

Neyse, Levent; Johannesson, Magnus; Dreber, Anna

Working Paper

2D:4D does not predict economic preferences: Evidence from a large, representative sample

SOEPpapers on Multidisciplinary Panel Data Research, No. 1086

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Neyse, Levent; Johannesson, Magnus; Dreber, Anna (2020) : 2D:4D does not predict economic preferences: Evidence from a large, representative sample, SOEPpapers on Multidisciplinary Panel Data Research, No. 1086, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/219006>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

SOEPpapers

on Multidisciplinary Panel Data Research

2D:4D does not predict economic preferences: Evidence from a large, representative sample

Levent Neyse, Magnus Johannesson, Anna Dreber

SOEPPapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPPapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPPapers are available at <http://www.diw.de/soeppapers>

Editors:

Jan **Goebel** (Spatial Economics)
Stefan **Liebig** (Sociology)
David **Richter** (Psychology)
Carsten **Schröder** (Public Economics)
Jürgen **Schupp** (Sociology)
Sabine **Zinn** (Statistics)

Conchita **D'Ambrosio** (Public Economics, DIW Research Fellow)
Denis **Gerstorff** (Psychology, DIW Research Fellow)
Katharina **Wrohlich** (Gender Economics)
Martin **Kroh** (Political Science, Survey Methodology)
Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Fellow)
Thomas **Siedler** (Empirical Economics, DIW Research Fellow)
C. Katharina **Spieß** (Education and Family Economics)
Gert G. **Wagner** (Social Sciences)

ISSN: 1864-6689 (online)

German Socio-Economic Panel (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: soeppapers@diw.de



2D:4D does not predict economic preferences: Evidence from a large, representative sample

Levent Neyse¹, Magnus Johannesson², and Anna Dreber^{2,3,*}

¹SOEP at DIW, Berlin, Germany

²Stockholm School of Economics, Department of Economics, Stockholm, Sweden

³University of Innsbruck, Department of Economics, Innsbruck, Austria

*Corresponding Author: Stockholm School of Economics, Department of Economics, Box 6501

SE-113 83 Stockholm, Sweden. E-mail: Anna.Dreber@hhs.se

Abstract

The digit ratio (2D:4D) is considered a proxy for testosterone exposure in utero, and there has been a recent surge of studies testing whether 2D:4D is associated with economic preferences. Although the results are not conclusive, previous studies have reported statistically significant correlations between 2D:4D and risk taking, altruism, positive reciprocity, negative reciprocity and trust. Many “researcher degrees of freedom” and small sample sizes are important limitations of previous studies. We present results from a pre-registered large sample study testing if 2D:4D is associated with economic preferences. Data were collected in a representative sample of adults in the German Socioeconomic Panel-Innovation Sample (SOEP-IS), in a sample of about 3,450 respondents (about 5 times larger than the previously largest study in this field). We find no statistically significant association between 2D:4D and economic preferences in the largest study to this date on the topic.

Keywords: Economic behavior, prenatal hormones, testosterone, digit ratio

JEL Codes: D03, D87

1 Introduction

There is substantial variation in economic preferences between individuals. Some of this variation has been linked to cultural or situational factors such as gender norms (e.g. Gneezy et al., 2009) or reference points (e.g. Thaler and Johnson, 1990). Relatively recently, the potential role of hormones has received substantial interest. Hormones have organizational effects on the brain during fetal development (Arnold, 2009; Lombardo et al., 2012; Phoenix et al., 1959), and these effects are often hypothesized to have long lasting effects on preferences and behaviors later in life. In particular testosterone, the sex steroid, has been hypothesized to be associated with a wide range of economic preferences and decisions.

While some studies have explored circulating levels of testosterone or administered testosterone (e.g. Bos et al., 2010; Eisenegger et al., 2011; Sapienza et al., 2009; Zethraeus et al., 2009), with mixed results (see Dreber and Johannesson, 2019 for further discussion), others have focused on prenatal testosterone exposure in utero. Such exposure is hypothesized to impact brain development of the fetus and could thus potentially explain some of the individual heterogeneity in preferences (Baron-Cohen, 2002). Since testosterone exposure in utero cannot be measured directly after birth, researchers have looked for proxies of this exposure. The ratio of the length of the 2nd digit to the length of the 4th digit (2D:4D) on each hand (Manning et al., 1998) has been used as such a proxy.

2D:4D has been linked to a number of traits such as personality, sexual orientation and various cognitive abilities, but the results are often contradictory and there is also mixed evidence on the role of publication bias (see, e.g. Grimbos et al., 2010; Hönekopp and Schuster, 2010; Hönekopp and Watson, 2011; Puts et al., 2008; Voracek and Loibl, 2009; Voracek et al., 2011). Recently, there has been a surge in the number of papers testing for an association between 2D:4D and economic preferences (e.g. Coates et al., 2009; Van Honk et al., 2012). Just on risk preferences alone, we count 18 papers. Such relationships suggest that biological factors may partially affect economic activities through decisions. Several studies also investigated direct associations with real economic outcomes. For example, Coates et al., 2009 reported an association between 2D:4D and the profitability

of financial traders and Nicolaou et al., 2017 reported an association between 2D:4D and self-employment. However, most of the papers in this literature are subject to many “researcher degrees of freedom” (Gelman & Loken, 2013; Simmons et al., 2011) and suffer from small sample sizes.

We contribute to this literature by testing the associations between 2D:4D and five economic preference measures in a large, representative sample. To do so, we integrated a 2D:4D and Economic Preferences module in the German Socioeconomic Panel-Innovation Sample (SOEP-IS) (Richter & Schupp, 2015). Our sample contains 3,433 adult respondents, which is about 5 times larger than the previously largest study in this field (Brañas-Garza et al., 2018). Our five measures of economic preferences are i) risk taking, ii) altruism, iii) positive reciprocity, iv) negative reciprocity and v) trust. The questions used to measure these economic preferences were adapted from the preference survey module of the Global Preference Survey (Falk et al., 2018) (see the Methods section for further details about the preference measures and Supplementary Information Text for the complete list of items). We chose to focus on these five economic preference measures because of previous attempts to link these measures to 2D:4D, although a large proportion of this literature has focused on risk taking and altruism. All our tests are pre-registered to avoid researcher degrees of freedom affecting the results.

Our hypotheses about the direction of the association between 2D:4D and each economic preferences are based on the previous literature. For risk taking, the hypothesis is a negative relation to 2D:4D, which would indicate a positive association between prenatal testosterone exposure and risk taking. However, even though previous literature has also had this hypothesis, the results in this literature are far from conclusive (see Parslow et al., 2019 for a review). For example, the first paper on this topic reports a negative correlation between risk taking and left hand 2D:4D but finds no statistically significant correlation with the right hand 2D:4D, and the significant correlation can only be found in one sample and not another (Dreber & Hoffman, 2007). There are other examples of statistically significant negative correlations (e.g. Barel, 2019; Bönthe et al., 2016; Brañas-Garza et al., 2018; Brañas-Garza and Rustichini, 2011; Garbarino et al., 2011; Stenstrom

et al., 2011; Sytsma, 2014), though many of these only find statistically significant correlations in a subset of the population studied. There are thus also several null results (e.g. Alonso et al., 2018; Apicella et al., 2008; Aycinena et al., 2014; Barel, 2019; Bönthe et al., 2016; Brañas-Garza et al., 2018; Chicaiza-Becerra and Garcia-Molina, 2017; Drichoutis and Nayga, 2015; Lima de Miranda et al., 2018; Neyse et al., 2020; Sapienza et al., 2009; Schipper, 2014; Stenstrom et al., 2011; Sytsma, 2014) and also one finding of a statistically significant positive correlation between 2D:4D and risk taking in a female subsample in one of the papers that found a statistically significant negative correlation in another subsample (Brañas-Garza & Rustichini, 2011). Based on the previous literature we test the hypothesis that there is a statistically significant negative association between 2D:4D and risk taking.

There are also several studies testing the association between 2D:4D and altruism, measured as giving in the dictator game (Brañas-Garza et al., 2019; Brañas-Garza et al., 2013; Buser, 2012; Galizzi & Nieboer, 2015; Millet & Dewitte, 2010). Although these previous results have been mixed, the main hypothesis in the previous literature has been a positive association between 2D:4D and altruism - this is also the hypothesis that we test. In the exploratory analyses we also test for a non-linear relationship between 2D:4D and altruism by also including the squared 2D:4D as done in some of the previous studies (e.g. Brañas-Garza et al., 2013; Galizzi and Nieboer, 2015).

