I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12059

Gender Differences in Preferences Increase with Economic Development and Gender Equality

Armin Falk
Johannes Hermle

DECEMBER 2018

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12059

Gender Differences in Preferences Increase with Economic Development and Gender Equality

Armin Falk
University of Bonn, briq and IZA

Johannes Hermle
University of California, Berkeley and IZA

DECEMBER 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Gender Differences in Preferences Increase with Economic Development and Gender Equality*

Preferences – concerning time, risk and social interactions – systematically shape human behavior, and contribute to differential economic and social outcomes between the genders. Here, we present a global investigation of gender differences in six fundamental preferences. Our data consist of 80,000 individuals in 76 representative country samples with measures on willingness to take risks, patience, altruism, positive and negative reciprocity as well as trust. Gender differences in preferences were positively related to economic development and gender equality. This suggests that greater availability of and equal access to material and social resources for both genders favor the manifestation of gender-differentiated preferences across countries.

JEL Classification: C91, D91, D63, D64

Keywords: gender, preferences, cross-country variation

Corresponding author:
Armin Falk
briq - Institute on Behavior & Inequality
Schaumburg-Lippe-Str. 5-9
53113 Bonn
Germany
E-mail: armin.falk@briq-institute.org

* This is the authors' version of the work. The definitive version was published in *Science*, 19 October 2018: Vol. 362, Issue 6412, eaas9899 [DOI: 10.1126/science.aas9899].

We thank P. Andre, E. Avis, S. Fiedler, K. Jensen, M. Kosfeld, L. Shiferaw, and J. Tungodden for helpful comments. L. Henkel, R. Suchy, and S. Walter provided outstanding research assistance. We also thank M. Antony for his support in collecting and managing the GPS data, as well as M. Entner and S. Sauter for administrative support.

A. Falk acknowledges financial support from the European Research Council through ERC #209214. The authors declare no competing interests.

The data on preferences are deposited here: https://gps.briq-institute.org

Fundamental preferences, such as altruism, risk-taking, reciprocity, patience or trust, constitute the foundation of choice theories and govern human behavior. A growing literature in economics (*1*, *2*) and psychology (*3*) documents important differences in preferences between the genders. These differences provide a key explanation for differential choices and outcomes between women and men in contexts such as occupational choice, financial investment, or educational decisions (*4*, *5*), among many others. In understanding the origins of gender differences in preferences and their variability across countries and cultures, an extensive literature discusses biological and evolutionary determinants (*6*, *7*) and the role of the social environment (*8–10*).

Hypotheses

We contrast two competing hypotheses which make opposite predictions concerning the cross-country correlational patterns of gender differences in preferences with economic development and gender equality. Following social role theory, one may hypothesize that gender differences in preferences attenuate in more developed, gender-egalitarian countries (social role hypothesis). This hypothesis rests on two premises. First, economic development is a key determinant of societal progression towards gender equality (*11*, *12*), which is critical for the dissolution of traditional gender roles (*13*, *14*). Second, as discussed by a large body of literature (*8–10*), gender-specific roles instill distinct preferences in women and men and hence constitute a crucial component in explaining the gender preference gap. As a consequence, according to the social role hypothesis, higher economic development and gender equality, and the associated dissolution of traditional gender roles should lead to a narrowing of gender differences in preferences.

In contrast to the social role hypothesis, there is reason to expect that gender differences in preferences expand with economic development and gender equality (resource hypothesis). As suggested by post-materialist theory (*15*, *16*), a critical societal precondition for self-expression is the fulfillment of basic material needs. In line with this, existing research documents that the

unrestricted expression of preferences hinges on the availability of sufficient material and social resources (*17–20*). With respect to gender, differences in preferences should therefore manifest themselves only if both genders obtain sufficient access to these resources to independently develop and express their intrinsic preferences (*21*). Specifically, greater availability of material resources eliminates the gender-neutral goal of subsistence. This creates scope for attending to gender-specific ambitions and desires. As a consequence, economic development may facilitate the unfolding of differences between the genders. More developed countries also feature higher levels of gender equality in political, social and economic domains (*11*), which is a critical requirement for the acceptance of gender-specific desires and preferences. In particular, as women become less exposed and vulnerable to male influence, gender differentiation may be reinforced through women's greater opportunities of self-expression. In sum, greater availability of material and social resources to both genders may facilitate the independent development and expression of gender-specific preferences, and hence lead to an expansion of gender differences in more developed and gender-egalitarian countries.

Data and Measures

An empirical test of the two competing hypotheses requires data that meet three critical conditions: (i) reliability of preference measures, (ii) extensive cultural variation as well as comprehensive global coverage, and (iii) representativeness of country samples. Our investigation used the Global Preference Survey (GPS) (*22*, *23*). The GPS was collected as part of the Gallup World Poll 2012 and contains measures of six fundamental preferences with regards to social and non-social domains: willingness to take risks; patience, which captures preferences over the inter-temporal timing of rewards; altruism; trust (*24*); as well as positive and negative reciprocity, which capture the costly willingness to reward kind actions, or to punish unkind actions, respectively.

Before the launch of the international survey, multiple survey items were selected for these preferences through an ex-ante experimental validation (*25*). For each preference, subjects

responded to a large set of survey items and participated in incentivized choice experiments. The subset of survey items that maximized adjusted R-squared in predicting incentivized behavior in the corresponding experiment was selected for the international survey. The selected items, which are described in the Materials and Methods, comprise a combination of qualitative self-assessments and quantitative items that involve economic trade-off decisions. The qualitative items elicit participants' subjective assessment of their willingness to act in a certain way, such as whether participants are generally willing to take risks. Complementarily, the quantitative items provide revealed preference measures by using participants' choices in monetary tradeoff decisions. As an example, the quantitative item for risk taking provides the participants with a sequence of five interdependent choices between a fixed and a risky payment (lottery). This allows one to progressively approach the point of indifference between the fixed payment and the lottery, which serves as a revealed preference measure for risk taking behavior. The presence of both qualitative and quantitative items allows for robustness tests with respect to potential culture-specific response behavior. To make survey items comparable across cultures, all items were translated back and forth by professionals and monetary values mentioned in the survey questions were adjusted along median household income across countries. To guarantee cross-cultural validity, the survey items were pre-tested in 22 countries of various cultural heritage as part of the Gallup World Poll 2012 pre-test conducted in late 2011.

After the ex-ante experimental validation and pre-tests, the international survey was implemented in a total of 76 countries, representing about 90 percent of the global population and global GDP. To provide geographic representativeness as well as developmental and cultural variation, the countries were selected to include all continents and a very broad range of economic development levels. To allow generalizable inferences, for each country the data contain samples representative of the resident population aged 15 and older, with a median sample size of 1,000 participants per country. In total, the data include preference measures for about 80,000 participants.

After implementation of the worldwide survey, the measures for the six preferences were generated according to the following procedure. First, each of the survey items was standardized using the mean and variance of the entire worldwide sample. Then, to obtain the preference measures, the relevant z-scores were averaged using weights developed in the experimental validation. For further details on the data collection and construction of our measures see the Materials and Methods as well as Supplementary Online Material.

The data allow one to assess the existence and quantitative relevance of gender differences in preferences at the global level (*22*). For this purpose, global gender differences were calculated as follows: each preference measure was standardized at the global level to exhibit a mean of zero and a standard deviation of one. Then, for each preference, an OLS regression was performed on the worldwide sample using as independent variable a gender indicator in which male is the reference category, controlling for age, age squared, subjective math skills, education level, household income quintile, and country fixed effects. Standard errors were clustered at the country level. The estimated coefficient on the gender indicator served as the gender difference in the respective preference. On the global level, all six preferences featured significant gender differences (fig. S1): women tended to be more prosocial and less negatively reciprocal than men with differences in standard deviations of 0.106 for altruism ($p<0.0001$), 0.064 for trust ($p<0.0001$), 0.055 for positive ($p<0.0001$) and 0.129 for negative reciprocity ($p<0.0001$), respectively. Turning to non-social preferences, women were less risk taking by 0.168 standard deviations ($p<0.0001$), and less patient by 0.050 standard deviations ($p<0.0001$) (*26*). The observed differences in preferences set the stage for our analysis.

Analysis of Gender Differences in Preferences in Relation to Economic Development and Gender Equality

To test the competing hypotheses, we computed country-level gender differences for each preference. For this purpose, we standardized each preference measure at the country level to exhibit a mean of zero and a standard deviation of one. We then performed for each preference and country a separate OLS regression using as independent variable a gender indicator in which

male is the reference category. We also included several controls to isolate the gender effect from potentially confounding factors which differ between the genders. These controls are age, age squared, subjective math skills, education level, and household income quintile. The obtained coefficient on the gender indicator served as measure of the gender difference in the respective preference and country.

Using the country-level estimates of gender differences in preferences, we examined variation along levels of economic development and gender equality. As the measure of economic development, we used GDP per capita. To assess the role of gender equality, we created a Gender Equality Index as a joint measure of four indices of gender equality: (A) the Global Gender Gap Index of the World Economic Forum (WEF), (B) the Gender Equality Index of the United Nations (UN), (C) the ratio of female and male labor force participation rates, and (D) years since women's suffrage. The Gender Equality Index was constructed as the predicted main component from a principal component analysis of the four indices.

To study the effect of economic development, we first sorted the 76 countries into four bins according to their level of development, measured by GDP per capita. We then computed for each bin the average country-level gender difference in each preference. Gender differences in all six preferences increased with a country's level of development (Fig. 1A). The positive correlations between log GDP per capita and country-level gender differences were large and statistically significant for all six preferences (0.58 for altruism ($p<0.0001$), 0.59 for trust ($p<0.0001$), 0.31 for positive reciprocity ($p=0.0067$), 0.35 for negative reciprocity ($p=0.0017$), 0.37 for risk taking ($p=0.0011$), and 0.38 for patience ($p=0.0006$)) (fig. S2). We also analyzed a summary index of gender differences for all preferences jointly. For this purpose, we first performed a principal component analysis of the country-level gender differences in the six preferences. We then created an index of gender differences in preferences as the predicted first main component. This index exhibited a correlation of 0.67 ($p<0.0001$) with log GDP per capita (Fig. 1B) (*27*).

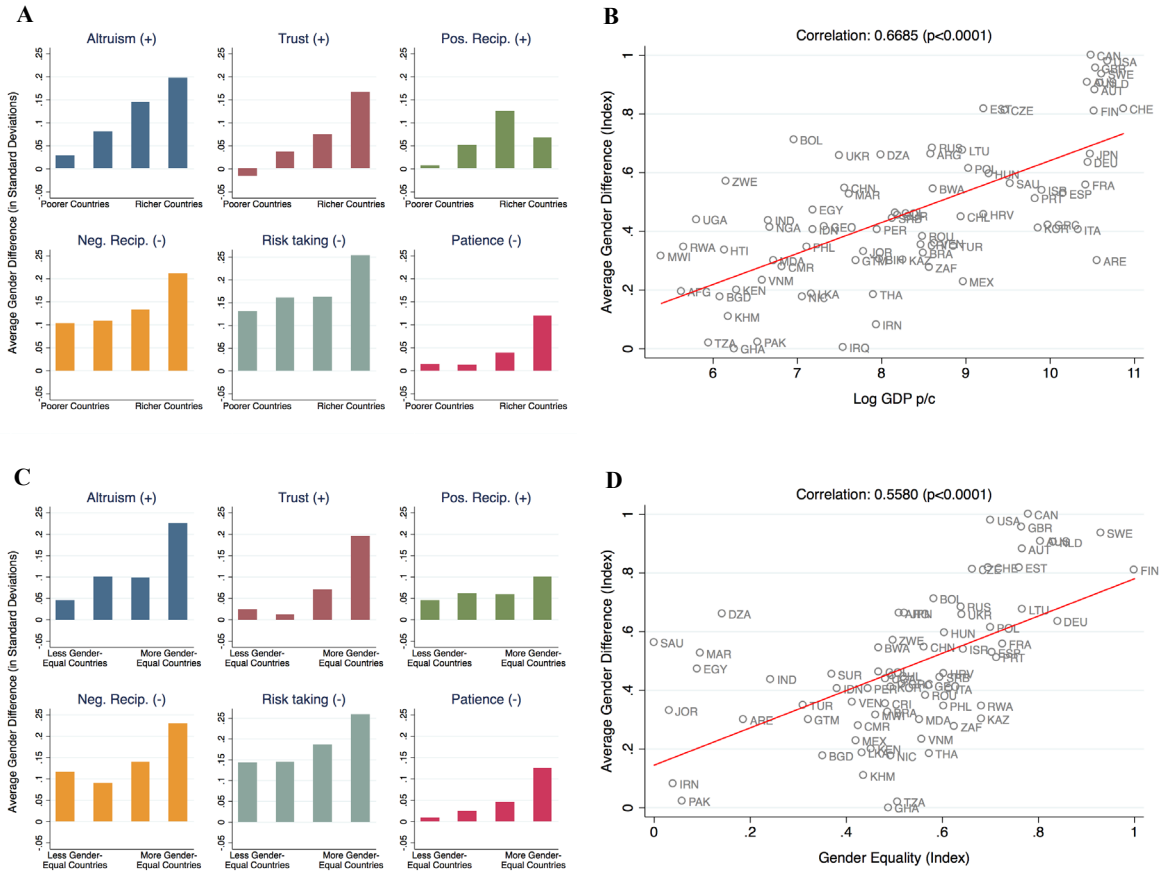


Fig 1. Analysis of gender differences in relation to economic development and gender equality. (A) Mean country-level gender difference in altruism, trust, positive reciprocity, negative reciprocity, risk taking, and patience by development level. Countries were sorted into 4 bins according to their GDP per capita quartile. The symbols (+)/(−) in the panel titles indicate the sign of the difference for each preference. (+) indicates that positive differences correspond to women exhibiting higher levels of the respective preference. (−) indicates that positive differences correspond to women exhibiting lower levels of the respective preference. (B) The relationship between the aggregate index of gender differences in all six preferences and log GDP per capita. (C and D) Same relationships for the Gender Equality Index. Country abbreviations are spelled out in the Supplementary Online Material.

To study the effect of gender equality, we ran the same analysis as for economic development using the Gender Equality Index as the explanatory variable. Gender differences in preferences were found to increase with gender equality both for each preference separately (Fig. 1C) as well as for the index of gender differences in preferences (Fig. 1D). For the individual preferences the correlation coefficients were 0.51 for altruism ($p<0.0001$), 0.41 for

trust (p=0.0005), 0.13 for positive reciprocity (p=0.2875), 0.40 for negative reciprocity (p=0.0005), 0.34 for risk taking (p=0.0036), and 0.43 for patience (p=0.0002) (fig. S3). The summary index of gender differences in preferences exhibited a correlation of 0.56 (p<0.0001) with the Gender Equality Index. Reassuringly, the positive relationship between the index of gender differences in preferences and gender equality was also found for the four individual indicators of gender equality (fig. S4).

Economic development and gender equality are strongly intertwined (*11*). To isolate the separate impacts of economic development and gender equality on gender differences in preferences, we therefore conducted a conditional analysis. We constructed partial regression plots illustrating the relationship between the index of gender differences in preferences and log GDP per capita conditional on the Gender Equality Index (Fig. 2A) and vice versa (Fig. 2B). The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. Hence, the slope coefficients can be interpreted as the standard deviation change in the dependent variable in response to a one standard deviation change in the independent variable.

