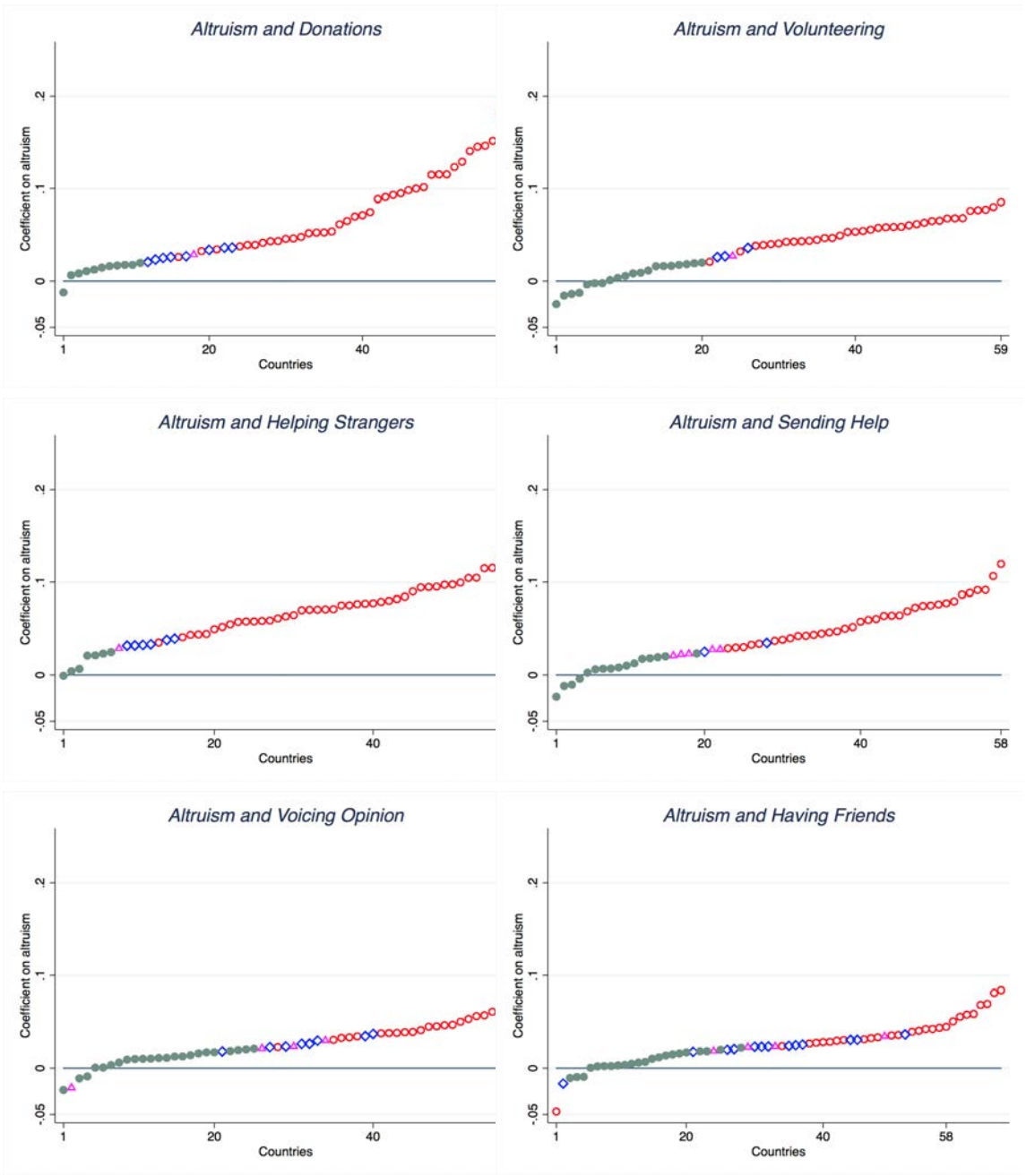

Figure 13: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable,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