When it comes to 2D:4D, positive reciprocity, negative reciprocity and trust, there are only a handful of earlier studies. Buser (2012), who explored the relation between 2D:4D and altruism, also studies the trust game and finds that low 2D:4D individuals send back statistically significantly less as the second player in the trust game, and thus are less reciprocal. This was the only previous study we were aware of testing for an association between 2D:4D and positive reciprocity before we wrote our pre-analysis plan.¹ Based on this study we test the hypothesis that there is a statistically significant positive association between 2D:4D and positive reciprocity. Responder behavior in the ultimatum

¹The recent paper by van Leeuwen et al. (2020), which is more thoroughly discussed in the Discussion section, also studies 2D:4D and among other things the economic behaviors trust, positive reciprocity and negative reciprocity and find no evidence for associations between these variables and 2D:4D.

game can be interpreted as a measure of negative reciprocity. Buser (2012) also tests for an association between 2D:4D and responder behavior in the ultimatum game, but finds no statistically significant association. However, another study finds a statistically significant negative correlation between right hand 2D:4D and the minimum acceptance level in the ultimatum game (a higher minimum acceptance level suggest stronger negative reciprocity) (Van den Bergh & Dewitte, 2006). These are the only previous studies we are aware of testing for an association between 2D:4D and negative reciprocity. Based on the study with the statistically significant negative correlation (Van den Bergh & Dewitte, 2006) we test the hypothesis that there is a statistically significant negative association between 2D:4D and negative reciprocity. Finally, there is only one previous study testing the relation between 2D:4D and trust. Buser, 2012 finds a statistically significant positive association between 2D:4D and trust in the first stage of the trust game. Based on this study we test the hypothesis that there is a statistically significant positive association between 2D:4D and trust.

The aforementioned studies are typically small and with many researcher degrees of freedom in the analysis (Gelman & Loken, 2013; Simmons et al., 2011) including scope for p-hacking and forking which leads to results with p-values that cannot be interpreted as false positive probabilities. The average total sample size for the 2D:4D studies discussed in Parslow et al. (2019) is $N=262$, with a range from 86 to 704 participants. However, most of these studies look at smaller subsamples based on gender or ethnicity, thus effective sample sizes tend to be substantially smaller. There is thus a need for a much larger study with a pre-registered analysis plan to rigorously test if 2D:4D is associated with economic preferences.

For our primary hypotheses tests, we focus on the right hand 2D:4D because it has been argued that the right hand 2D:4D is more sexually dimorphic than the left hand (Hönekopp & Watson, 2011). We test if the right hand 2D:4D is associated with the five measures of economic preferences in five regression analyses controlling for gender. We use the left hand 2D:4D and the average of the two hands in our robustness tests. Since previous literature sometimes reports finding correlations in one gender and not the other,

we perform a pre-registered exploratory analysis where we interact 2D:4D with gender to test if the associations differ between men and women. As some previous studies also test for a non-linear relationship between 2D:4D and economic preferences, we also carry out a pre-registered exploratory analysis adding 2D:4D squared to the regression analysis.

2 Methods

2.1 Data collection and sample

The study was pre-registered prior to starting the data collection (<https://osf.io/5vpdn/>). The pre-registration included the data, collection, measurement procedures and all statistical tests described below (including the exploratory tests and the robustness tests with the exception of the robustness tests using the corrected sample and the restricted sample, that were not pre-registered). The pre-registration did not include the instructions to the interviewers, but they are now posted on the OSF site. When we make any deviation from the pre-registration this is explicitly mentioned in the text. The data collection of the study was performed between September 2018 and December 2018 in the German Socioeconomic Panel-Innovation Sample (Goebel et al., 2019) and we received the data in May 2019. SOEP is a longitudinal survey study that has been running since 1984 and it has nearly 30,000 participants today. SOEP-IS, on the other hand, was established in 2011 (Richter & Schupp, 2015) and it is open for experimental and survey module suggestions from researchers every year. The survey committee decides on which of those suggested modules to include in the next wave. The selected modules are subsequently integrated to different sub-samples of SOEP-IS. The right to use the collected data belongs to applicant researchers for a year. After the one year embargo period, the data become available for the whole research community. The merged dataset also includes longitudinal socio-economic data, as well as previous years' modules. If a set of individuals participated in Modules A & B in 2014 and C & D in 2016, all their answers to these four modules can thus be merged into a single dataset. According to the 2018 release of SOEP-IS, it has a total number of 5722 participants from 3232 households.

We planned to survey the whole sample of SOEP-IS for the 2D:4D collection module. While the number of participants was predicted around 4,500 adults prior to the survey, the actual number was 4,860. As 2D:4D measurement was voluntary, the participants could skip the measurement and continue responding to the survey questions. As a result, our working sample consisted 3,482 participants with complete right or left hand measures (3,433 individuals with right, 3,454 with left, 3,405 with both). On top of 2D:4D measurements the participants were also asked to reply to a set of survey questions about their economic and social preferences. The gender distribution of the working sample is balanced (53.8% women) and the mean age is 54.22 (SD=18.39).

2.2 Economic preference measures

The 2D:4D module included a total number of 8 survey questions to elicit risk taking, altruism, positive reciprocity, negative reciprocity and trust (see Supplementary Information for the phrasing of these 8 survey questions). All questions were adapted from the preference survey module of the seminal paper of Falk et al. (2018), who followed both experimental tests and survey optimization procedures to construct their module (see also Falk et al. (2018) for a detailed description of the preference module construction procedure).

The risk question is the standard risk question that has been included in the past waves of SOEP as well, where participants are asked “How do you rate yourself personally? In general, are you someone who is ready to take risks or do you try to avoid risks (risk-averse)?”. For this question, participants are asked to indicate their willingness to take risks on a Likert scale from 0 to 10 where a higher number indicates more risk taking. The Global Preference Survey involves almost an identical self reported question (“In general how willing are you to take risks?”) and also a staircase procedure with hypothetical lotteries. We chose to use the standard self-reported question which was also experimentally validated in Dohmen et al. (2011).

We measured altruism with two questions and negative reciprocity with three questions (see Supplementary Information for more details). The multiple survey questions were

combined to a measure of altruism and negative reciprocity in the same way as done in the Global Preference Survey (Falk et al., 2018). They standardized responses on each survey question to z-scores and estimated weights for aggregating responses into the overall preference measure. We used the same approach and the same weights to aggregate survey questions into preference measures (see Supplementary Information for these weights). However, the index resulting from aggregating the z-scores for the individual questions will not in itself be a z-index (as the STD of this index will not be equal to 1). We therefore also standardized the index for altruism and the index for negative reciprocity into z-scores, to ease the interpretation of results. This standardization of the indices was not done in the previous study (Falk et al., 2018) using these two indices and it was not mentioned in the pre-analysis plan (but it does not affect the results per se; but only the units used to report results in). For comparability in results between our five preference measures we standardized responses to z-scores also for risk taking, positive reciprocity, and trust for the regression analyses testing the hypotheses.

To measure positive reciprocity, we integrated the self-assessment question (“When someone does me a favor, I am willing to return it.”) in the global preference module but did not integrate the hypothetical choice item in the data-collection. The respondents were asked whether they agreed with this expression on a 10 point Likert Scale. Finally, trust was also measured by a Likert Scale from 0 to 10, where a higher number is more agreement. The trust item we used was “People are basically honest”. Note that this phrasing of the trust question is slightly different than the one we proposed to include in the data collection and that we included in the pre-analysis plan (“As long as I am not convinced otherwise, I assume that people have only the best intentions”). As the SOEP-IS questionnaire already had the former item included in another module and also three other trust related items, the survey management decided to include the shorter version rather than the version we proposed. As the data for the module using the trust question were collected in only a sub-sample of SOEP-IS, we have a smaller number of observations for trust than planned and a smaller number of observations than for the other preference measures.

The distribution for each of the five economic preference measures are presented in the Supplementary Information (see Figures S6-S10).

2.3 2D:4D measurements

Both left and right hand 2D:4D's of the participants were measured during the household surveys with the help of digital calipers. Our main reasons for choosing a direct measurement method, instead of using scanners or mobile applications, were because of confidentiality of personal information as well as mobility and time efficiency. Scanned images of the hands would contain fingerprints of the participants, which may be considered as delicate personal information. Any discomfort or suspicion of the participants due to using scanned images may also result in dropping out of the study in further waves. Moreover, as the data was collected from a large number of households, digital calipers are easier to carry than flatbed scanners and they are also time and cost effective. While there are several mobile applications being developed for 2D:4D measurements, they are currently available only in demo versions.

263 interviewers were trained for the 2D:4D measurements. Their detailed hand measurement protocol has been posted as an Appendix to the pre-analysis plan (but it was not posted at the same time as the pre-analysis plan, but after the data collection was completed). This protocol was prepared following a set of tests on different types of digital calipers, measurement methods and multiple measurers. To test the protocol, two research assistants in SOEP measured a number of hands using the calipers and the proposed protocol. These measures collected by the research assistants were compared with digital measurements carried out by one of the authors of this study, LN. The measures obtained with the final protocol were almost identical to those obtained with digital scanners.