There was a quantitatively large and statistically significant association of gender differences with log GDP per capita conditional on the Gender Equality Index. The estimated slope coefficient was 0.53 (p<0.0001). Likewise, gender differences were strongly associated with the Gender Equality Index conditional on log GDP per capita with a somewhat smaller slope coefficient of 0.32 (p=0.0033) (see also column 7 in table S4). When conducting an F-test for equality of both coefficients, we failed to reject at p=0.2537, indicating that the strength of the relationships between the index of gender differences in preferences and log GDP per capita and the Gender Equality Index were not statistically different. These findings imply that both economic development and gender equality exhibited an independent and significant association with gender differences in preferences (*28*). Conditional on log GDP per capita, differences in preferences were also significantly and positively associated with the four individual measures of gender equality (Figs. 2C to F). Slope coefficients were 0.23 (p=0.0084) for the WEF Global

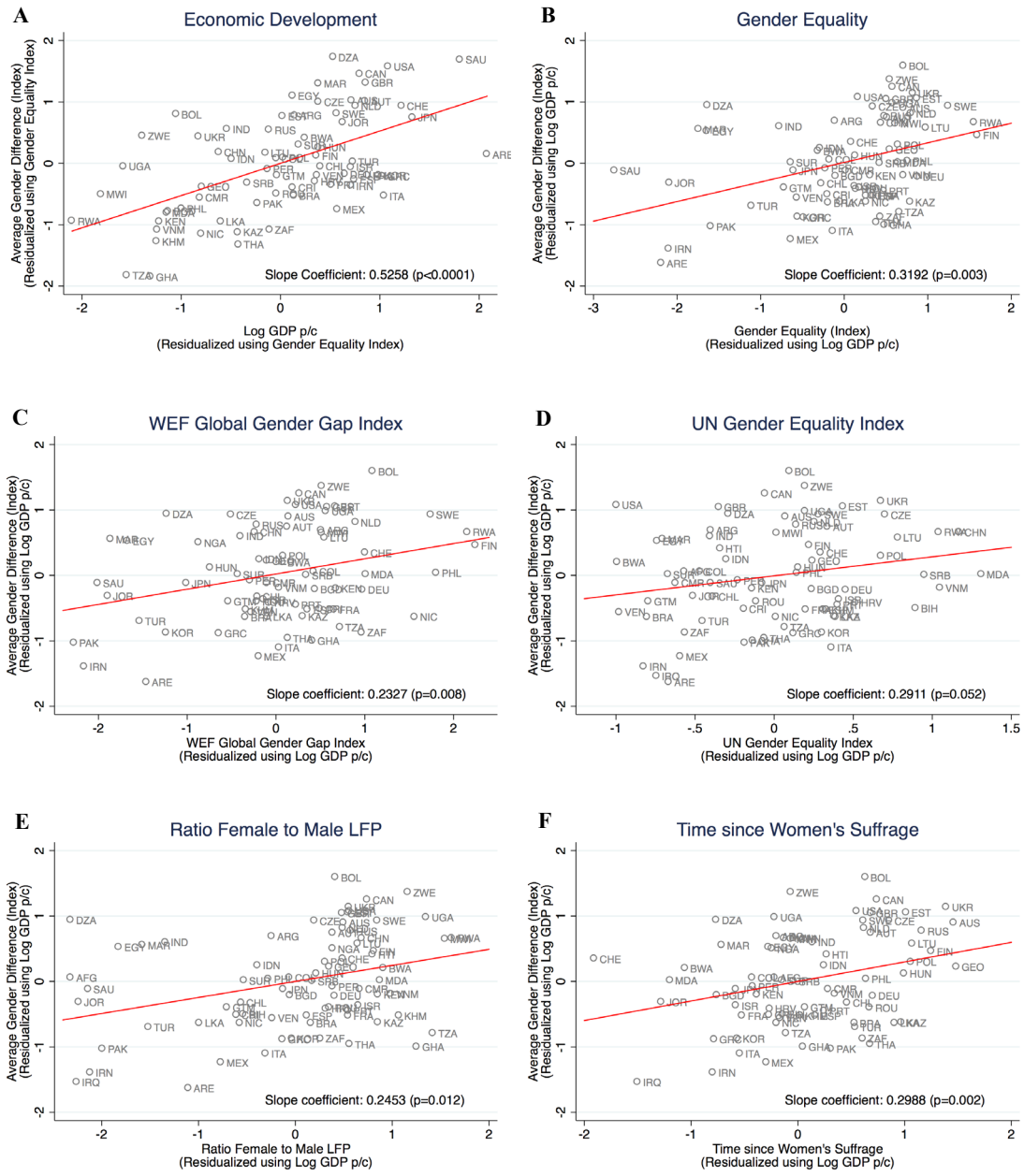


Fig. 2. Analysis of gender differences in preferences in relation to economic development conditional on gender equality and vice versa. Each panel depicts a partial regression plot. (A) The relationship between the aggregate index of gender differences in preferences and log GDP per capita after residualizing both variables with respect to the Gender Equality Index. (B to F) The relationship between the aggregate index of gender differences in preferences and five indices of gender equality after residualizing all variables with respect to log GDP per capita. Indices of gender equality are (B) the Gender Equality Index, (C) WEF Global Gender Gap Index, (D) UN Gender Equality Index, (E) ratio of female to male labor force participation, (F) years since women's suffrage. Country abbreviations are spelled out in the Supplementary Online Material. For corresponding regression evidence see table S4.

Gender Gap Index, 0.29 (p=0.0515) for the UN Gender Equality Index, 0.25 (p=0.0123) for ratio of female to male labor force participation, and 0.30 (p=0.0023) for years since women's suffrage.

In sum, these findings provide evidence in favor of the resource hypothesis that higher levels of economic development and gender quality are associated with stronger gender differentiation in preferences.

A potential concern regarding the reported results involves bias due to culture-specific survey response behavior (*29–32*). Note that our data contain two types of items, qualitative self-assessments and quantitative choice measures. Qualitative self-assessments might be affected by response biases such as scaling effects which might vary across cultures introducing systematic measurement error (*33*). In contrast, the quantitative items present trade-offs that are well-defined in terms of stakes and probabilities yielding revealed preferences measures that facilitate a culturally fair comparison. To test for robustness with regards to the elicitation method, we constructed two separate indices of gender differences using either qualitative or quantitative items only (in an analogous way as the main index). The correlations of the indices with log GDP per capita were found to be very similar, with values of 0.551 ($p<0.0001$) for qualitative and 0.516 ($p<0.0001$) for quantitative items (figs. S7A and B). A test of the null hypothesis of equality of the correlation coefficients failed to reject at conventional significance levels ($p=0.744$). Likewise, correlations with the Gender Equality Index were 0.480 ($p<0.0001$) for qualitative and 0.479 ($p<0.0001$) for quantitative items (figs. S7C and D). Testing equality of the coefficients failed to reject ($p=0.991$), thus providing no support that culture-specific response behavior contaminated the results.

To further test for the robustness of our results, we conducted several additional analyses. First, as trust reflects a composite trait that captures beliefs about others' behavior, prosocial preferences and preferences for risk taking, we repeated our analysis excluding the trust dimension. To do so, we constructed an alternative index of gender differences in preferences in a procedure parallel to the main index but using only the five remaining preferences. Similar to

our main results, this alternative index exhibited a quantitatively large association with economic development and measures of gender equality (tables S5 and S6). Second, we tested whether the level of standardization affected our results. We repeated our analysis employing preference measures standardized at the global rather than the country level. The results using preferences standardized at the global level were similar to our main results (tables S7 and S8, fig. S8). Third, we repeated our analysis without using individual-level controls when calculating gender differences, yielding similar results (tables S9 and S10, fig. S9). Fourth, a common concern in cross-country analysis involves measurement error. As the experimental validation was conducted in Germany, more linguistically similar countries might exhibit smaller measurement error. To test for robustness against this potential confound, we additionally controlled for linguistic distance to German, which left the results qualitatively unchanged (tables S11 and S12). Fifth, to address concerns of aggregation bias, we tested for the relationship between household income and gender differences in preferences in individual-level regressions finding a significant relationship for each preference (table S13). Finally, we tested for a non-linear relationship with economic development. A closer inspection of Fig. 1B suggested a non-linear, convex relationship, which is confirmed by regression analysis (column 2 in table S14). This pattern originated from the fact that richer countries are over-proportionally more gender-equal. Therefore, when we investigated the relationship between the index of gender differences in preferences and log GDP per capita after residualizing both variables with respect to the Gender Equality Index, the relationship was found to be linear (table S14). For details on the robustness tests, see Supplementary Text.

Concluding Remarks

The reported evidence indicates that higher levels of economic development and gender equality are associated with stronger gender differentiation in preferences. These findings may also relate to other personality traits, such as the Big Five *(34, 35)* or value priorities *(36)*. Our findings do not rule out an influence of gender-specific roles that drive gender differences in

preferences. They also do not preclude a role for biological or evolutionary determinants of gender differences *(37)*. Our results highlight, however, that theories not attributing a significant role to the social environment are incomplete *(38)*.

In this regard, our findings point towards the critical role of availability of and equal access to material and social resources for both genders in facilitating the independent formation and expression of gender-specific preferences across countries. As suggested by the resource hypothesis, greater availability of material resources removes the human need of subsistence, and hence provides the scope for attending to gender-specific preferences. A more egalitarian distribution of material and social resources enables both genders to independently express gender-specific preferences.

Materials and Methods

Extended Materials and Methods can be found in the Supplementary Materials.

Experimental selection of survey items and construction of preference measures

Survey items included in the GPS data were selected in an ex-ante experimental validation procedure at the Laboratory for Experimental Economics of the University of Bonn in winter 2010/2011. In this procedure, 402 subjects participated in incentivized choice experiments and responded to a large set of survey items which were either newly developed or taken from existing surveys *(25)*.

Incentivized choice experiments were conducted to obtain an incentivized behavioral measure for each preference: risk taking was measured as the average response to two multiple price lists in which subjects choose between a lottery and varying safe options. Patience was measured as the average response to two multiple price lists in which subjects choose between receiving a payment at the day of the experiment or a larger payment 12 days later. Trust was measured as the average amount sent as a first mover in two investment games. Altruism was measured as first mover behavior in a dictator game with a charitable organization as recipient.

Positive reciprocity was measured as the average amount sent back as a second mover in two investment games. Negative reciprocity was measured as the average amount invested into punishment after unilateral defection of the opponent in a prisoner's dilemma and the minimum acceptable offer in an ultimatum game.

For each preference, we selected those survey items for constructing the GPS which exhibited the highest predictive power for the corresponding incentivized behavioral measure *(25)*. Formally, for each preference the behavioral measure was regressed on different combinations of the survey items. The combination which maximized adjusted R-squared was then selected for the respective preference.

12 survey questions were selected for the GPS which comprised a mixture of qualitative items, measured on an 11-point Likert scale, and quantitative items involving economic tradeoff decisions: risk taking was elicited by (i) an item determining the indifference point between a lottery with 50% chance of winning and receiving a fixed certain payment and (ii) the response to the question *"Please tell me, in general, how willing or unwilling you are to take risks"*. Patience was elicited by (i) an item determining the indifference point between receiving a fixed monetary amount at the day of the survey and a larger amount 12 months later and (ii) the response to the question "*How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?"*. Positive reciprocity was elicited by (i) an item asking for the value of a thank-you gift the respondent is willing to give in return to help by a stranger and (ii) the response to the question *"When someone does me a favor I am willing to return it."*. Negative reciprocity was elicited by responses to the questions (i) *"If I am treated very unjustly, I will take revenge at the first occasion, even if there is a cost to do so."*, (ii) *"How willing are you to punish someone who treats you unfairly, even if there may be costs for you?"*, and (iii) *"How willing are you to punish someone who treats others unfairly, even if there may be costs for you?"*. Altruism was elicited by (i) the quantitative value in response to the question *"Imagine the following situation: Today you unexpectedly received 1,000 Euro. How much of this amount would you donate to a good cause?"* and (ii) the response to the question *"How*

willing are you to give to good causes without expecting anything in return?". Trust was elicited by the response to the question "*I assume that people have only the best intentions."*. For each preference, the final survey measure was given as the weighted average of the z-scores of the corresponding survey items. The weights were calculated as the coefficients in OLS regressions of the incentivized behavioral measures on the respective survey items.

Selection of countries, translation of survey items, and pretest

For the GPS, 76 countries were selected with the goal to provide representative coverage of the global population. As a key criterion, the selected countries covered all development levels and geographic regions, including 24 in Europe, 22 in Asia, 1 in Oceania, 14 in Africa and 15 in the Americas (for a comprehensive list of countries see Supplementary Materials). Further, the selection process aimed at maximizing variation along country characteristics such as language, historical, political, and ecological conditions and favored culturally distinct and non-neighboring countries.

For each country, the selected survey items were translated into the country's major languages involving at least three translators for each language. A first translator suggested, dependent on the region of the target language, an English, French, or Spanish version of the item. A second translator conducted the translation into the target language. A third translator conducted a translation back to the original language. If a discrepancy occurred, the process was iterated until all translators agreed. Furthermore, monetary amounts used in the survey questions were adjusted to correspond to the same share in the median income of the target countries.

The survey items were pretested as part of the Gallup World Poll 2012 pre-test, conducted at the end of 2011 in 22 countries with a sample size of 10 to 15 respondents per country. No respondent indicated problems in understanding the wording or the quantitative content of the survey items. Some respondents suggested rewording which was incorporated through minor adjustments of some survey items (for details see Supplementary Materials).

Sampling and selection of respondents

We included the GPS as part of the Gallup World Poll 2012 through the infrastructure of Gallup *(23)*. Respondents were sampled to achieve national representativeness of the resident population aged 15 and older. Telephone interviews were conducted where at least 80% of the country's population is covered by telephone or where it is the customary survey methodology. Otherwise, face-to-face interviews were conducted.

The selection of households in countries with telephone interviews employed either a random-digit-dialing method or nationally representative lists of phone numbers. In countries with face-to-face interviews, primary sampling units were stratified by population size and/ or geography. To select sampled households, a random-route procedure was employed. Respondents were selected randomly by either the latest birthday or Kish grid method.

References and Notes

1. M. Bertrand, "New Perspectives on Gender" in Handbook of Labor Economics, O. Ashenfelter, D. Card, Eds. (Elsevier, 2011), Vol. 4b, pp. 1543–1590.
2. R. Croson, U. Gneezy, Gender Differences in Preferences. *J. Econ. Lit.* **47**, 448–474 (2009).
3. E. E. Maccoby, C. N. Jacklin, *The psychology of sex differences* (Stanford University Press, 1974).
4. T. DeLeire, H. Levy, Worker Sorting and the Risk of Death on the Job. *J. Labor Econ.* **22**, 925–953 (2004).
5. T. Buser, M. Niederle, H. Oosterbeek, Gender, Competitiveness, and Career Choices. *Q. J. Econ.* **129**, 1409–1447 (2014).
6. D. M. Buss, Psychological sex differences: Origins through sexual selection. *Am. Psychol.* **50**, 164-8; discussion 169-71 (1995).
7. D. C. Geary, *Male, female: The evolution of human sex differences.* (American Psychological Association, Washington, 1998).
8. A. H. Eagly, *Sex differences in social behavior: A social-role interpretation.* (Lawrence Erlbaum Associates, Hillsdale, 1987).
9. A. H. Eagly, W. Wood, The origins of sex differences in human behavior: Evolved dispositions versus social roles. *Am. Psychol.* **54**, 408–423 (1999).
10. U. Gneezy, K. L. Leonard, J. A. List, Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society. *Econometrica*. **77**, 1637–1664 (2009).
11. E. Duflo, Women Empowerment and Economic Development. *J. Econ. Lit.* **50**, 1051–1079 (2012).

12. M. Doepke, M. Tertilt, Women's Liberation: What's in It for Men? *Q. J. Econ.* **124**, 1541–1591 (2009).

13. R. Fernandez, A. Fogli, C. Olivetti, Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *Q. J. Econ.* **119**, 1249–1299 (2004).

14. R. Fernández, A. Fogli, American Economic Association Culture: An Empirical Investigation of Beliefs, Work, and Fertility Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *Am. Econ. J. Macroecon.* **1**, 146–177 (2009).

15. R. Inglehart, *The Silent Revolution: Changing values and political styles among Western publics* (Princeton University Press, Princeton, 1977).

16. R. Inglehart, P. Norris, *Rising tide: Gender equality and cultural change around the world* (Cambridge University Press, Cambridge, 2003).

17. I. Almås, A. W. Cappelen, K. G. Salvanes, E. Ø. Sørensen, B. Tungodden, Willingness to Compete: Family Matters. *Manage. Sci.* **62**, 2149–2162 (2016).