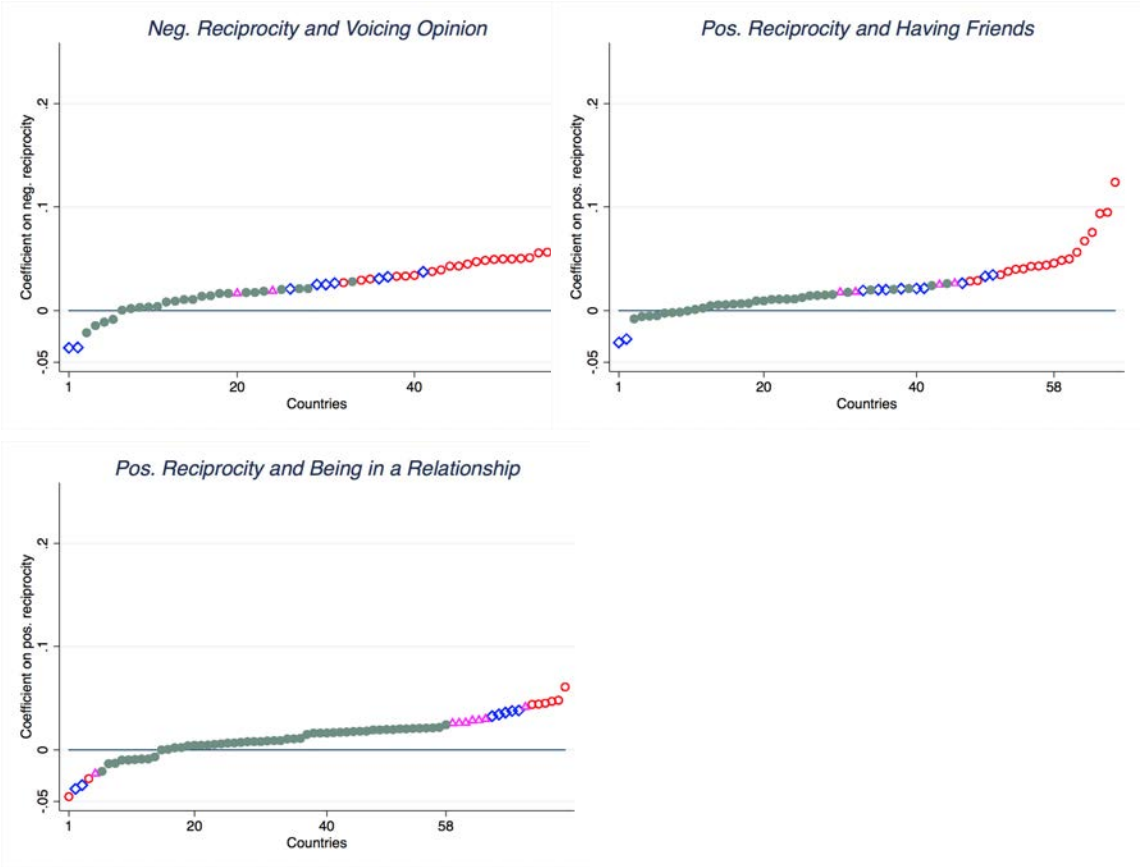

Figure 14: Correlations separately by country. Each panel plots the distribution of correlations across countries. That is, for each country, we regress the respective outcome on a preference and plot the resulting coefficients as well as their significance level. In order to make countries comparable,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