2.4 Statistical analysis

Following the recommendations by (Benjamin et al., 2018), we consider results with p-values below 0.005 to be statistically significant and results with p-values below 0.05 as

suggestive evidence. We conducted our main analysis with right hand 2D:4D and repeated the same analysis for left hand 2D:4D and the average of the two hands as robustness checks. A large body of the literature focuses on the right hand 2D:4D in their analyses based on the argument that this ratio shows greater sex differences (Hönekopp & Watson, 2011).

The five primary hypotheses were tested in OLS regressions estimated with robust (Huber-White) standard errors. The survey measure of economic preferences was the dependent variable and the right hand 2D:4D and the gender of the participant were the independent variables. The hypothesis tests tested if the coefficient of the right hand 2D:4D was significant in the hypothesized direction. For each hypothesis we included all participants with data on right hand 2D:4D and gender who responded to the survey question used to construct the preference measure tested in the hypothesis. For altruism and negative reciprocity, where the preference measures were based on combining answers to more than one survey question (see Supplementary Information), we included all participants who responded to all the survey questions used to construct the preference measure (thus we did not include potential participants who only answered to one question for these measures).

We included all participants with a recorded right hand 2D:4D measurement and we did not exclude any 2D:4D observations in our main results. However, as explained in more detail in the beginning of the Results section we added two not pre-registered robustness tests to test if our results are sensitive to outliers due to mismeasurement or injured fingers. These robustness tests are referred to as “the corrected sample” and “the restricted sample” and these results are reported in Supplementary Information Tables S2-S7, S11-S16, S20-S25.

2.4.1 Pre-registered exploratory analyses 1: Gender interactions

For the five primary hypotheses we tested if the association between right hand 2D:4D and the survey measures differed between men and women. These tests were done by adding an interaction between gender and the right hand 2D:4D in the five OLS regressions used to

test the five primary hypotheses. We tested if this interaction coefficient was statistically significant, but without specifying a hypothesized direction of the interaction coefficient as these analyses were exploratory and would need confirmation in other studies to carry much weight. We also reported the statistical significance of the right hand 2D:4D for men and women separately as part of these analyses, but these tests are only relevant if the interaction coefficient is statistically significant.

2.4.2 Pre-registered exploratory analyses 2: Non-linear relationships

Some studies in the literature have reported non-monotonic findings between 2D:4D and economic preferences, by adding a squared 2D:4D term to a linear regression (see for example Brañas-Garza et al., 2013; Galizzi and Nieboer, 2015). As exploratory analyses we therefore added the squared right hand 2D:4D to the OLS regression equations used for testing the five primary hypotheses, to test for non-linear effects of 2D:4D. In these tests we assessed the significance of the squared right hand 2D:4D, but without specifying a hypothesized direction of the squared term as these analyses are exploratory.

2.4.3 Pre-registered robustness tests: Left hand 2D:4D and average of left and right hand 2D:4D

As a robustness test we also estimated all the analyses above using the left hand 2D:4D and the average of the left hand and the right hand 2D:4D instead of the right hand 2D:4D (in the existing literature the right hand 2D:4D, the left hand 2D:4D and the average of the two have been used). We included all participants with data on left hand 2D:4D in the analyses for left hand 2D:4D in that robustness test (and with data on gender and the preference measure tested in each hypothesis as above) and all participants with data on both left and right hand 2D:4D in that robustness test (and data on gender and the preference measure tested in each hypothesis as above).

3 Results

In this section, we report our results in accordance with our pre-registered pre-analysis plan. In the pre-analysis plan, we did not include any exclusion criteria for excluding outliers (we wrote that all observations would be included). But the instructions to the interviewers for measuring 2D:4D in SOEP included some steps to prevent outliers due to mismeasurement or injured fingers. The interviewers were told that the 2D:4D typically lies between 0.8 and 1.1, and that if the measured ratio significantly exceeded this range they were told to repeat the measurement. For some individuals there therefore exists two recorded measurements of the digit ratio (one first measurement and a second measurement due to a suspected first mismeasurement). As we did not write anything about replacing the first measurement with the second measurement in these cases in the pre-analysis plan, we report results based on the first measurement for all individuals as our main results below. We supplement these results with an additional not pre-registered robustness test where we replace the first measurement with the second measurement for all individuals where a second measurement is available in the data (a second measurement of the right/left hand digit ratio was available for 90/149 individuals in the data). We carry out this robustness test for the primary hypotheses, the pre-registered exploratory analyses and the pre-registered robustness tests reported below and these results are reported in Tables S2-S4, S11-S13, and S20-S22. We refer to this robustness test as “the corrected sample”.

If the interviewee had a missing or severely injured second digit (2D) or fourth digit (4D), the interviewers were instructed not to measure the injured hand and continue with the other hand. But from comments added by interviewers in the data set, it is evident that in some cases they measured the injured hand and then added a comment about an injured finger. In our main results below, we include all the measured digit ratios even if there was a comment about an injured finger in the data set (as that is most consistent with our pre-analysis plan). But we have added an additional not-pre-registered robustness test where we have excluded digit ratios outside the range of 0.8-1.2 (this range corresponds to about +/- four STDs away from the mean in our data), as it is very unusual to observe

values outside this range in previous studies (in the data set we observed 14/30 digit ratios outside this range for the right/left hand digit ratio). This will imply that digit ratios that are outliers due to injured fingers will be excluded from the analysis. We chose this approach instead of trying to identify every recorded measurement in the data set due to injured fingers as it is not always obvious from the interviewers' comments if fingers are severely injured or not. We carry out this robustness test for the primary hypotheses, the pre-registered exploratory analyses and the pre-registered robustness tests reported below and these results are reported in Tables S5-S7, S14-S16, and S23-S25. We refer to this robustness test as "the restricted sample". For our restricted sample mean 2D:4D analyses we generated the dependent variable, m2D:4D, by calculating the average of restricted right and restricted left hand 2D:4Ds.

When we below report our results for the pre-registered tests, we comment in the text if the findings differ importantly in the not pre-registered tests in "the corrected sample" or "the restricted sample" (and we refer to the results below as "the full sample"). Following the recommendations by a recent paper (Benjamin et al., 2018), we consider results with p-values below 0.005 to be statistically significant and results with p-values below 0.05 as suggestive evidence; and all statistical tests are based on two-sided tests (these p-value cut-offs were mentioned already in the pre-registration).

3.1 Descriptives

Descriptives of the dependent and independent variables are shown in Table 1 (and in Table S1 for the corrected sample and the restricted sample), with results reported separately for men and women (the results for the preference variables are based on the $n=3,482$ individuals with at least one of the right or left hand 2D:4D data in the full sample). We also report the p-value of the gender difference of each variable based on an independent samples t-test, but these tests were not part of our pre-registered hypothesis tests. As conjectured in the 2D:4D literature, 2D:4D tends to be lower for men than for women. This difference is statistically significant for the left hand digit ratio and the mean of the left hand and right hand digit ratio, but not for the right hand digit ratio

in the full sample. In the corrected sample and the restricted sample there is suggestive evidence ($p < 0.05$) for a lower digit ratio for men also for the right hand digit ratio.

In line with previous literature on gender differences in economic preferences (Croson & Gneezy, 2009), men are on average more risk taking than women and less altruistic. Using the same measures as we do, Falk et al. investigated the economic and social preferences in 76 countries and with almost 80,000 participants. They found systematic gender differences in positive reciprocity, negative reciprocity and trust, with women on average being more likely to positively reciprocate, less likely to negatively reciprocate and more trusting (Falk et al., 2018). We confirm that finding for negative reciprocity ($p < 0.001$) and positive reciprocity ($p < 0.05$) whereas trust do not vary statistically significantly between men and women in our data.

Table 1: Descriptives by Gender (Full Sample)

	All	Men		Women	p-value	e-size
R2D:4D	1.000 (0.055)	0.999 (0.060)	<	1.002 (0.051)	0.151	0.049
L2D:4D	1.001 (0.068)	0.997 (0.062)	<	1.004 (0.073)	0.001	0.112
Mean 2D:4D	1.001 (0.049)	0.998 (0.047)	<	1.003 (0.050)	0.003	0.102
Risk	4.291 (2.320)	4.746 (2.332)	>	3.899 (2.238)	<0.001	0.371
Altruism	0.000 (0.769)	-0.094 (0.784)	<	0.081 (0.746)	<0.001	0.229
Positive Reciprocity	8.948 (1.558)	8.891 (1.550)	<	8.996 (1.564)	0.046	0.068
Negative Reciprocity	0.000 (0.768)	0.104 (0.791)	>	-0.090 (0.737)	<0.001	0.255
Trust	5.555 (1.989)	5.641 (1.987)	>	5.479 (1.989)	0.126	0.081

Note: Standard deviations are in parentheses. Reported tests are based on t-tests; effect sizes (*cohen's d*) are reported in e-size column. Reported values of Altruism and Neg. Reciprocity are the indices calculated from z-scores of the variables. All values are based on full sample without any exclusion criteria such as hand injuries or outliers.