18. J. Haushofer, E. Fehr, On the psychology of poverty. *Science.* **344**, 862–867 (2014).

19. T. Tanaka, C. F. Camerer, Q. Nguyen, Risk and time preferences: Experimental and household survey data from Vietnam. *Am. Econ. Rev.* **100**, 557–571 (2010).

20. V. McLoyd, Socioeconomic disadvantage and child development. *Am. Psychol.* **53**, 185–204 (1998).

21. One may be agnostic about the ultimate determinants of gender-specific preferences. They may either be acquired through cross-culturally universal gender roles or due to biological and evolutionary differences between the genders.

22. A. Falk, A. Becker, T. Dohmen, B. Enke, D. Huffmann, U. Sunde, *Global Evidence on Economic Preferences* (Working Paper No. 23943, National Bureau of Economic

Research, 2017).

23. See also Material and Methods and Supplementary Online Materials for a detailed description of the data. For further details on the Gallup World Poll see http://www.gallup.com/analytics/213704/world-poll.aspx.

24. We note that trust is not a preference but a composite trait, including beliefs about others' behavior, prosocial preferences and preferences for risk tasking. Given its importance, however, we decided to include it in our analysis. The results remained unchanged when excluding trust from the analysis, see Supplementary Text.

25. A. Falk, A. Becker, T. Dohmen, D. Huffman, U. Sunde, "The preference survey module: A validated instrument for measuring risk, time, and social preferences" (IZA Discussion Paper 9674, 2016).

26. The raw gender differences without controls are qualitatively very similar (table S1).

27. This relationship remained robust when controlling for country-specific differences such as geographic, demographic, and historical factors (table S2). A complementary analysis further showed that gender differentiation in preferences was driven by those components of economic development which disproportionately benefit females relatively to males (table S3).

28. We also investigated, separately for each preference, the relationship between gender differences and economic development conditional on gender equality and vice versa. For each preference, gender differences were found to be strongly associated with log GDP per capita conditional on the Gender Equality Index (fig. S5). Likewise, gender differences were found to be highly associated with the Gender Equality Index conditional on log GDP per capita (fig. S6).

29. M. Yuki, Intergroup Comparison versus Intragroup Relationships: A Cross-Cultural Examination of Social Identity Theory in North American and East Asian Cultural Contexts. *Soc. Psychol. Q.* **66**, 166–183 (2003).

30. S. Guimond, A. Chatard, D. Martinot, R. J. Crisp, S. Redersdorff, Social comparison, self-stereotyping, and gender differences in self-construals. *J. Pers. Soc. Psychol.* **90**, 221–242 (2006).

31. S. Guimond *et al.*, Culture, Gender, and the Self: Variations and Impact of Social Comparison Processes. *J. Pers. Soc. Psychol.* **92**, 1118–1134 (2007).

32. S. Guimond, Psychological similarities and differences between women and men across cultures. *Soc. Personal. Psychol. Compass*. **2**, 494–510 (2008).

33. A specific concern involves cross-cultural differences in gender-specific social comparison *(29-32)*: survey respondents in less developed, less gender-egalitarian countries may be inclined to define their self relative to members of their own gender (intra-group social comparison). In contrast, those in more developed and gender-egalitarian countries may be likely to compare themselves to members of both genders (inter-group social comparison). As a consequence, gender differences estimated through self-assessments may be understated in less developed countries.

34. P. T. Costa, A. Terracciano, R. R. McCrae, Gender differences in personality traits across cultures: Robust and surprising findings. *J. Pers. Soc. Psychol.* **81**, 322–331 (2001).

35. D. P. Schmitt, A. Realo, M. Voracek, J. Allik, Why Can't a Man Be More Like a Woman? Sex Differences in Big Five Personality Traits Across 55 Cultures. *J. Pers. Soc. Psychol.* **94**, 168–182 (2008).

36. S. H. Schwartz, T. Rubel, Sex differences in value priorities: Cross-cultural and

multimethod studies. *J. Pers. Soc. Psychol.* **89**, 1010–1028 (2005).

37. An example for the biological role in shaping human preferences comes from twin studies *(39-42)*.

38. For instance, our results can be accommodated by approaches suggesting interactions of evolutionary or biological determinants with the social environment *(43).*

39. D. Cesarini, C. T. Dawes, M. Johannesson, P. Lichtenstein, B. Wallace, Genetic variation in preferences for giving and risk taking. *Q. J. Econ.* **124**, 809-842 (2009).

40. N. Zethraeus, L. Kocoska-Maras, T. Ellingsen, B. O. Von Schoultz, A. L. Hirschberg, M. Johannesson, A randomized trial of the effect of estrogen and testosteron on economic behavior. *Proc. Natl. Acad. Sci.* **106**, 6535-6538 (2009).

41. D. Cesarini, C.T. Dawes, J. H. Fowler, M. Johannesson, P. Lichtenstein, B. Wallace, Heritability of cooperative behavior in the trust game. *Proc. Natl. Acad. Sci.* **105**, 3721-3726 (2008).

42. B. Wallace, D. Cesarini, P. Lichtenstein, M. Johannesson, Heriability of ultimatum game responder behavior. *Proc. Natl. Acad. Sci.* **104**, 15631-15634 (2008).

43. D. C. Geary, Evolution and developmental sex differences. *Current Directions in Psychological Science* **8**, 115-120 (1999).

44. http://www.gallup.com/178667/gallup-world-poll-work.aspx

45. J. Berg, J. Dickhaut, K. McCabe, Trust, Reciprocity, and Social History. *Games Econ. Behav.* **10.1**, 122-142 (1995).

46. R.Selten, "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperimentes" in Beiträge zur experimentellen Wirtschaftsforschung, H. Sauermann, Ed. (J. C. B. Mohr, 1967).

47. W. Güth, R. Schmittberger, B. Schwarze, An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* **3**, 367–388 (1982).

48. A. Falk, E. Fehr, U. Fischbacher, Driving Forces Behind Informal Sanctions. *Econometrica.* **73**, 2017-2030 (2005).

49. Q. Ashraf, O. Galor, The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic develoment. *Am. Econ. Rev.* **103**, 1-46 (2013).

50. W. Nordhaus, Geography and macroeconomics: New data and new findings. *Proc. Natl. Acad. Sci.* **103**, 3510-3517 (2006).

51. S. Michalopoulos, The origins of ethnolinguistic diversity, *Am. Econ. Rev.* **102**, 1508-1539 (2012).

52. A. Alesina, A. Devleeschauwer, W. Easterly, S. Kurlat, R. Waczlarg, Fractionalization, *J. Econ. Growth.* **8**, 155-194 (2003).

Supplementary Materials

Content

Extended Materials and Methods

Extended Materials and Methods

Overview and infrastructure

Materials and Methods contain details on Global Preference Survey (GPS) data collection on altruism, trust, positive and negative reciprocity, risk taking, and patience which was conducted as part of the Gallup World Poll 2012 through the infrastructure of Gallup *(23).* Prior to implementing the GPS, a total of 12 survey items was selected through an ex-ante experimental validation. The survey items were then translated and made internationally comparable. At the end of 2011, a pre-test of the survey items was conducted in 22 countries as part of the Gallup World Poll 2012 pretest. After receiving feedback, minor adjustments were made to the survey items. The GPS was then implemented for a total of 76 countries as part of the Gallup World Poll 2012. For further details on the experimental validation and data collection see *(22, 25, 44).* The individual-level data on preferences are publically available and can be found here: (link will be provided).

Experimental selection and validation of survey items

The experimental selection and validation of survey items through laboratory experiments took place at the Laboratory for Experimental Economics at the University of Bonn in winter 2010/2011. 402 subjects took part in incentivized laboratory experiments and answered survey questions for each of the six preferences. The survey questions which performed as the best joint predictors of incentivized behavior were selected as items for the respective preference in the GPS. The following paragraphs contain details on the experimental validation.

Choice experiments, preference measures, and survey items in the validation

The following section describes the set of incentivized choice experiments and the experimental measure of each preference. An overview table is presented below.

In order to isolate social preferences from repeated game motives, all experiments with social interaction were one-shot. Following a perfect stranger random matching protocol, it was ensured that subjects never interacted more than once with the same person.

Risk taking was elicited through two multiple price lists in which subjects chose between a lottery and varying safe options. The average of the two switching rows served as experimental measure of risk taking. Using the average of two choices relative to the choice in only one experiment reduces measurement error. In a parallel way the measure for patience was elicited through two multiple price lists in which subjects chose between receiving a payment at the day of the experiment or a larger payment 12 months later.

Trust and positive reciprocity were elicited as first and second mover behavior in two investment games *(45)* where the amount sent was either doubled or tripled. Hence, each subject took part in four investment games, twice as first mover, twice as second mover. The contingent response method *(46)* was applied for second mover behavior. The average of choices as first or second mover served as experimental measures of trust and reciprocity, respectively.

Altruism was elicited as donation amount in a dictator game with a charitable organization as recipient.

Negative reciprocity was elicited through two different experiments: a subject's minimum acceptable offer in an ultimatum game *(47)* and a subject's investment into punishment after unilateral defection of their opponent in a prisoner's dilemma *(48).* Both choices were standardized

to account for differences in response scales and averaged to obtain the experimental measure of negative reciprocity.

The choice experiments were accompanied by a large set of qualitative and quantitative survey items. Goal of the experimental validation was to select those survey items for the GPS which were the best predictors of incentivized behavior in the choice experiments. Candidate survey items were taken from existing surveys, others were newly designed for the experimental selection and validation. The full list of survey items can be found in (*25*).

Preference	**Experiment**	**Measure**
Risk Taking	Two multiple price lists in which subjects choose between a lottery and varying safe options	Average of rows in both price lists in which subjects switch from preferring the lottery to the safe option
Time Discounting	Two multiple price lists in which subjects choose between a payment "today" and a larger payment "in 12 months"	Average of rows in two price lists in which subjects switch from preferring the early to the delayed payment
Trust	First mover behavior in two investment games	Average amount sent as a first mover in both investment games
Altruism	First mover behavior in a dictator game with a charitable organization as recipient	Amount of donation
Positive Reciprocity	Second mover behavior in two investment games (contingent response method)	Average amount sent back in both investment games
Negative Reciprocity	Investment into punishment after unilateral defection of the opponent in a prisoner's dilemma (contingent response method) and minimum acceptable offer in an ultimatum game	Average score: amount invested into punishment and minimum acceptable offer in an ultimatum game

Selection of survey items

For each preference, the survey items were selected as the best joint predictors of incentivized behavior. Each experimental preference measure was regressed via OLS on different combinations of the survey items. The best combination in terms of explanatory power, measured by adjusted R-squared, was then identified and selected for the international survey.

Wording of survey items and construction of preference measures

Survey items

Following the experimental validation, a set of 12 survey items was selected for the GPS. For each preference, the exact wording of the corresponding survey items is given below. As indicated below, survey items were either qualitative or quantitative.

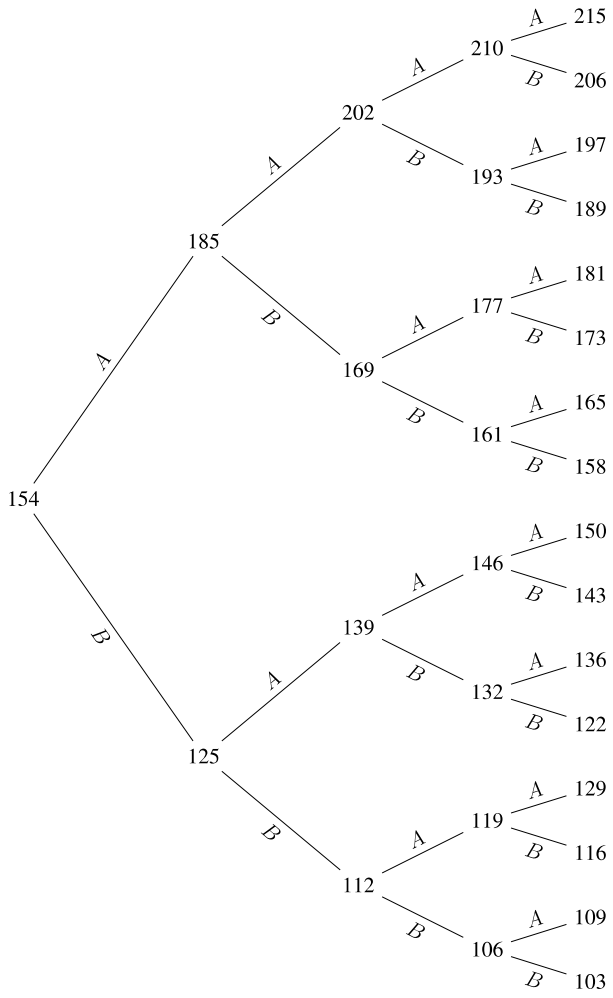
"Willingness to act" survey items indicate the following introduction: "*We now ask for your willingness to act in a certain way in four different areas. Please again indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing 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."*

Likewise, "Self-assessment" survey items were 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."*

I. Patience

1. Sequence of five interdependent questions (quantitative): *"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 displayed below. Numbers correspond to the payment in 12 months. *"A"* indicates the choice of "100 euros today". *"B"* indicates the choice of "*x* euros in 12 months".



The staircase procedure worked as follows. First, each respondent indicated whether they would prefer to receive 100 euros today or 154 euros in 12 months from the date of the interview leftmost decision node). If the respondent chose the payment today ("*A*"), the payment in 12 months was adjusted upwards to 185 euros in the second question. If the respondent chose the payment in 12 months ("*B"*), the corresponding payment was adjusted down to 125 euros. The subsequent steps in the tree followed the same logic.

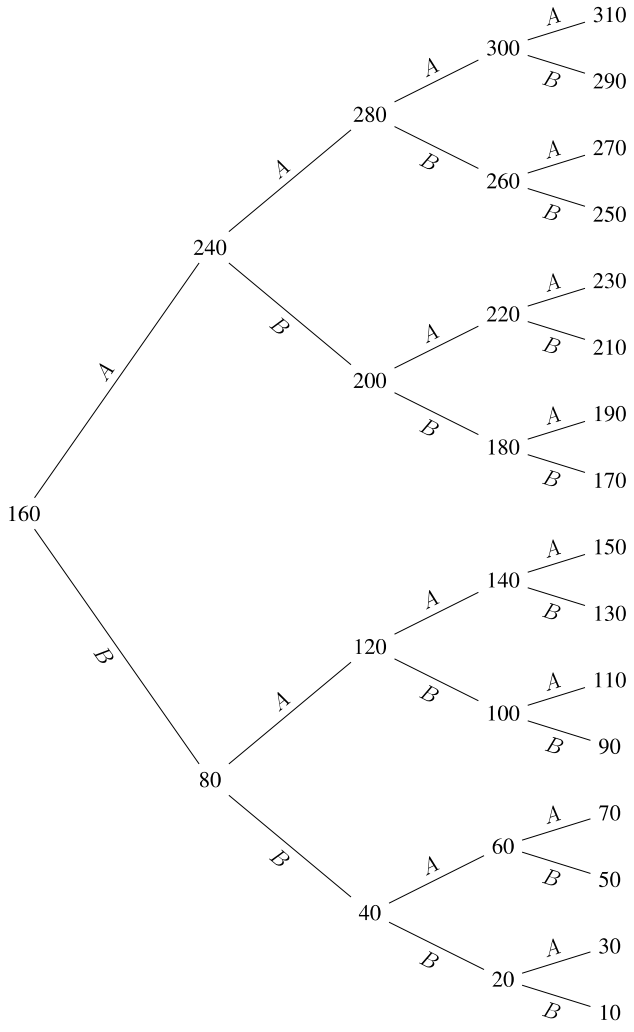
2. Willingness to act (qualitative): "*How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?"*

II. Risk Taking

3. Similar to self-assessment (qualitative): *"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."*

4. Sequence of five interdependent questions (quantitative): *"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 displayed below. Numbers correspond to the sure payment. *"A"* indicates choice of the draw. *"B"* indicates the safe payment option.