H Additional Results on Cross-Country Outcomes

Table 18: Country-level outcomes and preferences

	Dependent variable:									
	Entrepreneurship						Social outcomes			
	Patent applic. p/c		Scientific articles p/c		TFP		Volunt. & donat.		Armed conflicts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Risk taking	-0.11	-1.08	0.10**	-0.12***	0.20**	0.055	1.74**	0.86	-0.71	-0.92
	(0.54)	(0.71)	(0.05)	(0.04)	(0.10)	(0.11)	(0.77)	(0.55)	(0.49)	(0.57)
Prosociality	0.81*	0.68	0.051	0.0079	-0.00031	-0.034	1.26**	0.85*	0.078	0.039
	(0.44)	(0.41)	(0.05)	(0.04)	(0.07)	(0.07)	(0.53)	(0.48)	(0.39)	(0.41)
Negative reciprocity	1.35*	1.12	-0.035	-0.082	0.079	0.055	0.50	0.099	1.39***	1.33***
	(0.79)	(0.74)	(0.08)	(0.05)	(0.11)	(0.09)	(0.93)	(0.85)	(0.43)	(0.43)
Patience		1.57***		0.41***		0.25***		1.57**		0.38
		(0.50)		(0.05)		(0.07)		(0.70)		(0.43)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	61	61	67	67	59	59	32	32	73	73
R^2	0.70	0.74	0.45	0.78	0.49	0.57	0.56	0.65	0.35	0.35

OLS estimates, robust standard errors in parentheses. The dependent variables in columns (1)–(2) and (3)–(4) are the logs of the number of patent applications p/c and the number of scientific articles p/c, respectively. In columns (7)–(8), the dependent variable is volunteering and donation as a fraction of GDP. Frequency of conflicts is measured by the log of conflicts according to PRIO, in the Quality of Government dataset. Prosociality is the first principal component of altruism, positive reciprocity, and trust. Controls include distance to equator, average temperature, average precipitation, the share of the population living in (sub-)tropical zones, terrain ruggedness, average distance to nearest waterway, and an island dummy. See Online Appendix J for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I Results on Candidate Preference Proxies in the WVS

Table 19: 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)
Age	0.35	-0.11	-2.71***	-2.28***	-0.47	-0.15	0.71***	0.28
	(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)
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)
Constant	-0.21***	-0.36***	1.14***	1.19***	0.17**	0.22***	-0.26***	-0.21***
	(0.06)	(0.05)	(0.07)	(0.08)	(0.08)	(0.06)	(0.07)	(0.05)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	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 20: 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	-0.0033*	0.0016	-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		-0.42***		-0.46		1.42***
		(0.11)		(0.73)		(0.18)
Age squared		0.33***		-2.16***		-1.25***
		(0.09)		(0.58)		(0.15)
Female		-0.020***		-0.19***		-0.072***
		(0.00)		(0.03)		(0.01)
Constant	0.42***	0.100**	5.18***	3.83***	0.11***	-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	277317	220856	296130	230555	151767	131698
R^2	0.082	0.170	0.107	0.271	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 J for additional information about the variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J Description and Data Sources of Outcome Variables

J.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.

J.2 Country-Level Variables

Number of patent applications Number of patent applications per capita, according to the World Bank Development Indicators, averaged 2003–2012.

Scientific Articles The mean, over the period 1981-2000, of the annual number of scientic articles per capita, calculated as the total number of scientific and technical articles published in a given year divided by the total population in that year.

Total factor productivity TFP average 2003-2012, Penn World tables.

Conflicts The number of conflicts according to PRIO are taken from the Quality of Government dataset.

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.

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).

Terrain ruggedness. Taken from Nunn and Puga (2012).

Mean distance from nearest waterway. Distance from GIS grid cell to nearest icefree coastline or sea-navigable river, averaged across cells. 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).

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).

Geographic and biological conditions Taken from Spolaore and Wacziarg (2013), originally based on Olsson and Hibbs (2005).

Crop suitability of land Taken from Galor and Özak (2016).

Family ties Constructed from WVS following Alesina and Giuliano (2013).

Future time reference Classification adapted from Chen (2013) with minor additions and changes. First, we set Persian to missing after corresponding with him (he originally classified Persian as strong FTR, which is open to discussion). Second, we managed to classify Moroccan Arabic (strong), Fula (strong), and Khmer (weak).

Pronoun drop Classification based on World Atlas of Languages (WALS).