3.2 Primary Hypotheses

The primary hypothesis tests, robustness tests, and exploratory analyses reported below were pre-registered prior to starting the data collection. In Table 2 we present the results for our primary hypotheses in regressions with the economic preference measure as a function of 2D:4D and gender. Note that in the regressions we have standardized all the five preference measures to z-scores so that one unit in the dependent variable represents a standard deviation of the preference measure. We cannot reject the null hypothesis of no association between 2D:4D and economic preferences for any of the five preference measures; and the sign of the coefficient is in the opposite direction of the hypothesis for three of the five hypotheses. The female coefficients confirm the gender differences in preferences observed in the descriptives data.

In Figure 1 we graphically illustrate our results for the five primary hypotheses, by showing 95% and 99.5% confidence intervals of the estimated regression coefficients for the association between 2D:4D and economic preferences. To ease the interpretation of effects we report these results as standardized regression coefficients so that the results show how many standard deviation units economic preferences change for a one standard deviation change in 2D:4D. For risk taking, altruism, positive reciprocity and negative reciprocity the upper and lower limit of the 99.5% confidence interval is within about a +/- 0.05 standard deviation change in economic preferences for a one standard deviation change in 2D:4D. The confidence intervals for trust are somewhat wider due to the lower sample size for trust (the trust question was only included in a sub-sample of SOEP-IS). These results thus show precisely estimated null results, suggesting that 2D:4D is not associated with important effects on economic preferences. These results are reinforced by our high power to find significant associations. Based on the observed standard errors we have 80% power to detect a standardized regression coefficient at the 0.005 (0.05) p-value level of 0.066 (0.051) for risk taking, 0.057 (0.044) for altruism, 0.060 (0.046) for positive reciprocity, 0.062 (0.048) for negative reciprocity, and 0.103 (0.079) for trust.

Table 2: Regression Analysis: Primary Hypotheses

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
R2D:4D	0.040 (0.326)	0.251 (0.282)	0.063 (0.294)	0.037 (0.307)	-0.234 (0.512)
female	-0.361** (0.034)	0.222** (0.035)	0.066 (0.034)	-0.247** (0.034)	-0.085 (0.054)
constant	0.157 (0.326)	-0.372 (0.283)	-0.099 (0.295)	0.095 (0.309)	0.273 (0.512)
<i>N</i>	3431	3350	3422	3378	1391
adj. R^2	0.032	0.012	0.001	0.015	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

3.3 Exploratory Analyses 1: Gender Interactions

We pre-registered two exploratory analyses. The first of these tests for a gender difference in the association between 2D:4D and economic preferences by adding an interaction between gender and 2D:4D in the regressions. In testing for the gender interaction coefficient we also report if 2D:4D is statistically significantly associated with economic preferences among men or women. The results, including the interaction effect tests, are reported in Table 3. The gender interaction is not statistically significant in any of the five regressions ($p > 0.05$). For negative reciprocity there is suggestive evidence ($p = 0.033$) of a positive gender interaction with a negative point estimate of the association between 2D:4D and positive reciprocity for men and a positive point estimate for women. However, these point estimates do not provide statistically significant or suggestive evidence for an association between 2D:4D and positive reciprocity among either men or women. The suggestive evidence for the gender interaction is thus likely to be due to randomness. This is reinforced by that there is no statistically significant or suggestive evidence of this gender interaction in the corrected sample or the restricted sample. In the restricted sample there is instead suggestive evidence of a positive gender interaction for risk taking with a negative point estimate of the association between 2D:4D and risk taking for men and a positive point estimate for women. However, also in this case the point estimate is

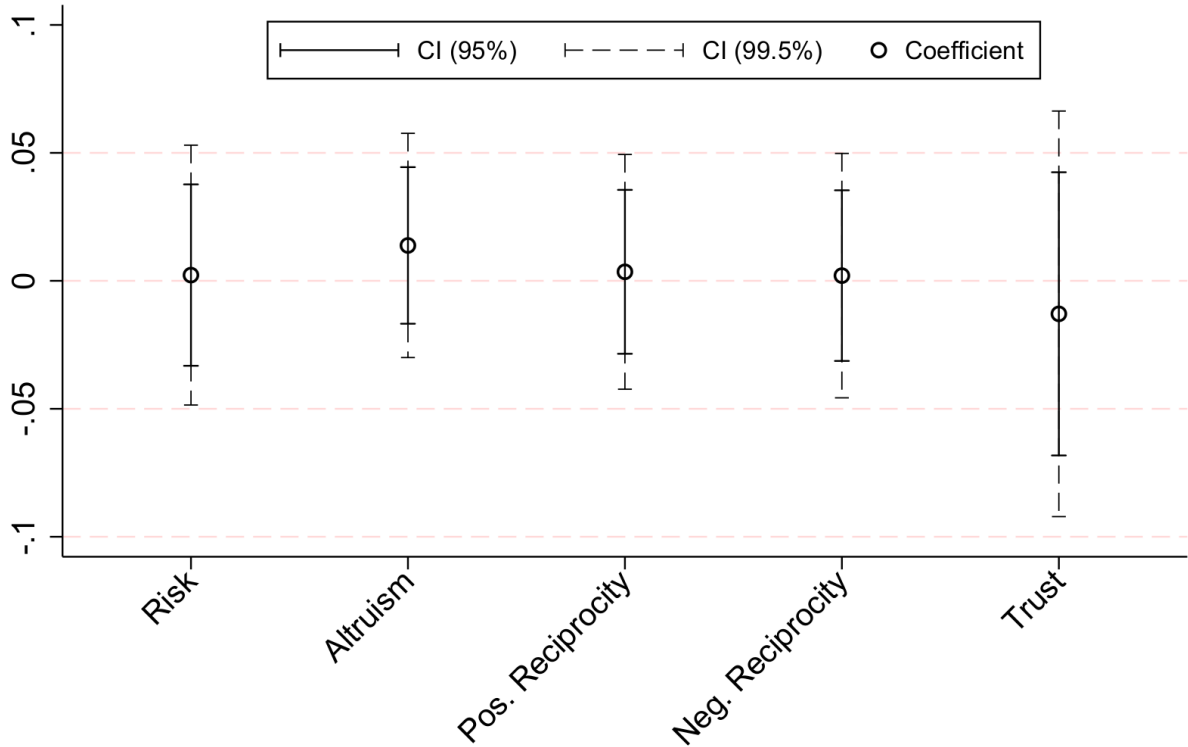


Figure 1: Standardized Regression Coefficients of Primary Hypotheses

not statistically significant for either men or women (and there is no suggestive evidence of an association either) (see Table S6).

3.4 Exploratory Analyses 2: Non-Linear Relationships

Our second exploratory analysis tests for a non-linear relationship between 2D:4D and economic preferences, by adding 2D:4D squared to the regression equation. Such non-linear relationships have been tested in several previous studies (e.g. Brañas-Garza et al., 2013; Galizzi and Nieboer, 2015; Lima de Miranda et al., 2018; Parslow et al., 2019). These results are reported in Table 4. The squared 2D:4D coefficient is not statistically significant in the five regression equations, with the exception of negative reciprocity. For negative reciprocity the squared 2D:4D term is statistically significant with a negative coefficient (and the 2D:4D coefficient has a positive coefficient). This result is also observed in the corrected sample, but not in the restricted sample suggesting that this result is due to outliers due to injured fingers (see Tables S4 and S7). In checking the data, there is one extreme right hand digit ratio of 2.175 in the full sample and in the corrected sample,

Table 3: Regression Analysis: Exploratory Analysis 1

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
R2D:4D	-0.551 (0.525)	0.128 (0.358)	-0.079 (0.410)	-0.540 (0.366)	0.622 (0.717)
female	-1.638* (0.700)	-0.043 (0.573)	-0.240 (0.600)	-1.489* (0.583)	1.661 (1.011)
R2D:4D X female	1.277 (0.699)	0.265 (0.572)	0.306 (0.598)	1.241* (0.581)	-1.749 (1.011)
constant	0.748 (0.524)	-0.250 (0.359)	0.043 (0.410)	0.671 (0.370)	-0.579 (0.716)
R2D:4D (F)	0.726 (0.463)	0.394 (0.446)	0.228 (0.436)	0.702 (0.451)	-1.128 (0.712)
N	3431	3350	3422	3378	1391
adj. R^2	0.033	0.012	0.000	0.016	0.002

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,
** $p < 0.005$. R2D:4D (F) presents the 2D:4D effects for women
which were calculated by summing the coefficients of R2D:4D and
R2D:4D X female variables.

but not in the restricted sample. According to the comments by the interviewers this is due to an injury of the 4D finger. If this outlier is removed the squared 2D:4D term is no longer statistically significant in the full sample (and there is no suggestive evidence either). We thus find no robust evidence of a non-linear relationship between 2D:4D and economic preferences. Without this extreme outlier the suggestive evidence for the 2D:4D gender interaction and negative reciprocity in Table 3 above also disappears .