   The staircase procedure worked as follows. First, each respondent indicated whether they would prefer to receive 160 euros for sure or whether they preferred a 50:50 chance of receiving 300 euros or nothing. If the respondent chose the safe option ("B"), the safe amount of money being offered in the second question decreased to 80 euros. If the

respondent chose the gamble ("*A*"), the safe amount was increased to 240 euros. The subsequent steps in the tree followed the same logic.

III. Positive Reciprocity

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

6. Choice (quantitative): *"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."*

IV. Negative Reciprocity

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

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

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

V. Altruism

10. Choice (quantitative): *"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.)"*

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

VI. Trust

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

Preference measures

To create the individual-level preference measures the following procedure was employed. First, for each of the 12 survey items z-scores were computed at the individual level. Second, for each preference the respective z-scores were averaged using weights developed in the experimental validation. Technically, these weights had been computed as coefficients in OLS regressions of observed choices in the experimental validation on the respective survey items, restricting the sum of coefficients to one. Weights are given by:

Patience	=	0.7115185 × Staircase patience + 0.2884815 × Willingness to give up something today
Risk taking	=	0.4729985 × Staircase risk + 0.5270015 × Willingness to take risks
Positive reciprocity	=	0.4847038 × Willingness to return favor + 0.5152962 × Size of gift
Negative reciprocity	=	0.6261938 / 2 × Willingness to punish if oneself is treated unfairly + 0.6261938 / 2 × Willingness to punish if other is treated unfairly + 0.3738062 × Willingness to take revenge
Altruism	=	0.6350048 × Willingness to give to good causes + 0.3649952 × Size of donation
Trust	=	1 × Belief people have best intentions

As explained in the context of the global pre-test (see below), the original survey item for negative reciprocity was split up into two items: the first asking for the willingness to punish if oneself was treated unfairly and the second asking for the willingness to punish if someone was treated unfairly. To apply the weighting procedure from the experimental validation, the corresponding weight was divided by two and applied to the two new modified items.

Pretest

The global survey was pre-tested in the Gallup World Poll 2012 pre-test, conducted in the end of 2011. The pre-test was conducted in 22 countries, including 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) with country-sample sizes between 10 and 15 respondents. Goal of the pretest was to receive feedback on whether survey items were understandable and/or whether there were cultural differences in the interpretation of survey items. Pre-test respondents were instructed to indicate difficulties in understanding the survey items and were invited to offer suggestions for rewording.

With regards to the quantitative items, no respondent had any problem in understanding the wording and probabilities used in the survey items. With regards to qualitative items, most respondents understood the survey items when being asked to rephrase the respective item in their own words. Some few respondents made suggestions for rewording of the items which led to an adjustment of four items compared to the original (experimentally validated) items.

1. For a few Muslim participants the word "lottery" was problematic. As a consequence, the word "lottery" was replaced by "draw".

2. In some Eastern European and Central Asian countries, the word "charity" was not well understood and hence replaced by "good cause".

3. Some respondents asked for clarification with regards to the item about one's willingness to punish unfair behavior. As a consequence, this item was split up into two items, one asking for one's willingness to punish unfair behavior towards others, the other for one's willingness to punish unfair behavior towards oneself.

4. In the context of the item eliciting choices between monetary amounts today versus one year later, some few respondents, especially in countries with high inflation rates, were stating that their answer would depend on the inflation rate. As a consequence, the following clarification phrase was added "Please assume there is no inflation, i.e., future prices are the same as today's prices."

In addition, the format of the survey questions was made consistent with the Gallup World Poll questionnaire style.

Selection of countries

Countries were selected to provide representative coverage of the global population. A key objective of the selection process was to include all geographic regions and development levels. Additionally, the selection aimed at maximizing variation along country characteristics such as language, historical and political conditions, and ecological features. Furthermore, the selection process aimed to include non-neighboring and culturally distinct countries. The following table lists the sampled countries (including abbreviations), sample sizes for each country, and interview modes.

Abbreviation	Country	Sample Size	Interview Mode
AFG	Afghanistan	1000	Face-to-Face
ARE	United Arab Emirates	1000	Face-to-Face
ARG	Argentina	1000	Face-to-Face
AUS	Australia	1002	Landline/Cellular Phone
AUT	Austria	1001	Landline/Cellular Phone
BGD	Bangladesh	999	Face-to-Face
BIH	Bosnia and Herzegovina	1004	Face-to-Face
BOL	Bolivia	998	Face-to-Face
BRA	Brazil	1003	Face-to-Face
BWA	Botswana	1000	Face-to-Face
CAN	Canada	1001	Landline/Cellular Phone
CHE	Switzerland	1000	Landline/Cellular Phone
CHL	Chile	1003	Face-to-Face
CHN	China	2574	Face-to-Face, Landline Phone
CMR	Cameroon	1000	Face-to-Face
COL	Colombia	1000	Face-to-Face
CRI	Costa Rica	1000	Face-to-Face
CZE	Czech Republic	1005	Face-to-Face
DEU	Germany	997	Landline/Cellular Phone
DZA	Algeria	1022	Face-to-Face
EGY	Egypt	1020	Face-to-Face
ESP	Spain	1000	Landline/Cellular Phone
EST	Estonia	1004	Face-to-Face
FIN	Finland	1000	Landline/Cellular Phone
FRA	France	1001	Landline/Cellular Phone
GBR	United Kingdom	1030	Landline/Cellular Phone
GEO	Georgia	1000	Face-to-Face
GHA	Ghana	1000	Face-to-Face
GRC	Greece	1000	Face-to-Face
GTM	Guatemala	1000	Face-to-Face
HRV	Croatia	992	Face-to-Face
HTI	Haiti	504	Face-to-Face
HUN	Hungary	1004	Face-to-Face
IDN	Indonesia	1000	Face-to-Face
IND	India	2539	Face-to-Face
IRN	Iran	2507	Landline/Cellular Phone
IRQ	Iraq	1000	Face-to-Face

ISR	Israel	999	Face-to-Face

Abbreviation	Country	Sample Size	Interview Mode
ITA	Italy	1004	Landline/Cellular Phone
JOR	Jordan	1000	Face-to-Face
JPN	Japan	1000	Landline Phone
KAZ	Kazakhstan	999	Face-to-Face
KEN	Kenya	1000	Face-to-Face
KHM	Cambodia	1000	Face-to-Face
KOR	South Korea	1000	Landline/Cellular Phone
LKA	Sri Lanka	1000	Face-to-Face
LTU	Lithuania	999	Face-to-Face
MAR	Morocco	1000	Face-to-Face
MDA	Moldova	1000	Face-to-Face
MEX	Mexico	1000	Face-to-Face
MWI	Malawi	1000	Face-to-Face
NGA	Nigeria	1000	Face-to-Face
NIC	Nicaragua	1000	Face-to-Face
NLD	Netherlands	1000	Landline/Cellular Phone
PAK	Pakistan	1004	Face-to-Face
PER	Peru	1000	Face-to-Face
PHL	Philippines	1000	Face-to-Face
POL	Poland	999	Face-to-Face
PRT	Portugal	998	Landline/Cellular Phone
ROU	Romania	994	Face-to-Face
RUS	Russian Federation	1498	Face-to-Face
RWA	Rwanda	1000	Face-to-Face
SAU	Saudi Arabia	1035	Face-to-Face
SRB	Serbia	1023	Face-to-Face
SUR	Suriname	504	Face-to-Face
SWE	Sweden	1000	Landline/Cellular Phone
THA	Thailand	1000	Face-to-Face
TUR	Turkey	1000	Face-to-Face
TZA	Tanzania	1000	Face-to-Face
UGA	Uganda	1000	Face-to-Face
UKR	Ukraine	1000	Face-to-Face
USA	United States	1072	Landline/Cellular Phone
VEN	Venezuela	999	Face-to-Face

VNM	Vietnam	1000	Face-to-Face
ZAF	South Africa	1000	Face-to-Face
ZWE	Zimbabwe	1000	Face-to-Face

Survey item translation and cross-country adjustment of monetary amounts

Survey items were translated into the languages of each country according to the following procedure. To make sure that no idiosyncratic errors occurred, at least three translators were involved for each translation of an item in a specific target language. A first translator proposed, depending on the region, an English, French, or Spanish version of the item. A second translator proficient in English, French, or Spanish and the target language conducted the translation to the target language. A third translator translated the item back to the original language. If discrepancies between the original item and the back-translated item occurred, the procedure was repeated until all translators came to an agreement.

Monetary amounts in the quantitative items were made comparable across countries. To do so, monetary amounts were adjusted to correspond to the same share in median income (in the local currency) as the share in German median income (in the original item that was experimentally validated). To avoid cross-country differences in comprehensibility and to preserve simplicity of the items, monetary amounts were rounded.

Sampling and selection of respondents

The within-country sampling of respondents was conducted to achieve national representativeness of the resident population aged 15 and older. The area of coverage generally included the entire country. Exceptions in this regard included areas where the safety of the survey interviewers was endangered and, in some countries, scarcely populated islands. Interviews were either conducted via landline/cellular phone or face-to-face. Telephone interviews were conducted where telephone coverage represents 80% or more of the country's population or is the customary survey methodology.

Depending on the interview mode, the selection of respondents was conducted as follows. In countries where telephone interviews were conducted, either a random-digit-dialing method or nationally representative lists of phone numbers were used. At least three attempts were taken to reach a person in each household. In countries where face-to-face interviews were conducted, primary sampling units were first identified. Primary sampling units, consisting of clusters of households, were stratified by population size and/ or geography. To select sampled households a random-route procedure was employed. Selected households were contacted up to three times (at different times of the day or on different days). A substitution method was employed if the initially sampled household could not be interviewed. In both face-to-face and telephone interviews respondents were selected randomly by either the latest birthday or Kish grid method.

Definition of additional individual-level variables

Education level. Variable ranges from 1 to 3 according to the following classification. 1: Completed elementary education or less (up to 8 years of basic education). 2: Secondary to 3-year tertiary education and some education beyond secondary education (9-15 years of education). 3: Completed four years of education beyond high school and/or received a 4-year college degree.

Household income quintile. Variable ranges from 1 to 5 according to the respondent's household income quintile within the country.

Subj. math skills. Self-assessment of the statement *"I am good at math"* on an 11-point Likert scale.

Definition of country-level variables (including sources)

Time since women's suffrage. Taken from the Inter-Parliamentary Union Website (http://www.ipu.org/wmn-e/suffrage.htm#Note1). For countries where data were missing data were added from the World Economic Forum Global Gender Gap Report 2006 (http://www3.weforum.org/docs/WEF_GenderGap_Report_2006.pdf).

WEF Global Gender Gap Index. Taken from the World Economic Forum Global Gender Gap Report 2015 (http://reports.weforum.org/global-gender-gap-report-2015/rankings/).

UN Gender Inequality Index. Taken from the Human Development Report 2015 (http://hdr.undp.org/en/composite/GII). Values inverted to create an index of equality.

Ratio of female and male labor force participation. Average International Labour Organization estimates from 2003 to 2012 taken from the World Bank database (http://data.worldbank.org/indicator/SL.TLF.CACT.FM.ZS).

Male and female GNI p/c. Taken from the Human Development Report 2015 (http://hdr.undp.org/en/composite/GDI).

Longitude, absolute latitude, area. Taken from the CEPII geo database.

Mean of elevation. Elevation in km above sea level, taken from (*49*). Data originally based on geospatial elevation data reported by the G-ECON project (*50*).

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

Percentage of arable land. Fraction of land within a country which is arable, taken from the World Bank Development Indicators.

Land suitability for agriculture. Index of the suitability of land for agriculture based on ecological indicators of climate suitability for cultivation, such as growing degree days and the ratio of actual to potential evapotranspiration, as well as eco-logical indicators of soil suitability for cultivation, such as soil carbon density and soil pH, taken from *(51)*.

Precipitation. Average monthly precipitation of a country in mm per month, 1961-1990, taken from (*49*). Data originally based on geospatial average monthly precipitation data for this period reported by the G-ECON project *(50)*.

Temperature. Average monthly temperature of a country in degree Celsius, 1961- 1990, taken from (*49*). Data originally based on geospatial average monthly temperature data for this period reported by the G-ECON project *(50)*.

Percentage at risk of malaria. The percentage of population in regions of high malaria risk (as of 1994), multiplied by the proportion of national cases involving the fatal species of the malaria pathogen, P. falciparum. This variable was originally constructed by (*49*) and is part of Columbia University's Earth Institute data set on malaria. Data taken from (*50*).

Predicted genetic diversity. Predicted genetic diversity of the contemporary population, adjusted for post-Columbian migration flows and genetic distance between ethnic groups. See (*49*).

Median age. Taken from the World Bank database.

Ethnic, linguistic, and religious fractionalization. Indices due to (*52*) capturing the probability that two randomly selected individuals from the same country will be from different ethnic (religious) groups.

Linguistic distance to Germany. Computed as the linguistic distance of a country's major language to German based on the ASJP Database, version 18 (http://asjp.clld.org).

Share of atheists. Source: Religion Adherence Data by Robert Barro (http://scholar.harvard.edu/barro/publications/religion-adherence-data).

Colonization indicator. Indicator equal to one if the respective country had at least one colonizer over a long period of time and with substantial participation in governance. Source: the CEPII geo database.

Years of civil and interstate conflict between 1800 and 2007. Taken from the Correlates of War database.

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

Details on statistical analysis

This section describes details of the statistical analysis. We first describe the construction of measures of gender differences in preferences. Then, we provide details on the construction of figures using residualized variables.

Computation of country-level gender differences in preferences

On the country level, gender differences for each of the six preferences (p) were computed as follows. First, each preference was standardized at the country level. Second, for each preference the following individual-level OLS regression was performed separately for each country c,

$$p_i = \beta_1^c female_i + \beta_2^c age_i + \beta_3^c age_i^2 + \beta_4^c education\ level_i + \beta_5^c income\ quintile_i + \beta_6^c\ subjective\ math\ skills_i + \varepsilon_i$$

The obtained coefficient β_1^c on the dummy for female ($female_i$) served as measure of the country-level gender difference for country c in the respective preference. Including controls in the estimation isolates the gender difference from potentially confounding factors which differ between the genders.

Summary index of country-level gender differences in preferences

The country-level summary index of country-level gender differences in preferences was computed as follows. First, we performed a principal component analysis of the country-level gender differences in the six preferences. The predicted first main component then served as the summary index of average gender differences in preferences.

Global gender differences in preferences

On the global level, gender differences for each of the six preferences (p) and associated confidence intervals (fig. S1) were computed as follows. First, each preference was standardized at the global level. Second, for each preference the following individual-level OLS regression with country fixed effects (c_i) was performed on the global sample,

$$p_i = \beta_1 female_i + \beta_2 age_i + \beta_3 age_i^2 + \beta_4 education\ level_i + \beta_5 income\ quintile_i + \beta_6 subjective\ math\ skills_i + c_i + \varepsilon_i \quad .$$

The obtained coefficient β_1 on the dummy for female ($female_i$) served as measure of the global gender difference in the respective preference. Including controls in the estimation isolates the gender difference from potentially confounding factors which differ between the genders. Confidence intervals were computed from standard errors clustered at the country-level. In alternative specifications we calculated unconditional gender differences in a parallel way without using controls. Gender differences obtained from this alternative approach were found to be similar (table S1).

Construction of partial regression plots

The visualization of results employed partial regression plots which show the relationship of residualized variables. Intuitively, a partial regression plot of residual values of variables y and x using for the residualization variable z shows the relationship between variables y and x controlling for z. Technically, for constructing such a figure, we first performed two OLS regressions regressing y on z and x on z. We then calculated the residuals $r_x = x - \hat{x}$ and $r_y = y - \hat{y}$, where $\hat{x}$ and $\hat{y}$ are the predicted values based on the OLS regressions. The partial regression plot of residual values of variables y and x using for the residualization variable z then shows the relationship of r_y and r_x.

Supplementary Text

This section describes the details of the supplemental analysis. There were two main purposes of the supplemental analysis: first, to further analyze the relationship with economic development and gender equality for the six preference measures separately, and second, to test for robustness against potential confounds.

Results on individual preferences

For all preferences, gender differences featured a quantitatively large and significant relationship with log GDP p/c (fig. S2). The correlations were particularly large for trust (0.5918, $p<0.0001$) and altruism (0.5847, $p<0.0001$). The correlations were smaller but statistically significant for positive reciprocity (0.3086, $p=0.0067$), negative reciprocity (0.3542, $p=0.0017$), risk taking (0.3685, $p=0.0011$), and patience (0.3837, $p=0.0006$).

We also investigated the relationship of gender differences in preferences with the Gender Equality Index (fig. S3). The correlations were large and significant for five out of six preferences: trust (0.4050, $p=0.0005$), altruism (0.5073, $p<0.0001$), negative reciprocity (0.4035, $p=0.0005$), risk taking (0.3412, $p=0.0036$), patience (0.4257, $p=0.0002$). The correlation was smaller and insignificant for positive reciprocity (0.1280, $p=0.2875$).

To separate the impacts of economic development and gender equality, we conducted a residual analysis. We first conducted this analysis for economic development residualizing with respect to the Gender Equality Index. To do so, we first regressed the country-level gender differences in the respective preference on the Gender Equality Index. We then predicted the residual values of the gender differences in the respective preference. Next, we regressed log GDP p/c on the Gender Equality Index and predicted the residual values of log GDP p/c. The correlation between the residualized values of gender differences and log GDP p/c represents the relationship controlling for the Gender Equality Index. Similar to the unconditional results, they were particularly large for trust (0.4574, $p=0.0001$) and altruism (0.4751, $p<0.0001$). Correlations were found to be smaller but statistically significant for positive reciprocity (0.2771, $p=0.0193$), negative reciprocity (0.2444, $p=0.0400$), risk taking (0.2868, $p=0.0153$), and patience (0.2621, $p=0.0273$) (fig. S5).

In an analogous way, we conducted a residual analysis for the Gender Equality Index. To do so, we residualized the gender differences in each preference as well as the Gender Equality Index with respect to log GDP p/c. The correlations of residualized values (fig. S6) were positive and statistically significant (at least at the 10% level) for trust (0.2050, $p=0.0863$), altruism (0.3304, $p=0.0049$), negative reciprocity (0.2788, $p=0.0185$), risk taking (0.1973, $p=0.0991$), and patience (0.2967, $p=0.0120$). Positive reciprocity exhibited no systematic correlation (-0.0115, $p=0.9242$).

Results excluding trust

Trust is by definition not a preference but a joint measure capturing beliefs about others' behavior as well as prosocial preferences and preferences for risk taking. However, given its importance we included it in our main analysis. To test for robustness, we created a country-level summary index of gender differences in preferences excluding trust. This alternative index was constructed in a parallel way as the main index but using gender differences for the five remaining preferences only (excluding trust).

Results on the relationship with economic development and gender equality using this alternative index (tables S5 and S6) confirmed our main findings and led to results similar both in terms of the size of the coefficients as well as in terms of statistical significance.

Results using preferences standardized at the global level

In the main specifications, country-level gender differences for each preference were calculated after standardizing each preference on the country level. In alternative specifications, we calculated country-level gender differences after standardizing each preference on the global level. The relationship between these alternative estimates and log GDP p/c (fig. S8) was similar to our main results in terms of magnitude and statistical significance for all preferences: trust (0.5787, $p<0.0001$), altruism (0.5505, $p<0.0001$), positive reciprocity (0.2819, $p=0.0136$), negative reciprocity (0.2980, $p=0.0089$), risk taking (0.2974, $p=0.0091$), and patience (0.4391, $p=0.0001$).

Using these alternative estimates of gender differences, we additionally constructed an alternative summary index of gender differences in preferences in a parallel way as the main index. Results on the relationship with economic development and gender equality using this alternative index (tables S7 and S8) confirmed our main findings and led to results similar both in terms of the size of the coefficients as well as in terms of statistical significance.

Results without controls

In the main specifications, country-level gender differences for each preference were calculated conditional on individual-level controls. In alternative specifications, we calculated country-level gender differences without using individual-level controls.

The relationship between these alternative estimates and log GDP p/c (fig. S9) was similar to our main results in terms of magnitude and statistical significance for all preferences: trust (0.5434, $p<0.0001$), altruism (0.5808, $p<0.0001$), positive reciprocity (0.2748, $p=0.0163$), negative reciprocity (0.4038, $p=0.0003$), risk taking (0.3860, $p=0.0006$), and patience (0.4830, $p<0.0001$).

Using these alternative estimates of gender differences, we additionally constructed an alternative summary index of gender differences in preferences in a parallel way as the main index. Results on the relationship with economic development and gender equality using this alternative index (tables S9 and S10) confirmed our main findings and led to results similar both in terms of the size of the coefficients as well as in terms of statistical significance.

Results controlling for linguistic distance to Germany

In further specifications, we tested whether results were driven by linguistic differences with Germany, where the experimental validation of survey items took place. Therefore, we repeated our analysis controlling for a country's linguistic distance to Germany. The results were found to be qualitatively very similar (tables S11 and S12).

Results from individual-level regressions

The main analysis was conducted on the country level. To address concerns of aggregation bias, we conducted additional individual-level analysis. In particular, we regressed each preference (p), standardized at the country-level, on a gender indicator with male as the reference category, log household income per capita, and their interaction. Log household income per capita was standardized to exhibit a mean of zero and standard deviation of one. Furthermore, we included as controls age, age squared, subjective math skills, education level, and country fixed effects c_i.

$$p_i = \beta_1 female_i + \beta_2 female_i \times \log household\ income\ p/c_i + \beta_3 \log household\ income\ p/c_i + \beta_4 age_i + \beta_5 age_i^2 + \beta_6 subjective\ math\ skills_i + \beta_7 education\ level_i + c_i + \varepsilon_i.$$

Standard errors were clustered at the country level. Results from the individual-level regressions (table S13) were similar to the country-level results: for the average individual, gender differences were 0.072 ($p<0.001$) for trust, 0.110 ($p<0.001$) for altruism, 0.056 ($p<0.001$) for positive reciprocity, -0.137 ($p<0.001$) for negative reciprocity, -0.179 ($p<0.001$) for risk taking, and -0.049 ($p<0.001$) for patience.

Most importantly, gender differences were found to significantly increase with an increase in household income per capita. In particular, a one-standard deviation increase in log household income per capita magnified gender differences in standard deviations by 0.069 ($p<0.001$) for trust, 0.060 ($p<0.001$) for altruism, 0.017 ($p=0.066$) for positive reciprocity, 0.024 ($p=0.028$) for negative reciprocity, 0.028 ($p=0.025$) for risk taking, and 0.040 ($p<0.001$) for patience.

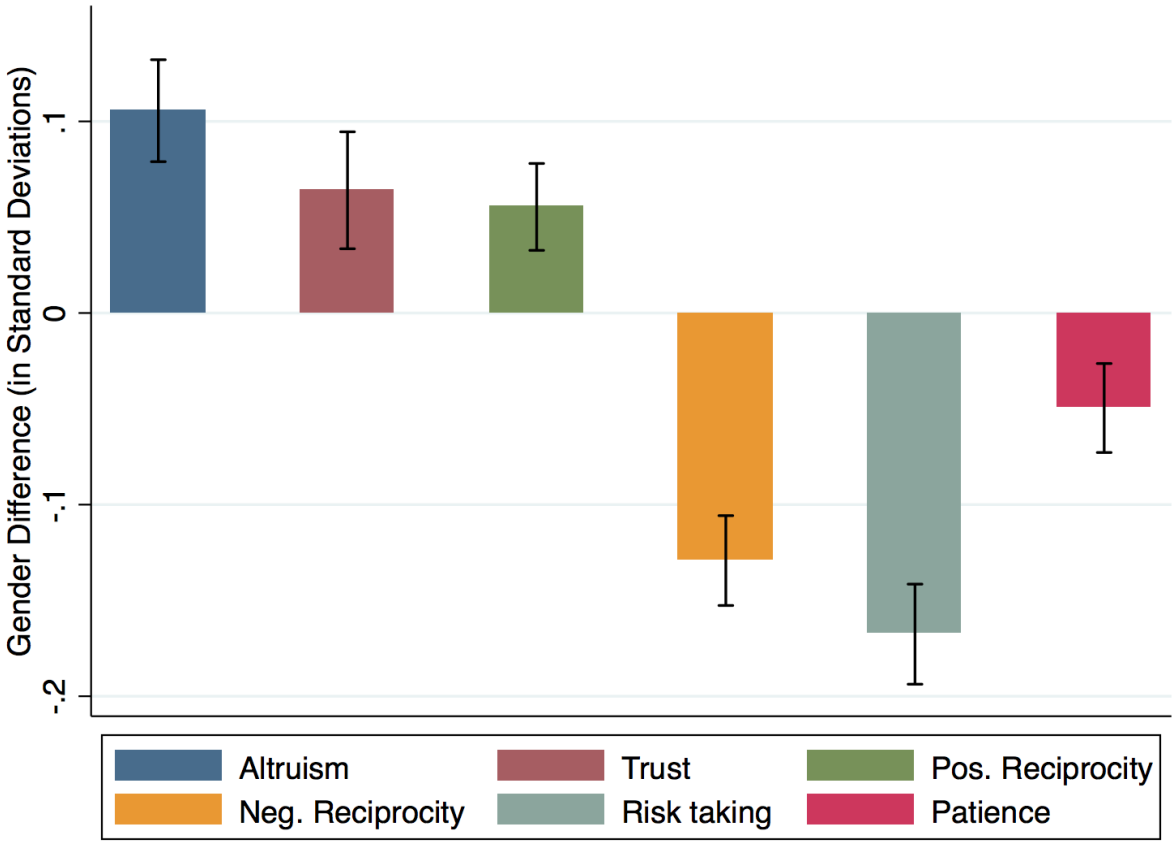


Fig. S1. Gender differences in preferences on the global level.

Positive values indicate that women exhibited higher levels of the respective preference, negative values indicate that women exhibited lower levels of the respective preference. For each preference, the gender difference was calculated as the coefficient on a gender indicator with male as the reference category in an OLS regression of the respective preference on the gender indicator, controlling for age, age squared, subjective math skills, education level, household income quintile, and country fixed effects on the worldwide sample. Error bars indicate 95% confidence intervals obtained from standard errors clustered at the country level.

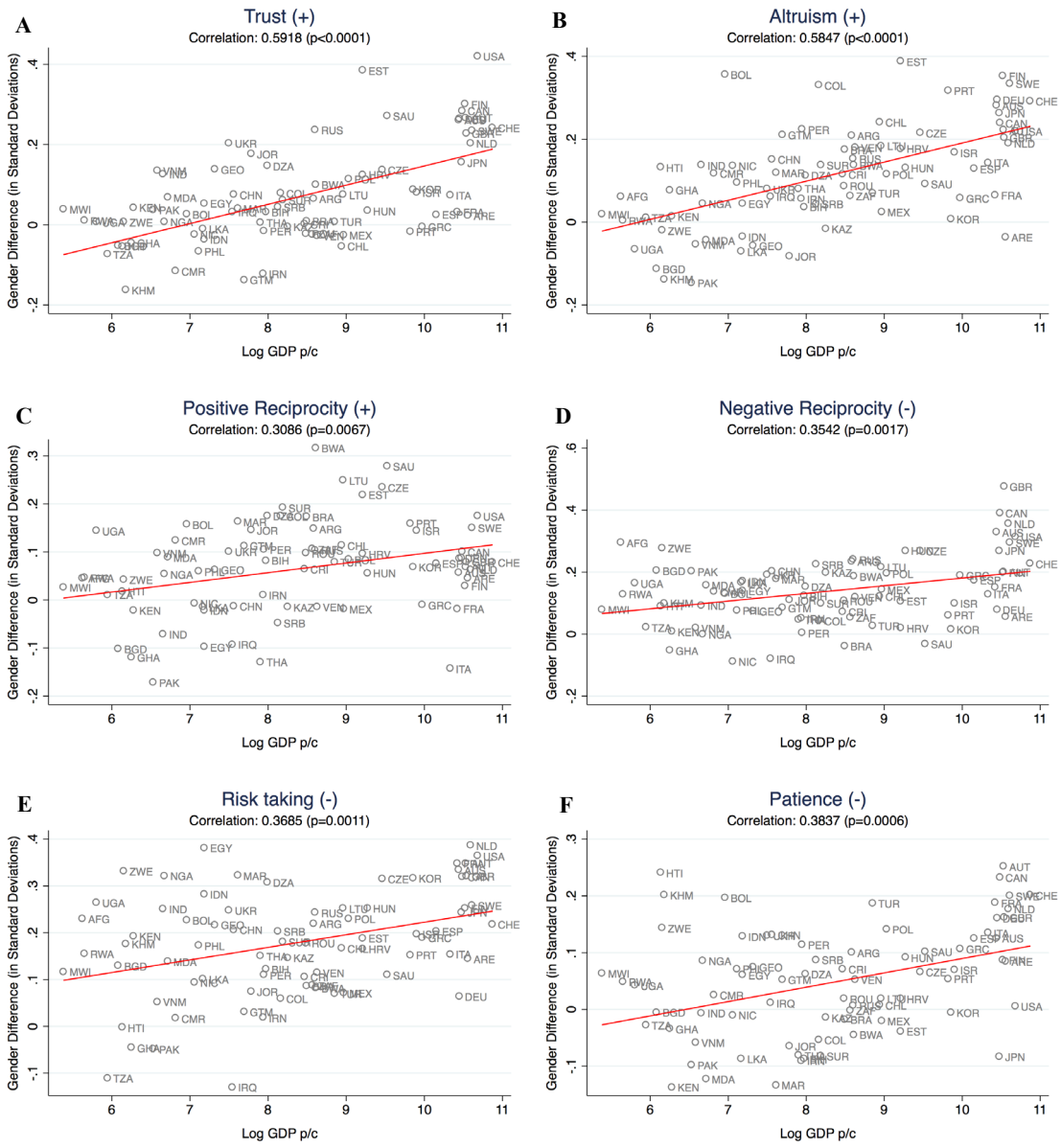


Fig. S2. Gender differences and economic development by preference and country.

Each panel shows the relationship between country-level gender differences in a particular preference and the level of economic development measured by log GDP p/c, including a linear fit. For each preference, the symbols (+)/(-) in the panel titles indicate the direction of the difference. (+) indicates that women exhibited higher levels of the respective preference if the difference was positive. (-) indicates that men exhibited higher levels of the respective preference if the difference was positive. For each preference and country, the gender difference was calculated as the coefficient on a gender indicator with male as the reference category in an OLS regression of the respective preference on the gender indicator, controlling for age, age squared, subjective math skills, education level, household income quintile for the particular country sample.