3.5 Robustness Tests: Left Hand 2D:4D and Average of Right and Left Hand 2D:4D

As pre-registered robustness tests we estimated the primary hypotheses and exploratory analyses also for the left hand 2D:4D and the average of the left hand and right hand 2D:4D. For the full sample there is no statistically significant or suggestive evidence of an association between 2D:4D and any of the economic preferences in any of the robustness tests for the primary hypotheses or the exploratory analyses (see Tables S8-S10 and S17-

Table 4: Regression Analysis: Exploratory Analysis 2

	(1)	(2)	(3)	(4)	(5)
	Risk	Altruism	Positive Recip.	Negative Recip.	Trust
R2D:4D	-0.768 (1.233)	0.316 (0.982)	-1.023 (0.951)	2.293* (0.993)	1.467 (5.709)
female	-0.361** (0.034)	0.222** (0.035)	0.067 (0.034)	-0.249** (0.034)	-0.085 (0.054)
R2D:4D ²	0.370 (0.500)	-0.030 (0.346)	0.498 (0.340)	-1.032** (0.346)	-0.832 (2.756)
constant	0.594 (0.755)	-0.408 (0.644)	0.488 (0.624)	-1.125 (0.654)	-0.593 (2.965)
<i>N</i>	3431	3350	3422	3378	1391
adj. <i>R</i> ²	0.032	0.012	0.000	0.015	-0.000

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

S19).

We also carried out the analyses using the left hand 2D:4D and the average of the left hand and right and 2D:4D using the corrected sample and the restricted sample (see Tables S11-S16 and S20-S25). As noted above these additional robustness tests were not pre-registered. In these analyses there is no statistically significant or suggestive evidence in the robustness tests of the primary hypotheses or the exploratory tests of a gender interaction. But in three cases the squared 2D:4D coefficient is statistically significant, in the robustness tests of a non-linear relationship. For negative reciprocity the squared 2D:4D term for the average 2D:4D of both hands is statistically significant in the corrected sample with a negative coefficient (and the 2D:4D coefficient has a positive coefficient) (see Table S22). As above for right hand 2D:4D this result seem to be due to outliers as the squared term is not statistically significant in the restricted sample (see Table S25), and as above the result is not robust to removing the one extreme outlier due to an injured finger noted above (there is not even suggestive evidence for the squared term without this observation).

The other two cases with a statistically significant squared term is for risk taking in the restricted sample; here the squared term is statistically significant in both the

regression for left hand 2D:4D and the regression for the average of the left hand and right hand 2D:4D (see Tables S16 and S25). In both these regressions the squared term has a negative coefficient (and the 2D:4D has a positive coefficient). This implies that risk taking initially increases with 2D:4D until the function peaks, and then it decreases. For left hand 2D:4D the function peaks at 0.97 and for the average of the right hand and the left hand 2D:4D the function peaks at 0.98. The hypothesized sign of 2D:4D and risk taking is a negative sign, which implies that for a sizeable part of the 2D:4D distribution the direction of the association is in the opposite direction of this hypothesis in these models with the squared term (the lower part of the 2D:4D distribution until the function peaks). The results for these two regressions with the squared term are thus not fully consistent with the hypothesized negative association, and we think they are likely to be chance findings. These analyses are also part of the exploratory analyses, and would need confirmation in future confirmatory studies to provide substantive evidence of a non-linear association between 2D:4D and risk taking.

4 Discussion

Our null results suggest that variation in 2D:4D does not explain any important variation in economic preferences between individuals. The point estimates for the five measures of economic preferences are close to zero and are precisely estimated with 99.5% confidence intervals within about a ± 0.05 standard deviation change in economic preferences for a one standard deviation change in 2D:4D (with the exception of trust, which has somewhat wider confidence intervals due to the smaller sample size).

Why do we find null results while several smaller studies have found indications of 2D:4D correlating with economic preferences? This is probably due to the “usual suspects” increasing the risk of false positive results; small sample sizes, publication bias and researcher degrees of freedom. For example, in most previous papers there is room to try whether there is a statistically significant association between the studied behavior for the right hand, or for the left hand, or for the average of the two hands, or for men only, or for women only, or for the pooled sample of men and women, etc. Our sample

size was about 5 times larger than the previously largest study in this field, and the pre-registration of our hypothesis tests leaves little room for researcher degrees of freedom affecting our results.

Another interpretation of our findings is that testosterone exposure in utero is unimportant for economic preferences. Recent evidence by van Leeuwen et al. (2020) suggest that this may be the case – using direct measures of testosterone and estrogen from umbilical cord blood at birth they find no evidence for an association of these hormones and economic preferences in a sample of 217 adults. In this study the authors also find no evidence for associations between 2D:4D and economic preferences in a sample of 597 adults. Another possibility is that 2D:4D might not be a valid proxy for testosterone exposure in utero. The evidence of a link between 2D:4D and testosterone exposure, while coming from several strands of literature, is mainly indirect and the results are often inconsistent. First, it has been argued that 2D:4D is lower in men than in women. While this tends to be the case also in our study the difference is small, and as reported in ABC, this result is not always found (for example in a large sample of the Hadza population Apicella et al., 2015). Second, the most direct evidence comes from a study on the testosterone-to-estradiol ratio in amniotic fluid in a sample of 29 children (Lutchmaya et al., 2004). They report a statistically significant negative correlation between this ratio and right hand 2D:4D, even after controlling for gender, while the correlation for the left hand is reported as insignificant. Similarly, another study finds a weak negative correlation between the testosterone levels in the mother’s plasma and 2D:4D of newborns for both hands and both sexes (Ventura et al., 2013). Yet, in a considerably larger study (106 children) the result is not statistically significant. Third, based on the theory of sex hormone transfer in utero where the idea is that female fetuses receive transfers of testosterone from their male co-twin (Miller, 1994), some studies compare the 2D:4D of same sex and opposite sex twins. While one study finds that women with a male co-twin have lower left hand 2D:4D than women with a female co-twin, there is no statistically significant result for the right hand (van Anders et al., 2006). Another study similarly reports a significant result for mean 2D:4D (Voracek & Dressler, 2007), whereas studies with larger sample

sizes report no significant differences (Hiraishi et al., 2012; Medland et al., 2008).

There are also studies that have attempted to establish a link between 2D:4D and congenital adrenal hyperplasia (CAH, a disease that results in an excess production of testosterone) as well as CAG repeat polymorphism (which affects the transcriptional activity of the androgen receptor) - these results also consist of a mix of positive and null results (see, e.g. Brown et al., 2002; Buck et al., 2003; Okten et al., 2002). Finally, men with Klinefender’s Syndrome, in which testosterone deficiency is one of the main features, have been shown to have higher 2D:4D than male controls (Manning et al., 2013).

In sum, while there is some evidence in support of a link between 2D:4D and testosterone exposure in utero, the results do not provide strong evidence and it should be a research priority to establish if such a link actually exists based on a large sample pre-registered study. We would also encourage pre-registered large sample replication studies on the association between 2D:4D and other traits previously linked to 2D:4D such as personality, sexual orientation and spatial ability.

5 Open practices statement

The pre-analysis plan and hypotheses of this study were formally pre-registered in Open Science Framework prior to data collection (<https://osf.io/5vpdn/>). The analysis code can be found in the same link. The dataset from SOEP-IS is open to whole research community. Please contact LN (lneyse@diw.de) for more dataset and further information

References

- Alonso, J., Di Paolo, R., Ponti, G., & Sartarelli, M. (2018). Facts and Misconceptions about 2d:4d, Social and Risk Preferences. *Frontiers in Behavioral Neuroscience*, *12*. <https://doi.org/10.3389/fnbeh.2018.00022>
- Apicella, C. L., Dreber, A., Campbell, B., Gray, P. B., Hoffman, M., & Little, A. C. (2008). Testosterone and financial risk preferences. *Evolution and Human Behavior*, *29*(6), 384–390. <https://doi.org/10.1016/j.evolhumbehav.2008.07.001>

- Apicella, C. L., Tobolsky, V. A., Marlowe, F. W., & Miller, K. W. (2015). Hadza hunter-gatherer men do not have more masculine digit ratios (2d:4d). *American Journal of Physical Anthropology*, *159*(2), 223–232. <https://doi.org/10.1002/ajpa.22864>
- Arnold, A. P. (2009). The organizational–activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues. *Hormones and Behavior*, *55*(5), 570–578. <https://doi.org/10.1016/j.yhbeh.2009.03.011>
- Aycinena, D., Baltaduonis, R., & Rentschler, L. (2014). Risk preferences and prenatal exposure to Sex hormones for ladinos. *PLOS ONE*, *9*(8), e103332. <https://doi.org/10.1371/journal.pone.0103332>
- Barel, E. (2019). 2d: 4d, optimism, and risk taking. *Current Psychology*, *38*(1), 204–212. <https://doi.org/10.1007/s12144-017-9598-4>
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*(6), 248–254. [https://doi.org/10.1016/S1364-6613\(02\)01904-6](https://doi.org/10.1016/S1364-6613(02)01904-6)
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Et al. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6. <https://doi.org/doi.org/10.1038/s41562-017-0189-z>
- Bönte, W., Procher, V. D., & Urbig, D. (2016). Biology and selection into entrepreneurship—the relevance of prenatal testosterone exposure. *Entrepreneurship Theory and Practice*, *40*(5), 1121–1148. <https://doi.org/10.1111/etap.12165>
- Bos, P. A., Terburg, D., & Van Honk, J. (2010). Testosterone decreases trust in socially naive humans. *Proceedings of the National Academy of Sciences*, *107*(22), 9991–9995. <https://doi.org/10.1073/pnas.0911700107>
- Brañas-Garza, P., Espin, A. M., Garcia-Muñoz, T., & Kovářik, J. (2019). Digit ratio (2d:4d) and prosocial behaviour in economic games: No direct correlation with generosity, bargaining or trust-related behaviours. *Biology Letters*, *15*(8), 20190185. <https://doi.org/10.1098/rsbl.2019.0185>
- Brañas-Garza, P., Galizzi, M. M., & Nieboer, J. (2018). Experimental and self-reported measures of risk taking and digit ratio (2d: 4d): Evidence from a large, systematic

- study. *International Economic Review*, 59(3), 1131–1157. <https://doi.org/doi.org/10.1111/iere.12299>
- Brañas-Garza, P., Kovářík, J., & Neyse, L. (2013). Second-to-Fourth Digit Ratio Has a Non-Monotonic Impact on Altruism. *PLOS ONE*, 8(4), e60419. <https://doi.org/10.1371/journal.pone.0060419>
- Brañas-Garza, P., & Rustichini, A. (2011). Organizing Effects of Testosterone and Economic Behavior: Not Just Risk Taking. *PLOS ONE*, 6(12), e29842. <https://doi.org/10.1371/journal.pone.0029842>
- Brown, W. M., Hines, M., Fane, B. A., & Breedlove, S. M. (2002). Masculinized finger length patterns in human males and females with congenital adrenal hyperplasia. *Hormones and Behavior*, 42(4), 380–386. <https://doi.org/10.1006/hbeh.2002.1830>
- Buck, J. J., Williams, R. M., Hughes, I. A., & Acerini, C. L. (2003). In-utero androgen exposure and 2nd to 4th digit length ratio-comparisons between healthy controls and females with classical congenital adrenal hyperplasia. *Human Reproduction (Oxford, England)*, 18(5), 976–979.
- Buser, T. (2012). Digit ratios, the menstrual cycle and social preferences. *Games and Economic Behavior*, 76(2), 457–470. <https://doi.org/10.1016/j.geb.2012.07.006>
- Chicaiza-Becerra, L. A., & Garcia-Molina, M. (2017). Prenatal testosterone predicts financial risk taking: Evidence from Latin America. *Personality and Individual Differences*, 116, 32–37. <https://doi.org/10.1016/j.paid.2017.04.021>
- Coates, J. M., Gurnell, M., & Rustichini, A. (2009). Second-to-fourth digit ratio predicts success among high-frequency financial traders. *Proceedings of the National Academy of Sciences*, 106(2), 623–628. <https://doi.org/10.1073/pnas.0810907106>
- Crosan, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2), 448–74. <https://doi.org/10.1257/jel.47.2.448>
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522–550. <https://doi.org/10.1111/j.1542-4774.2011.01015.x>

- Dreber, A., & Hoffman, M. (2007). *Portfolio selection in utero* (Mimeo).
- Dreber, A., & Johannesson, M. (2019). Sex Hormones and Economic Decision Making in the Lab: A Review of the Causal Evidence, In *Routledge International Handbook of Social Neuroendocrinology*.
- Drichoutis, A. C., & Nayga, R. M. (2015). Do risk and time preferences have biological roots? *Southern Economic Journal*, *82*(1), 235–256. <https://doi.org/10.4284/0038-4038-2013.246>
- Eisenegger, C., Haushofer, J., & Fehr, E. (2011). The role of testosterone in social interaction. *Trends in Cognitive Sciences*, *15*(6), 263–271. <https://doi.org/10.1016/j.tics.2011.04.008>
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, *133*(4), 1645–1692. <https://doi.org/10.1093/qje/qjy013>
- Galizzi, M. M., & Nieboer, J. (2015). Digit ratio (2d:4d) and altruism: Evidence from a large, multi-ethnic sample. *Frontiers in Behavioral Neuroscience*, *9*. <https://doi.org/10.3389/fnbeh.2015.00041>
- Garbarino, E., Slonim, R., & Sydnor, J. (2011). Digit ratios (2d:4d) as predictors of risky decision making for both sexes. *Journal of Risk and Uncertainty*, *42*(1), 1–26. <https://doi.org/10.1007/s11166-010-9109-6>
- Gelman, A., & Loken, E. (2013). *The Garden of Forking Paths: Why Multiple Comparisons Can Be a Problem, Even When There Is No “Fishing Expedition” or “p-hacking” and the Research Hypothesis Was Posited Ahead of Time* (Working Paper).
- Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, *77*(5), 1637–1664. <https://doi.org/10.3982/ECTA6690>
- Goebel, J., Grabka, M. M., Liebig, S., Kroh, D., Martin and, Schröder, C., & Schupp, J. (2019). The German Socio-economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*, *239*(2), 345–360. <https://doi.org/10.1515/jbnst-2018-0022>

- Grimbos, T., Dawood, K., Burriss, R. P., Zucker, K. J., & Puts, D. A. (2010). Sexual orientation and the second to fourth finger length ratio: A meta-analysis in men and women. *Behavioral Neuroscience*, *124*(2), 278. <https://doi.org/10.1037/a0018764>
- Hiraishi, K., Sasaki, S., Shikishima, C., & Ando, J. (2012). The second to fourth digit ratio (2d:4d) in a Japanese twin sample: Heritability, prenatal hormone transfer, and association with sexual orientation. *Archives of Sexual Behavior*, *41*(3), 711–724. <https://doi.org/10.1007/s10508-011-9889-z>
- Hönekopp, J., & Schuster, M. (2010). A meta-analysis on 2d: 4d and athletic prowess: Substantial relationships but neither hand out-predicts the other. *Personality and Individual Differences*, *48*(1), 4–10. <https://doi.org/10.1016/j.paid.2009.08.009>
- Hönekopp, J., & Watson, S. (2011). Meta-analysis of the relationship between digit-ratio 2d: 4d and aggression. *Personality and Individual Differences*, *51*(4), 381–386. <https://doi.org/10.1016/j.paid.2010.05.003>
- Lima de Miranda, K., Neyse, L., & Schmidt, U. (2018). Risk Preferences and Predictions about Others: No Association with 2d:4d Ratio. *Frontiers in Behavioral Neuroscience*, *12*, 9. <https://doi.org/10.3389/fnbeh.2018.00009>
- Lombardo, M. V., Ashwin, E., Auyeung, B., Chakrabarti, B., Taylor, K., Hackett, G., Bullmore, E. T., & Baron-Cohen, S. (2012). Fetal Testosterone Influences Sexually Dimorphic Gray Matter in the Human Brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *32*(2), 674–680. <https://doi.org/10.1523/JNEUROSCI.4389-11.2012>
- Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., & Manning, J. (2004). 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development*, *77*(1), 23–28. <https://doi.org/10.1016/j.earlhumdev.2003.12.002>
- Manning, J. T., Kilduff, L. P., & Trivers, R. (2013). Digit ratio (2d:4d) in Klinefelter's syndrome. *Andrology*, *1*(1), 94–99. <https://doi.org/10.1111/j.2047-2927.2012.00013.x>
- Manning, J. T., Scutt, D., Wilson, J., & Lewis-Jones, D. I. (1998). The ratio of 2nd to 4th digit length: A predictor of sperm numbers and concentrations of testos-