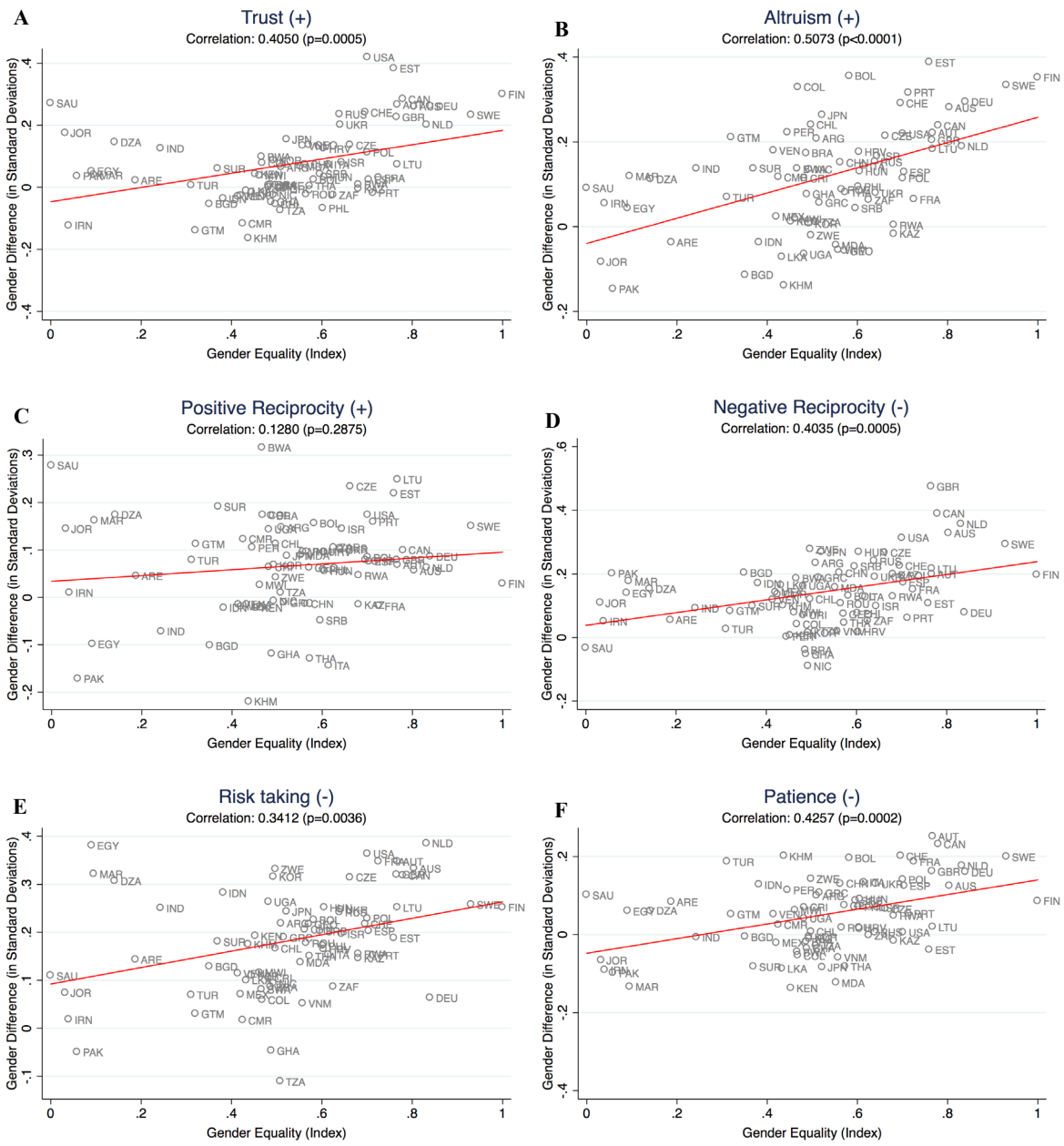


Fig. S3. Gender differences and gender equality by preference and country.

Each panel shows the relationship between country-level gender differences in a particular preference and the Gender Equality Index, including a linear fit. For each preference, the symbols (+)/(-) in the panel titles indicate the direction of the difference. (+) indicates that women exhibited higher levels of the respective preference if the difference was positive. (-) indicates that men exhibited higher levels of the respective preference if the difference was positive. For each preference and country, the gender difference was calculated as the coefficient on a gender indicator with male as the reference category in an OLS regression of the respective preference on the gender indicator, controlling for age, age squared, subjective math skills, education level, household income quintile for the particular country sample.

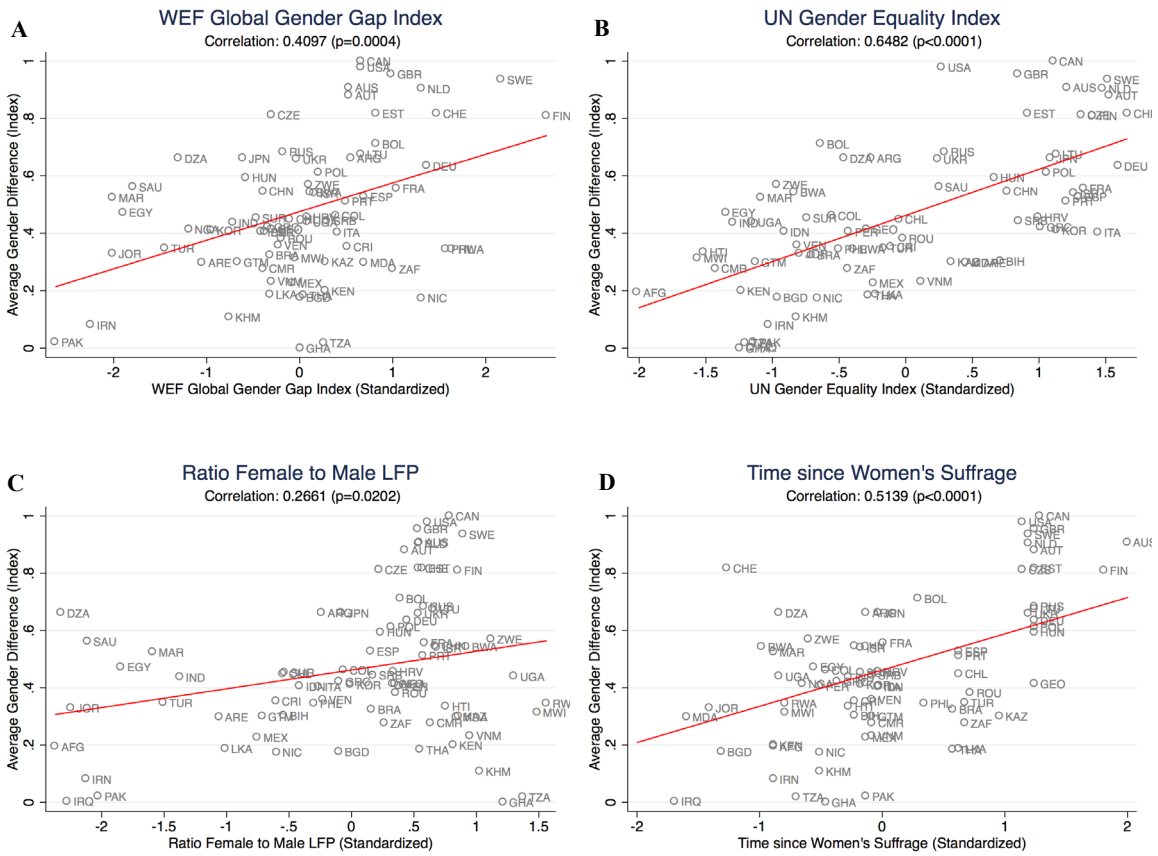


Fig. S4. Gender differences and gender equality by equality index.

Each panel shows the relationship between the index of gender differences in preferences and an indicator for gender equality, including a linear fit. (A) the Global Gender Gap Index of the World Economic Forum (WEF), (B) the Gender Equality Index of the United Nations (UN), (C) the ratio of female and male labor force participation rates, and (D) years since women's suffrage.

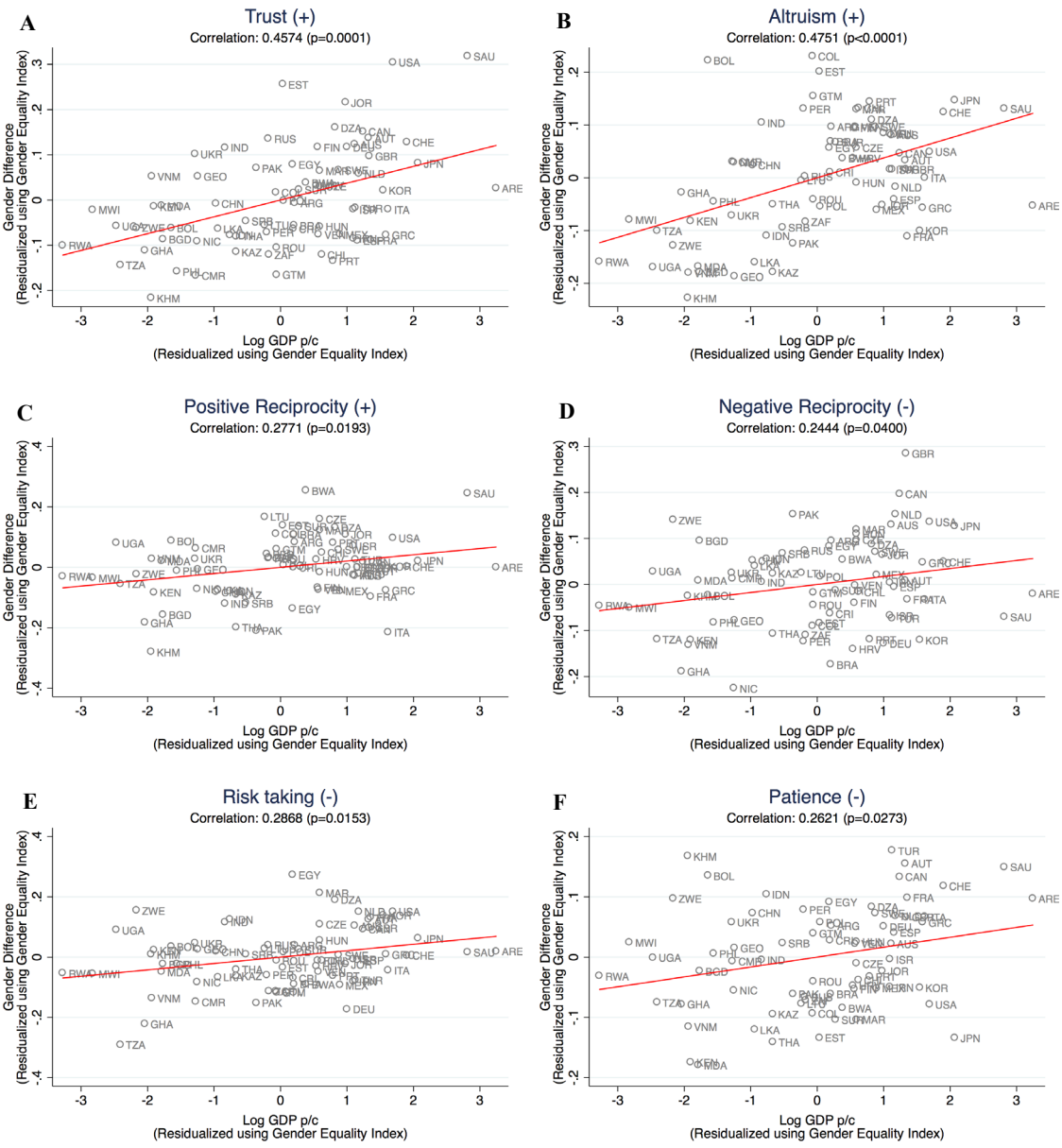


Fig. S5. Gender differences and economic development conditional on gender equality by preference and country.

Each panel shows the relationship between country-level gender differences in a particular preference and economic development measured by log GDP p/c, including a linear fit. Gender differences in preferences and log GDP p/c were residualized with respect to the Gender Equality Index. For each preference, the symbols (+)/(-) in the panel titles indicate the general direction of the difference. (+) indicates that women generally exhibited higher levels of the respective preference. (-) indicates that men generally exhibited higher levels of the respective preference. For each preference and country, the gender difference was calculated as the coefficient on a gender indicator with male as the reference category in an OLS regression of the respective preference on the gender indicator, controlling for age, age squared, subjective math skills, education level, household income quintile for the particular country sample.

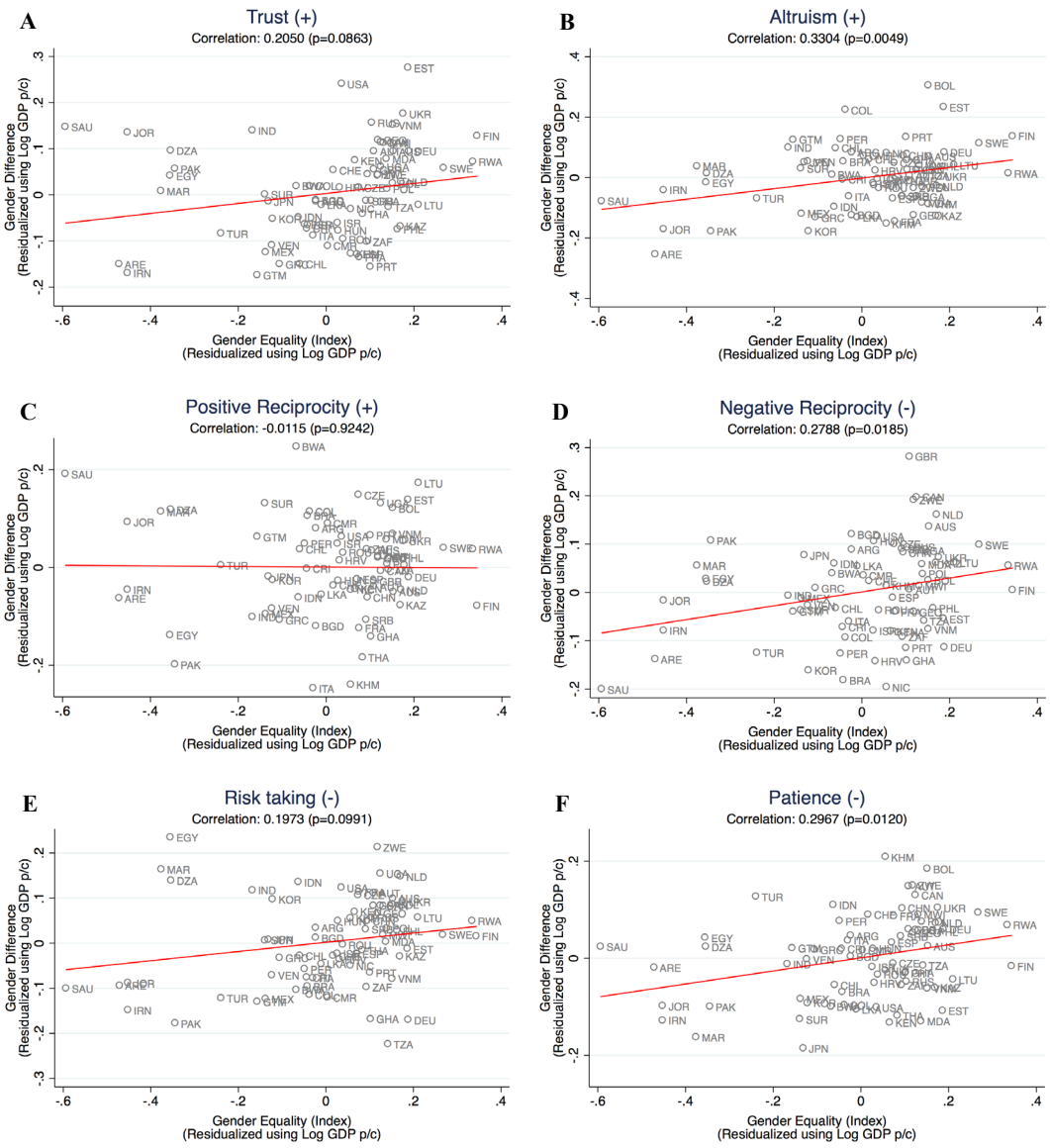


Fig. S6. Gender differences and gender equality conditional on economic development by preference and country.

Each panel shows the relationship between country-level gender differences in a particular preference and the Gender Equality Index, including a linear fit. Gender differences in preferences and the Gender Equality Index were residualized with respect to log GDP p/c. For each preference, the symbols (+)/(-) in the panel titles indicate the general direction of the difference. (+) indicates that women generally exhibited higher levels of the respective preference. (-) indicates that men generally exhibited higher levels of the respective preference. For each preference and country, the gender difference was calculated as the coefficient on a gender indicator with male as the reference category in an OLS regression of the respective preference on the gender indicator, controlling for age, age squared, subjective math skills, education level, household income quintile for the particular country sample.

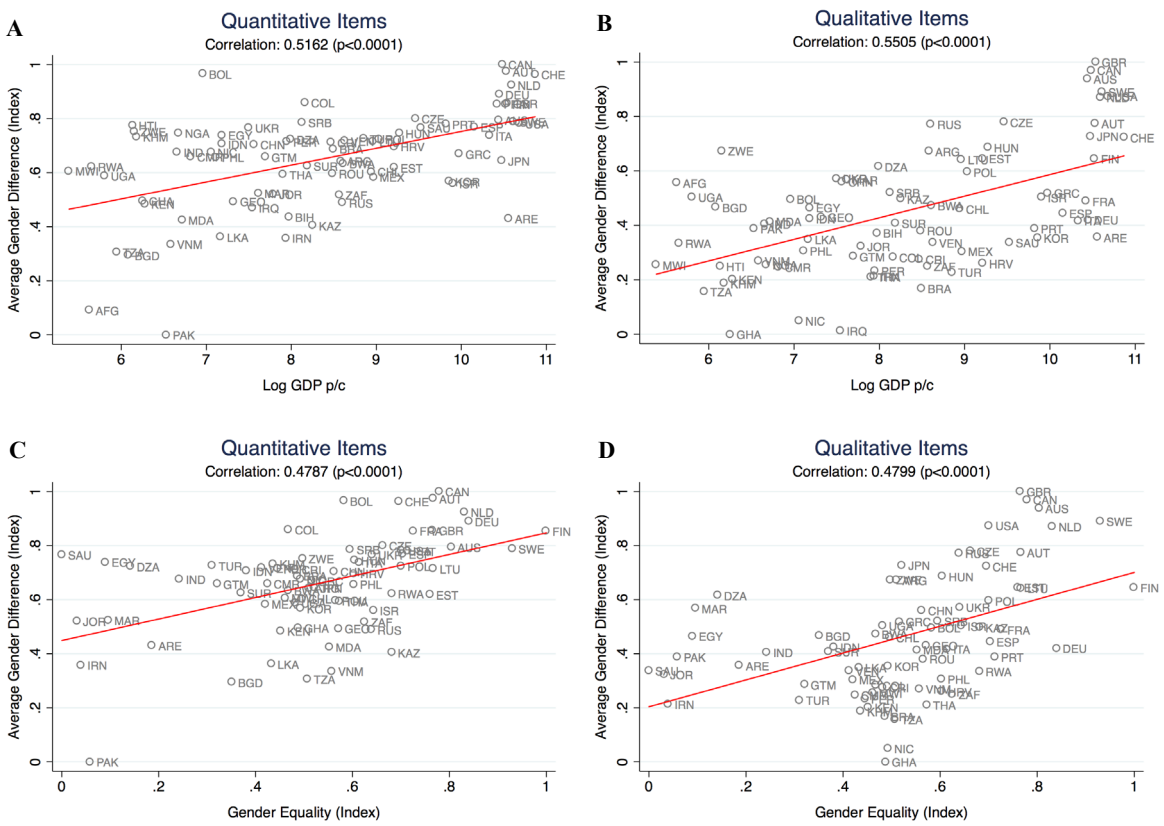


Fig. S7. Gender differences in responses to qualitative and quantitative items in relation to economic development and gender equality by country.

(A and B) The relationship between indices of gender differences in responses to quantitative and qualitative items and economic development, measured by log GDP p/c, including a linear fit. (C and D) The relationship between indices of gender differences in responses to quantitative and qualitative items and the Gender Equality Index including a linear fit. The indices of gender differences in quantitative and qualitative items were obtained as the predicted first main component from a principal component analysis of the country-level gender differences in the respective survey items.

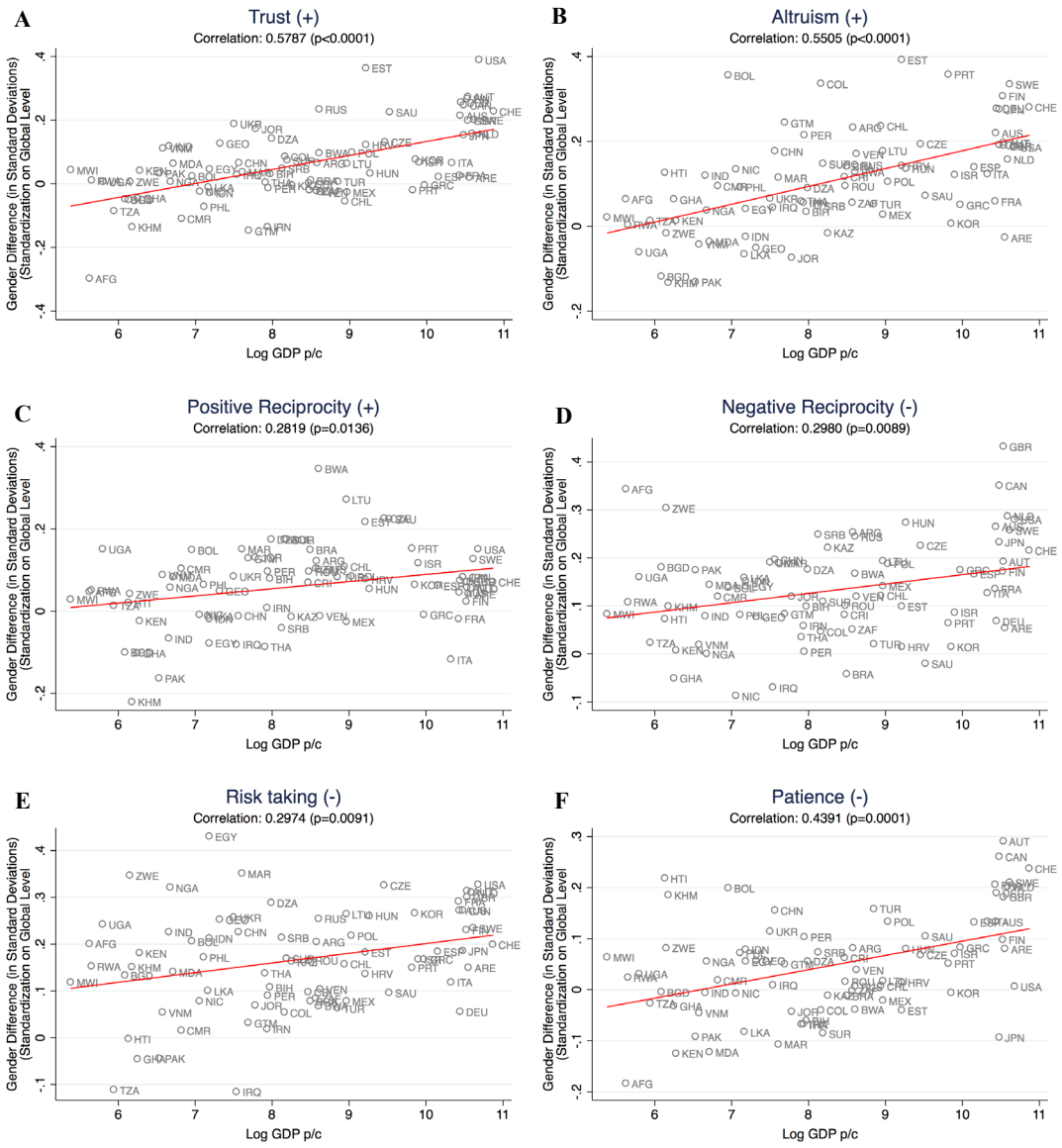


Fig. S8. Gender differences and economic development by preference and country using preferences standardized at the global level.

Each panel shows the relationship between country-level gender differences in a particular preference and the level of economic development measured by log GDP p/c, including a linear fit. For each preference, the symbols (+)/(-) in the panel titles indicate the direction of the difference. (+) indicates that women exhibited higher levels of the respective preference if the difference was positive. (-) indicates that men exhibited higher levels of the respective preference if the difference was positive. Preference measures were standardized at the global instead of the country level. For each preference and country, the gender difference was then calculated as the coefficient on a gender indicator with male as the reference category in an OLS regression of the respective preference on the gender indicator, controlling for age, age squared, subjective math skills, education level, household income quintile for the particular country sample.

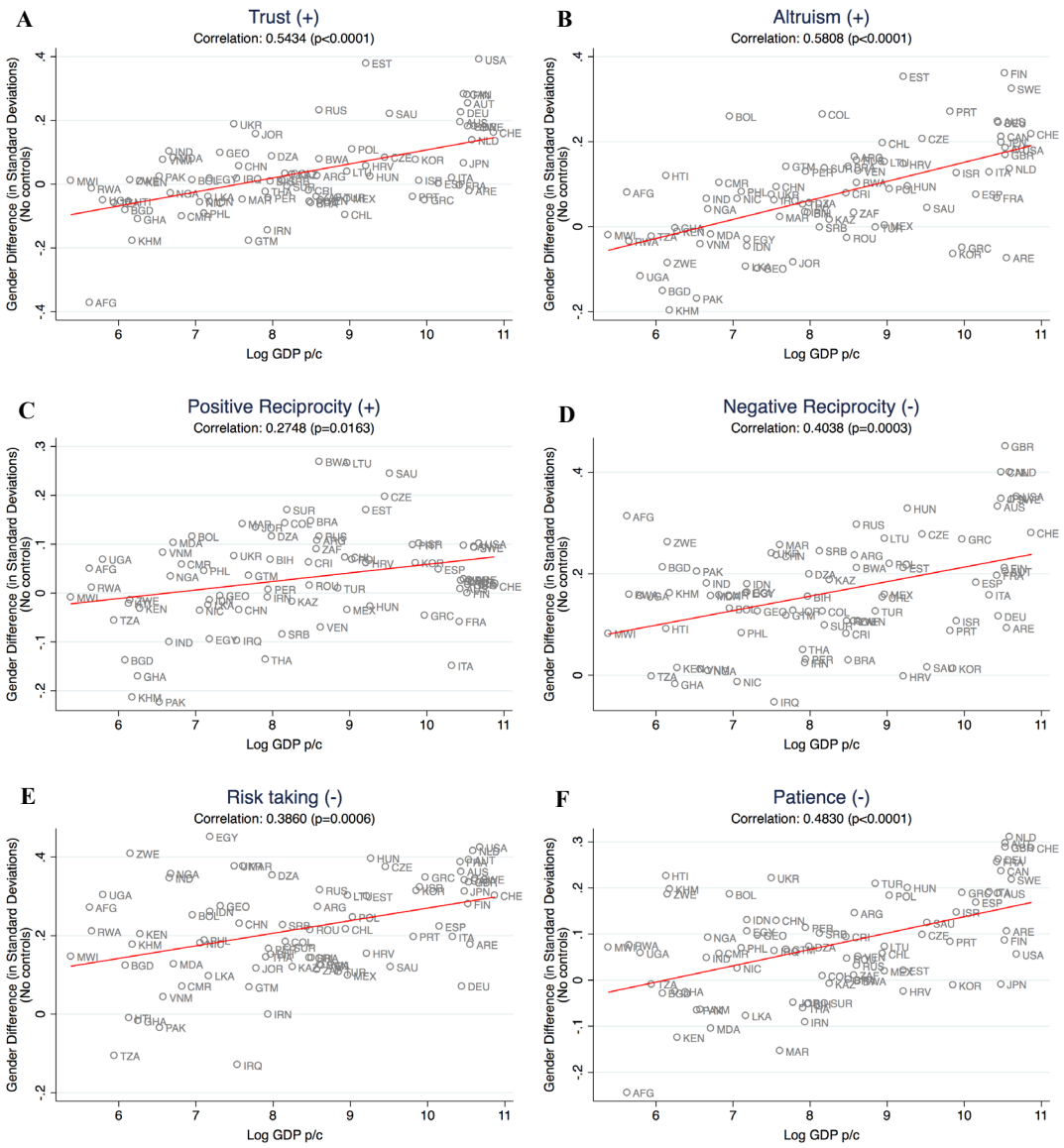


Fig. S9. Gender differences and economic development by preference and country without controls.

Each panel shows the relationship between country-level gender differences in a particular preference and the level of economic development measured by log GDP p/c, including a linear fit. For each preference, the symbols (+)/(-) in the panel titles indicate the direction of the difference. (+) indicates that women exhibited higher levels of the respective preference if the difference was positive. (-) indicates that men exhibited higher levels of the respective preference if the difference was positive. For each preference and country, the gender difference was calculated as the coefficient on a gender indicator with male as the reference category in an OLS regression of the respective preference on the gender indicator without controls for the particular country sample.

	Altruism	Trust	Pos. Recip.	Neg. Recip.	Risk taking	Patience
Conditional	0.106***	0.064***	0.055***	-0.129***	-0.168***	-0.050***
	(0.013)	(0.015)	(0.011)	(0.012)	(0.013)	(0.012)
Unconditional	0.066***	0.030*	0.042***	-0.170***	-0.227***	-0.077***
	(0.015)	(0.017)	(0.013)	(0.013)	(0.016)	(0.017)

Table S1. Global gender differences in preferences conditional on controls and unconditional.

Positive values indicate that women exhibited higher levels of the respective preference, negative values indicate that women exhibited lower levels of the respective preference. Gender differences were calculated as coefficients on a gender indicator with male as the reference category in an OLS regression of the respective preference on the worldwide sample. Conditional gender differences were calculated using as controls age, age squared, subjective math skills, education level, household income quintile, and country fixed effects. Unconditional gender differences were calculated without controls.

	(1)	(2)	(3)	(4)	(5)
	Average Gender Difference (Index)				
Log GDP p/c	0.668***	0.620***	0.567***	0.704***	0.703***
	(0.091)	(0.125)	(0.154)	(0.124)	(0.187)
Geographic Ctrls.	No	Yes	No	No	Yes
Demographic and Cultural Ctrls.	No	No	Yes	No	Yes
Historical Ctrls.	No	No	No	Yes	Yes
Observations	76	74	73	75	72
R-squared	0.447	0.713	0.518	0.449	0.759

Standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S2. Gender differences in preferences and economic development.

Country-level regressions of the index of gender differences in preferences on log GDP p/c and different sets of controls. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. Column (1) used no controls. Column (2) used geographic controls containing longitude, absolute latitude, log area, mean elevation, % living in (sub-)tropical zones, % arable land, land suitability for agriculture, mean precipitation, mean temperature, % at risk of malaria, predicted genetic diversity and its square, and continent fixed effects. Column (3) used demographic and cultural controls containing average age, ethnic fractionalization, linguistic fractionalization, religious fractionalization, and share of atheists. Column (4) used historical controls containing years of civil conflict 1800-2007, years of interstate conflict 1800-2007, and an indicator variable for colonization, with 1 indicating that the country was under colonial rule. Column (5) used all three sets of controls.

	(1)	(2)	(3)	(4)	(5)	(6)
	Gender Difference in					
	Altruism	Trust	Pos. Recip.	Neg. Recip.	Risk Taking	Patience
Female GNI p/c	0.123***	0.072**	0.021	0.088***	0.064**	0.040**
	(0.021)	(0.030)	(0.030)	(0.022)	(0.025)	(0.017)
Male GNI p/c	-0.055**	0.015	0.008	-0.042*	-0.016	0.004
	(0.021)	(0.031)	(0.033)	(0.024)	(0.024)	(0.016)
Observations	76	76	76	76	76	76
R-squared	0.403	0.461	0.077	0.241	0.193	0.187

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S3. Gender differences in preferences and gender-specific levels of economic development.

Country-level regressions of gender differences in preferences on standardized values of male and female Gross National Income p/c (by preference). As dependent variable, column (1) used the gender difference in altruism, column (2) used the gender difference in trust, column (3) used the gender difference in positive reciprocity, column (4) used the gender difference in negative reciprocity, column (5) used the gender difference in risk taking, column (6) used the gender difference in patience.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Average Gender Difference (Index)										
Log GDP p/c	0.668***						0.526***	0.596***	0.432**	0.661***	0.552***
	(0.091)						(0.101)	(0.095)	(0.165)	(0.085)	(0.093)
Gender Equality (Index)		0.556***					0.319***				
		(0.115)					(0.105)				
WEF Global Gender Gap Index			0.405***					0.233***			
			(0.104)					(0.086)			
UN Gender Equality Index				0.652***					0.291*		
				(0.085)					(0.147)		
Ratio Female to Male LFP					0.266**					0.245**	
					(0.121)					(0.096)	
Time since Women's Suffrage						0.514***					0.299***
						(0.135)					(0.095)
Observations	76	71	72	75	76	76	71	72	75	76	76
R-squared	0.447	0.311	0.168	0.420	0.071	0.264	0.528	0.494	0.475	0.507	0.523

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S4. Gender differences in preferences, gender equality, and economic development.