- terone, luteinizing hormone and oestrogen. *Human Reproduction (Oxford, England)*, *13*(11), 3000–3004. <https://doi.org/10.1093/humrep/13.11.3000>
- Medland, S. E., Loehlin, J. C., & Martin, N. G. (2008). No effects of prenatal hormone transfer on digit ratio in a large sample of same-and opposite-sex dizygotic twins. *Personality and Individual Differences*, *44*(5), 1225–1234. <https://doi.org/10.1016/j.paid.2007.11.017>
- Miller, E. M. (1994). Prenatal sex hormone transfer: A reason to study opposite-sex twins. *Personality and Individual Differences*, *17*(4), 511–529. [https://doi.org/10.1016/0191-8869\(94\)90088-4](https://doi.org/10.1016/0191-8869(94)90088-4)
- Millet, K., & Dewitte, S. (2010). The presence of aggression cues inverts the relation between digit ratio (2d:4d) and prosocial behaviour in a dictator game. *British Journal of Psychology*, *100*(1), 151–162. <https://doi.org/10.1348/000712608X324359>
- Neyse, L., Vieider, F. M., Ring, P., Probst, C., Kaernbach, C., van Eimeren, T., & Schmidt, U. (2020). Risk attitudes and digit ratio (2d: 4d): Evidence from prospect theory. *Journal of Risk and Uncertainty*, 1–23. <https://doi.org/10.1007/s11166-020-09321-w>
- Nicolaou, N., Patel, P. C., & Wolfe, M. T. (2017). Testosterone and tendency to engage in self-employment. *Management Science*, *64*(4), 1825–1841. <https://doi.org/10.1287/mnsc.2016.2664>
- Okten, A., Kalyoncu, M., & Yariş, N. (2002). The ratio of second- and fourth-digit lengths and congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Early Human Development*, *70*(1-2), 47–54. [https://doi.org/10.1016/S0378-3782\(02\)00073-7](https://doi.org/10.1016/S0378-3782(02)00073-7)
- Parslow, E., Raneyhill, E., Zethraeus, N., Blomberg, L., von Schoultz, B., Hirschberg, A. L., Johannesson, M., & Dreber, A. (2019). The digit ratio (2d:4d) and economic preferences: No robust associations in a sample of 330 women. *Journal of the Economic Science Association*. <https://doi.org/10.1007/s40881-019-00076-y>
- Phoenix, C. H., Goy, R. W., Gerall, A. A., & Young, W. C. (1959). Organizing action of prenatally administered testosterone propionate on the tissues mediating mating

- behavior in the female guinea pig. *Endocrinology*, *65*, 369–382. <https://doi.org/10.1210/endo-65-3-369>
- Puts, D. A., McDaniel, M. A., Jordan, C. L., & Breedlove, S. M. (2008). Spatial ability and prenatal androgens: Meta-analyses of congenital adrenal hyperplasia and digit ratio (2d: 4d) studies. *Archives of Sexual Behavior*, *37*(1), 100. <https://doi.org/10.1007/s10508-007-9271-3>
- Richter, D., & Schupp, J. (2015). The soep innovation sample (SOEP" IS). *Schmollers Jahrbuch: Journal of Applied Social Science Studies/Zeitschrift für Wirtschafts- und Sozialwissenschaften*, *135*(3), 389–400. <https://doi.org/10.3790/schm.135.3.389>
- Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences*, *106*(36), 15268–15273. <https://doi.org/10.1073/pnas.0907352106>
- Schipper, B. (2014). *Sex hormones and choice under risk* (tech. rep. No. 147). University of California, Davis, Department of Economics. Retrieved August 21, 2018, from <https://ideas.repec.org/p/cda/wpaper/14-7.html>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Stenstrom, E., Saad, G., Nepomuceno, M. V., & Mendenhall, Z. (2011). Testosterone and domain-specific risk: Digit ratios (2d:4d and rel2) as predictors of recreational, financial, and social risk-taking behaviors. *Digit Ratio (2D:4D) and Individual Differences Research*, *51*(4), 412–416. <https://doi.org/10.1016/j.paid.2010.07.003>
- Sytsma, T. (2014). *Handling Risk: Testosterone and Risk Preference, Evidence from Dhaka, Bangladesh* (Master thesis). <https://repository.usfca.edu/thes/85>
- Thaler, R. H., & Johnson, E. J. (1990). Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. *Management Science*,

36(6), 643–660. <https://doi.org/10.1287/mnsc.36.6.643>

doi: 10.1287/mnsc.36.6.643

- Van den Bergh, B., & Dewitte, S. (2006). Digit ratio (2d:4d) moderates the impact of sexual cues on men's decisions in ultimatum games. *Proceedings of the Royal Society B*, 273(1597), 2091–2095. <https://doi.org/10.1098/rspb.2006.3550>
- Van Honk, J., Montoya, E. R., Bos, P. A., Van Vugt, M., & Terburg, D. (2012). New evidence on testosterone and cooperation. *Nature*, 485(7399), E4. <https://doi.org/10.1038/nature11136>
- van Anders, S. M., Vernon, P. A., & Wilbur, C. J. (2006). Finger-length ratios show evidence of prenatal hormone-transfer between opposite-sex twins. *Hormones and Behavior*, 49(3), 315–319. <https://doi.org/10.1016/j.yhbeh.2005.08.003>
- van Leeuwen, B., Smeets, P., Bovet, J., Nave, G., Stieglitz, J., & Whitehouse, A. (2020). No link between sex hormones around birth and economic preferences. *mimeo*.
- Ventura, T., Gomes, M., Pita, A., Neto, M., & Taylor, A. (2013). Digit ratio (2d:4d) in newborns: Influences of prenatal testosterone and maternal environment. *Early Human Development*, 89(2), 107–112. <https://doi.org/10.1016/j.earlhumdev.2012.08.009>
- Voracek, M., & Dressler, S. G. (2007). Digit ratio (2d:4d) in twins: Heritability estimates and evidence for a masculinized trait expression in women from opposite-sex pairs. *Psychological Reports*, 100(1), 115–126. <https://doi.org/10.2466/pr0.100.1.115-126>
- Voracek, M., & Loibl, L. M. (2009). Scientometric Analysis and Bibliography of Digit Ratio (2d:4d) Research, 1998–2008. *Psychological Reports*, 104(3), 922–956. <https://doi.org/10.2466/PR0.104.3.922-956>
- Voracek, M., Pietschnig, J., Nader, I. W., & Stieger, S. (2011). Digit ratio (2d:4d) and sex-role orientation: Further evidence and meta-analysis. *Digit Ratio (2D:4D) and Individual Differences Research*, 51(4), 417–422. <https://doi.org/10.1016/j.paid.2010.06.009>
- Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., Von Schoultz, B., Hirschberg, A. L., & Johannesson, M. (2009). A randomized trial of the effect of estrogen and testos-

terone on economic behavior. *Proceedings of the National Academy of Sciences*,
106(16), 6535–6538. <https://doi.org/10.1073/pnas.0812757106>

Supplementary Information for

2D:4D does not predict economic preferences: Evidence from a large, representative sample

This PDF file includes:

Supplementary text

Figs. S1 to S10

Tables S1 to S25

SI References

Supporting Information Text

Preferences Module

To elicit the economic preferences, we integrated the following set of items to the SOEP-IS 2018 wave. These items were adapted from [Falk et al. \(2018\)](#), who experimentally validated them in their preference module. We use their weighting formulas to generate our altruism and negative reciprocity variables. The risk attitudes item was also experimentally validated by [Dohmen et al. \(2011\)](#). For positive reciprocity, we use only the self reported question of [Falk et al. \(2018\)](#). As our proposed trust item "As long as I am not convinced otherwise, I assume that people have only the best intentions" was similar to another trust item, "People are basically honest" in SOEP-IS 2018 wave, the survey management decided to include only the latter item. The items we used are as follows. For the items that require responses on the 10 point Likert Scale 0 refers to "I do not agree at all" and 10 refers to "I fully Agree".