Country-level regressions of the index of gender differences in preferences on measures of gender equality and log GDP p/c. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. As independent variable, column (1) used log GDP p/c, column (2) used the Gender Equality Index, column (3) used the WEF Global Gender Gap Index, column (4) used the UN Gender Equality Index, column (5) used the ratio of female to male labor force participation rates, column (6) used the time since women's suffrage. Columns (7) to (11) were analogous to columns (2) to (6) but additionally used log GDP p/c as an independent variable.

	(1)	(2)	(3)	(4)	(5)
	Average Gender Difference (Index without Trust)				
Log GDP p/c	0.613***	0.573***	0.534***	0.648***	0.749***
	(0.097)	(0.152)	(0.160)	(0.130)	(0.205)
Geographic Ctrls.	No	Yes	No	No	Yes
Demographic and Cultural Ctrls.	No	No	Yes	No	Yes
Historical Ctrls.	No	No	No	Yes	Yes
Observations	76	74	73	75	72
R-squared	0.376	0.642	0.430	0.383	0.710

Standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S5. Gender differences in preferences and economic development excluding trust.

Country-level regressions of an alternative index of gender differences in preferences excluding trust on log GDP p/c and different sets of controls. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. Column (1) used no controls. Column (2) used geographic controls containing longitude, absolute latitude, log area, mean elevation, % living in (sub-)tropical zones, % arable land, land suitability for agriculture, mean precipitation, mean temperature, % at risk of malaria, predicted genetic diversity and its square, and continent fixed effects. Column (3) used demographic and cultural controls containing average age, ethnic fractionalization, linguistic fractionalization, religious fractionalization, and share of atheists. Column (4) used historical controls containing years of civil conflict 1800-2007, years of interstate conflict 1800-2007, and an indicator variable for colonization, with 1 indicating that the country was under colonial rule. Column (5) used all three sets of controls.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Average Gender Difference (Index without Trust)										
Log GDP p/c	0.613***						0.481***	0.554***	0.409**	0.605***	0.485***
	(0.097)						(0.104)	(0.101)	(0.169)	(0.093)	(0.095)
Gender Equality (Index)		0.552***					0.336***				
		(0.106)					(0.102)				
WEF Global Gender Gap Index			0.405***					0.244**			
			(0.104)					(0.093)			
UN Gender Equality Index				0.596***					0.254		
				(0.095)					(0.156)		
Ratio Female to Male LFP					0.272**					0.253**	
					(0.125)					(0.110)	
Time since Women's Suffrage						0.518***					0.329***
						(0.117)					(0.082)
Observations	76	71	72	75	76	76	71	72	75	76	76
R-squared	0.376	0.310	0.168	0.350	0.074	0.268	0.491	0.452	0.399	0.440	0.468

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S6. Gender differences in preferences, gender equality, and economic development excluding trust.

Country-level regressions of an alternative index of gender differences in preferences excluding trust on measures of gender equality and log GDP p/c. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. As independent variable, column (1) used log GDP p/c, column (2) used the Gender Equality Index, column (3) used the WEF Global Gender Gap Index, column (4) used the UN Gender Equality Index, column (5) used the ratio of female to male labor force participation rates, column (6) used the time since women's suffrage. Columns (7) to (11) were analogous to columns (2) to (6) but additionally used log GDP p/c as an independent variable.

	(1)	(2)	(3)	(4)	(5)
	Average Gender Difference (Index using Preferences Standardized at Global Level)				
Log GDP p/c	0.642***	0.547***	0.521***	0.682***	0.648***
	(0.091)	(0.129)	(0.152)	(0.124)	(0.194)
Geographic Ctrls.	No	Yes	No	No	Yes
Demographic and Cultural Ctrls.	No	No	Yes	No	Yes
Historical Ctrls.	No	No	No	Yes	Yes
Observations	76	74	73	75	72
R-squared	0.413	0.688	0.490	0.418	0.741

Standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S7. Gender differences in preferences and economic development using preferences standardized at the global level.

Country-level regressions of an alternative index of gender differences in preferences on log GDP p/c and different sets of controls. The index was constructed in a parallel way to the main index but used preferences standardized at the global (instead of country) level. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. Column (1) used no controls. Column (2) used geographic controls containing longitude, absolute latitude, log area, mean elevation, % living in (sub-)tropical zones, % arable land, land suitability for agriculture, mean precipitation, mean temperature, % at risk of malaria, predicted genetic diversity and its square, and continent fixed effects. Column (3) used demographic and cultural controls containing average age, ethnic fractionalization, linguistic fractionalization, religious fractionalization, and share of atheists. Column (4) used historical controls containing years of civil conflict 1800-2007, years of interstate conflict 1800-2007, and an indicator variable for colonization, with 1 indicating that the country was under colonial rule. Column (5) used all three sets of controls.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Average Gender Difference (Index using Preferences Standardized at Global Level)										
Log GDP p/c	0.642***						0.503***	0.574***	0.393**	0.635***	0.526***
	(0.091)						(0.102)	(0.096)	(0.166)	(0.086)	(0.094)
Gender Equality (Index)		0.540***					0.314***				
		(0.115)					(0.107)				
WEF Global Gender Gap Index			0.392***					0.226**			
			(0.107)					(0.091)			
UN Gender Equality Index				0.634***					0.305**		
				(0.087)					(0.152)		
Ratio Female to Male LFP					0.253**					0.233**	
					(0.122)					(0.101)	
Time since Women's Suffrage						0.503***					0.298***
						(0.128)					(0.091)
Observations	76	71	72	75	76	76	71	72	75	76	76
R-squared	0.413	0.293	0.156	0.397	0.064	0.253	0.491	0.458	0.443	0.467	0.488

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S8. Gender differences in preferences, economic development and gender equality using preferences standardized at the global level.

Country-level regressions of an alternative index of gender differences in preferences on log GDP p/c and measures of gender equality. The index was constructed in a parallel way to the main index but used preferences standardized at the global (instead of country) level. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. As independent variable, column (1) used log GDP p/c, column (2) used the Gender Equality Index, column (3) used the WEF Global Gender Gap Index, column (4) used the UN Gender Equality Index, column (5) used the ratio of female to male labor force participation rates, column (6) used the time since women's suffrage. Columns (7) to (11) were analogous to columns (2) to (6) but additionally used log GDP p/c as an independent variable.

	(1)	(2)	(3)	(4)	(5)
	Average Gender Difference (Index Using no Controls)				
Log GDP p/c	0.669***	0.549***	0.544***	0.698***	0.653***
	(0.087)	(0.124)	(0.159)	(0.122)	(0.187)
Geographic Ctrls.	No	Yes	No	No	Yes
Demographic and Cultural Ctrls.	No	No	Yes	No	Yes
Historical Ctrls.	No	No	No	Yes	Yes
Observations	76	74	73	75	72
R-squared	0.447	0.712	0.536	0.451	0.751

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S9. Gender differences in preferences and economic development using index without controls.

Country-level regressions of an alternative index of gender differences in preferences on log GDP p/c and different sets of controls. The index was constructed in a parallel way to the main index but country-level gender differences were calculated without using controls. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. Column (1) used no controls. Column (2) used geographic controls containing longitude, absolute latitude, log area, mean elevation, % living in (sub-)tropical zones, % arable land, land suitability for agriculture, mean precipitation, mean temperature, % at risk of malaria, predicted genetic diversity and its square, and continent fixed effects. Column (3) used demographic and cultural controls containing average age, ethnic fractionalization, linguistic fractionalization, religious fractionalization, and share of atheists. Column (4) used historical controls containing years of civil conflict 1800-2007, years of interstate conflict 1800-2007, and an indicator variable for colonization, with 1 indicating that the country was under colonial rule. Column (5) used all three sets of controls.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Average Gender Difference (Index Using no Controls)										
Log GDP p/c	0.669***						0.520***	0.597***	0.409**	0.662***	0.545***
	(0.087)						(0.095)	(0.090)	(0.162)	(0.082)	(0.089)
Gender Equality (Index)		0.574***					0.340***				
		(0.111)					(0.102)				
WEF Global Gender Gap Index			0.420***					0.247***			
			(0.099)					(0.082)			
UN Gender Equality Index				0.660***					0.318**		
				(0.084)					(0.149)		
Ratio Female to Male LFP					0.256**					0.235**	
					(0.116)					(0.091)	
Time since Women's Suffrage						0.530***					0.317***
						(0.130)					(0.090)
Observations	76	71	72	75	76	76	71	72	75	76	76
R-squared	0.447	0.326	0.177	0.430	0.065	0.281	0.534	0.497	0.480	0.503	0.533

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S10. Gender differences in preferences, gender equality, and economic development using index without controls.

Country-level regressions of an alternative index of gender differences in preferences on log GDP p/c and measures of gender equality. The index was constructed in a parallel way to the main index but country-level gender differences were calculated without using controls. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. As independent variable, column (1) used log GDP p/c, column (2) used the Gender Equality Index, column (3) used the WEF Global Gender Gap Index, column (4) used the UN Gender Equality Index, column (5) used the ratio of female to male labor force participation rates, column (6) used the time since women's suffrage. Columns (7) to (11) were analogous to columns (2) to (6) but additionally used log GDP p/c as an independent variable.

	(1)	(2)	(3)	(4)	(5)
	Average Gender Difference (Index)				
Log GDP p/c	0.584***	0.577***	0.472***	0.621***	0.663***
	(0.099)	(0.136)	(0.157)	(0.129)	(0.193)
Geographic Ctrls.	No	Yes	No	No	Yes
Demographic and Cultural Ctrls.	No	No	Yes	No	Yes
Historical Ctrls.	No	No	No	Yes	Yes
Ctrls. for Linguistic Distance to Germany	Yes	Yes	Yes	Yes	Yes
Observations	76	74	73	75	72
R-squared	0.489	0.721	0.549	0.496	0.772

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S11. Gender differences in preferences and economic development controlling for linguistic distance to Germany.

Country-level regressions of the index of gender differences in preferences on log GDP p/c and different sets of controls. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. All specifications used controls for linguistic distance to Germany containing the ASJP measure of linguistic distance to Germany as well as an indicator variable for German language, with 1 indicating that the country's major language is German. Column (1) used no additional controls. Column (2) additionally used geographic controls containing longitude, absolute latitude, log area, mean elevation, % living in (sub-)tropical zones, % arable land, land suitability for agriculture, mean precipitation, mean temperature, % at risk of malaria, predicted genetic diversity and its square, and continent fixed effects. Column (3) additionally used demographic and cultural controls containing average age, ethnic fractionalization, linguistic fractionalization, religious fractionalization, and share of atheists. Column (4) additionally used historical controls containing years of civil conflict 1800-2007, years of interstate conflict 1800-2007, and an indicator variable for colonization, with 1 indicating that the country was under colonial rule. Column (5) additionally used all three sets of further controls.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Average Gender Difference (Index)										
Log GDP p/c	0.584***						0.484***	0.547***	0.315*	0.597***	0.490***
	(0.099)						(0.111)	(0.102)	(0.165)	(0.094)	(0.096)
Gender Equality (Index)		0.443***					0.285***				
		(0.121)					(0.107)				
WEF Global Gender Gap Index			0.289**					0.197**			
			(0.113)					(0.088)			
UN Gender Equality Index				0.571***					0.323**		
				(0.092)					(0.144)		
Ratio Female to Male LFP					0.179					0.211**	
					(0.119)					(0.097)	
Time since Women's Suffrage						0.424***					0.271***
						(0.130)					(0.093)
Ctrls. for Linguistic Distance to Germany	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76.000	71.000	72.000	75.000	76.000	76.000	71.000	72.000	75.000	76.000	76.000
R-squared	0.489	0.380	0.277	0.492	0.238	0.372	0.544	0.517	0.519	0.532	0.550

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S12. Gender differences in preferences, gender equality, and economic development controlling for linguistic distance to Germany.

Country-level regressions of the index of gender differences in preferences on measures of gender equality and log GDP p/c. The dependent and independent variables were standardized to exhibit a mean of zero and a standard deviation of one. As independent variable, column (1) used log GDP p/c, column (2) used the Gender Equality Index, column (3) used the WEF Global Gender Gap Index, column (4) used the UN Gender Equality Index, column (5) used the ratio of female to male labor force participation rates, column (6) used the time since women's suffrage. Columns (7) to (11) were analogous to columns (2) to (6) but additionally used log GDP p/c as an independent variable. All specifications used additional controls for linguistic distance to Germany containing the ASJP measure of linguistic distance to Germany as well as an indicator variable for German language, with 1 indicating that the country's major language is German.

	(1)	(2)	(3)	(4)	(5)	(6)
	Trust	Altruism	Pos. Recip.	Neg. Recip.	Risk taking	Patience
1 if female	0.072***	0.110***	0.056***	-0.137***	-0.179***	-0.049***
	(0.014)	(0.012)	(0.012)	(0.012)	(0.013)	(0.011)
1 if female X Log [Household income p/c]	0.069***	0.060***	0.017*	-0.024**	-0.028**	-0.040***
	(0.012)	(0.010)	(0.009)	(0.011)	(0.012)	(0.011)
Log [Household income p/c]	-0.051***	0.021*	0.033***	0.038**	0.097***	0.068***
	(0.013)	(0.012)	(0.012)	(0.017)	(0.012)	(0.012)
Age	0.453**	-0.110	1.022***	-0.309*	-0.144	0.522***
	(0.218)	(0.147)	(0.191)	(0.186)	(0.200)	(0.168)
Age squared	-0.076	0.177	-1.132***	-0.529***	-1.138***	-1.207***
	(0.221)	(0.156)	(0.208)	(0.185)	(0.207)	(0.183)
Subj. math skills	0.062***	0.041***	0.036***	0.042***	0.042***	0.025***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.004)	(0.002)
Education level	-0.048***	0.076***	0.079***	-0.004	0.082***	0.091***
	(0.014)	(0.013)	(0.011)	(0.010)	(0.011)	(0.014)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	77072	77855	78086	76761	77673	77726
R-squared	0.030	0.022	0.018	0.040	0.081	0.027

Standard errors clustered at the country level in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S13. Gender differences in preferences and respondent-level income.

Individual-level regressions of preferences, standardized at the country-level, on a gender indicator with male as the reference category, log household income per capita, and their interaction controlling for age, age squared, subjective math skills, education level, and country-fixed effects. Log household income per capita was standardized to exhibit a mean of zero and standard deviation of one. As dependent variable, column (1) used trust, column (2) used altruism, column (3) used positive reciprocity, column (4) used negative reciprocity, column (5) used risk taking, column (6) used patience.

	Average Gender Difference (Index)			
	(1)	(2)	(3)	(4)
Log [GDP p/c PPP]	0.429***	-0.803	0.337***	-0.134
	(0.059)	(0.561)	(0.064)	(0.611)
Log [GDP p/c PPP] squared		0.074**		0.029
		(0.034)		(0.037)
Gender Equality (Index)			1.482***	1.374**
			(0.487)	(0.535)
F-statistic and p-value for F-test of zero impact of Log [GDP p/c PPP]	53.57 (p<0.0001)	29.56 (p<0.0001)	27.33 (p<0.0001)	13.31 (p<0.0001)
Observations	76	76	71	71
R-squared	0.447	0.475	0.528	0.531

Robust standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table S14. Gender differences in preferences, non-linear effects of economic development and gender equality.

Country-level regressions of the index of gender differences in preferences on log GDP p/c, log GDP p/c squared and the Gender Equality Index. As independent variables, column (1) used log GDP p/c, column (2) used log GDP p/c and log GDP p/c squared, column (3) used log GDP p/c and the Gender Equality Index, column (4) used log GDP p/c, log GDP p/c squared and the Gender Equality Index.