1. Risk Attitudes:

How do you rate yourself personally? In general, are you someone who is ready to take risks or do you try to avoid risks? [0,10] (0 refers to risk aversion and 10 refers to risk seeking)

2. Altruism Item 1:

Imagine the following situation: Today you unexpectedly received 1,000 Euro. How much of this amount would you donate to a good cause? [0,1000]

3. Altruism Item 2:

How willing are you to give to good causes without expecting anything in return? [0,10]

$$\text{Altruism} = 0.4649952 \times \text{Altruism Item 1} + 0.5350048 \times \text{Altruism Item 2}$$

4. Positive Reciprocity:

When someone does me a favor I am willing to return it. [0,10]

5. Negative Reciprocity Item 1:

If I am treated very unjustly, I will take revenge at the first occasion, even if there is a cost to do so. [0,10]

6. Negative Reciprocity Item 2:

How willing are you to punish someone who treats you unfairly, even if there may be costs for you? [0,10]

7. Negative Reciprocity Item 3:

How willing are you to punish someone who treats others unfairly, even if there may be costs for you? [0,10]

$$\begin{aligned} \text{Negative Reciprocity} &= 0.3738062 \times \text{Negative Reciprocity Item 1} \\ &+ 0.5261938/2 \times \text{Negative Reciprocity Item 2} + 0.5261938/2 \times \text{Negative Reciprocity Item 3} \end{aligned}$$

8. Trust:

People are basically honest. [0,10]

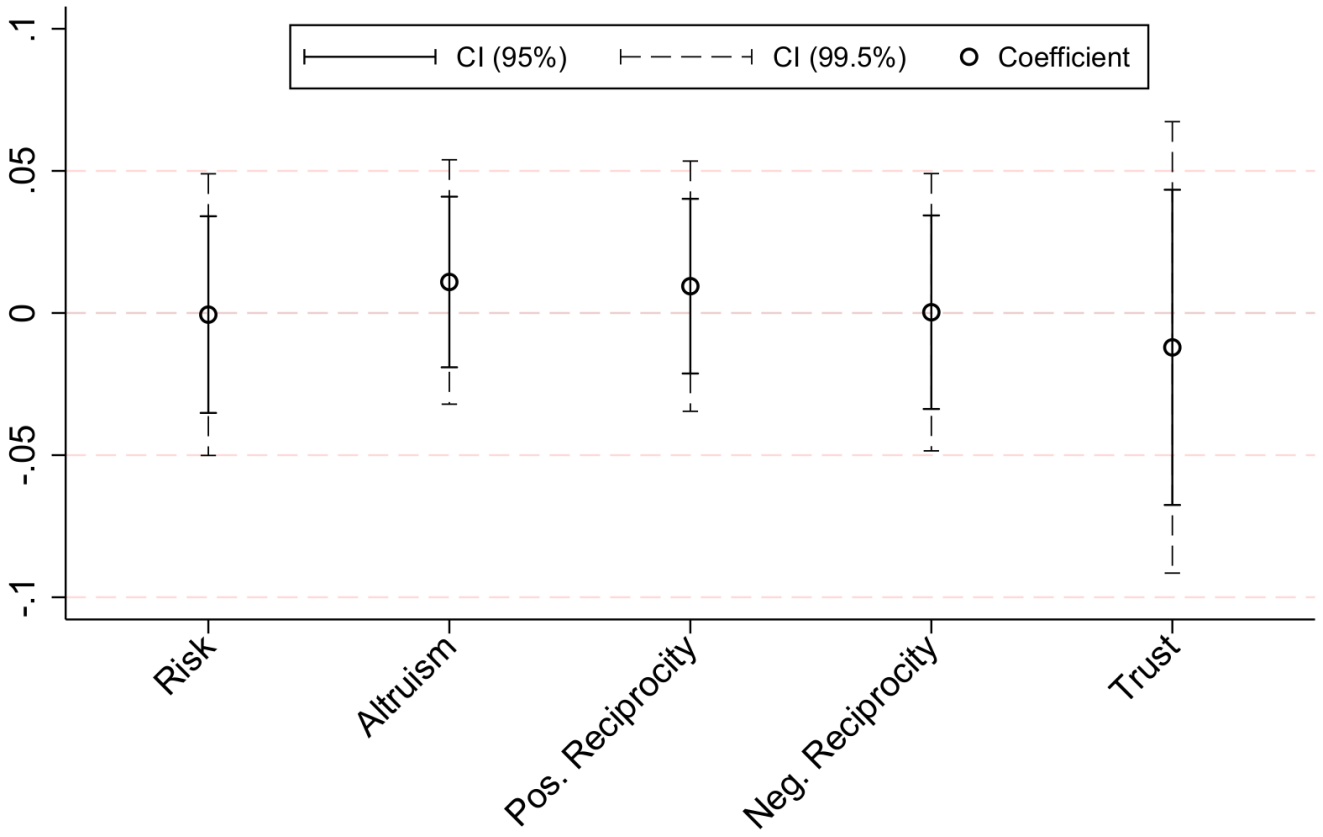


Fig. S1. Standardized regression coefficients of R2D:4D for primary hypotheses based on corrected sample in Table S2. Standardized R2D:4D (*y*-axis) was used to generate the figure. Interval bars refer to 95% and 99.5% confidence intervals.

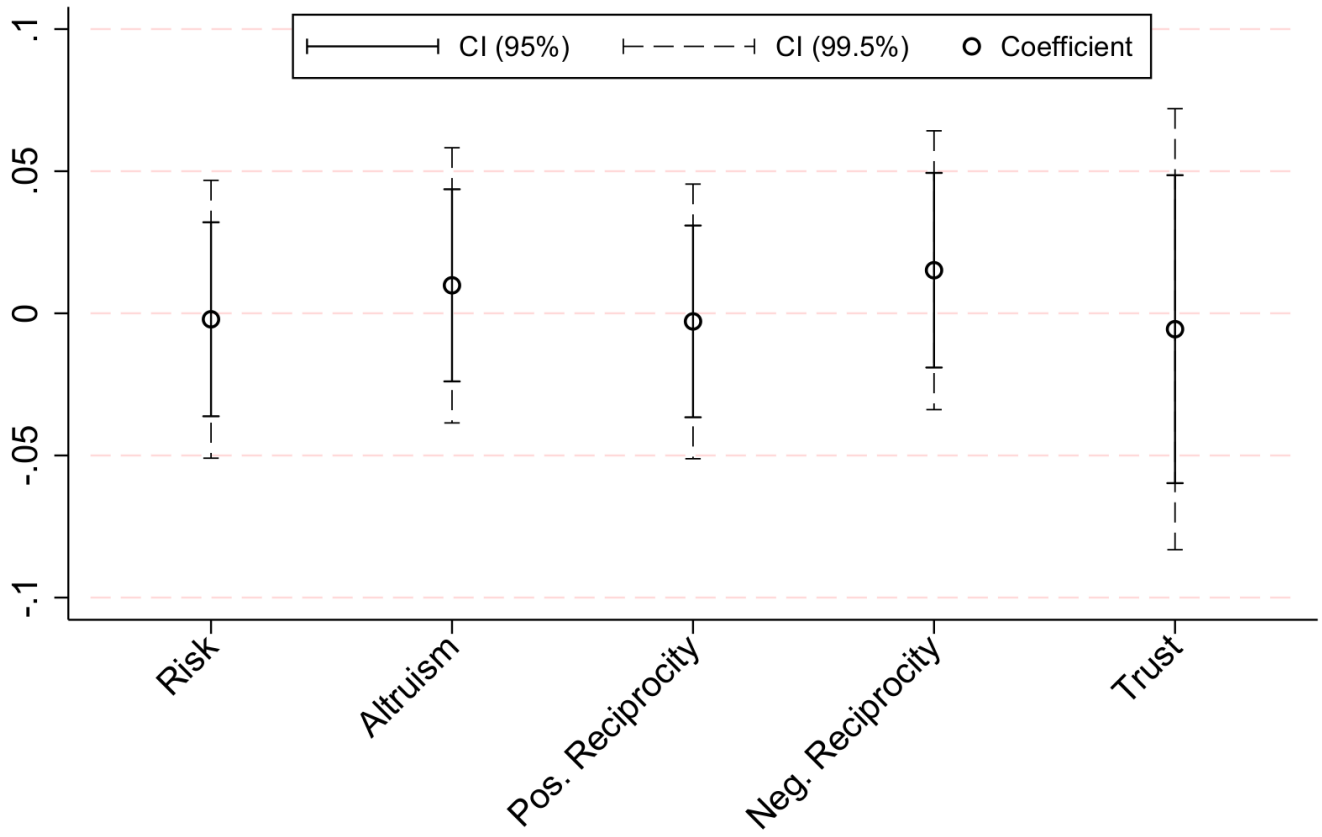


Fig. S2. Standardized regression coefficients of R2D:4D for primary hypotheses based on restricted sample in Table S5. Standardized R2D:4D (*y*-axis) was used to generate the figure. Interval bars refer to 95% and 99.5% confidence intervals.

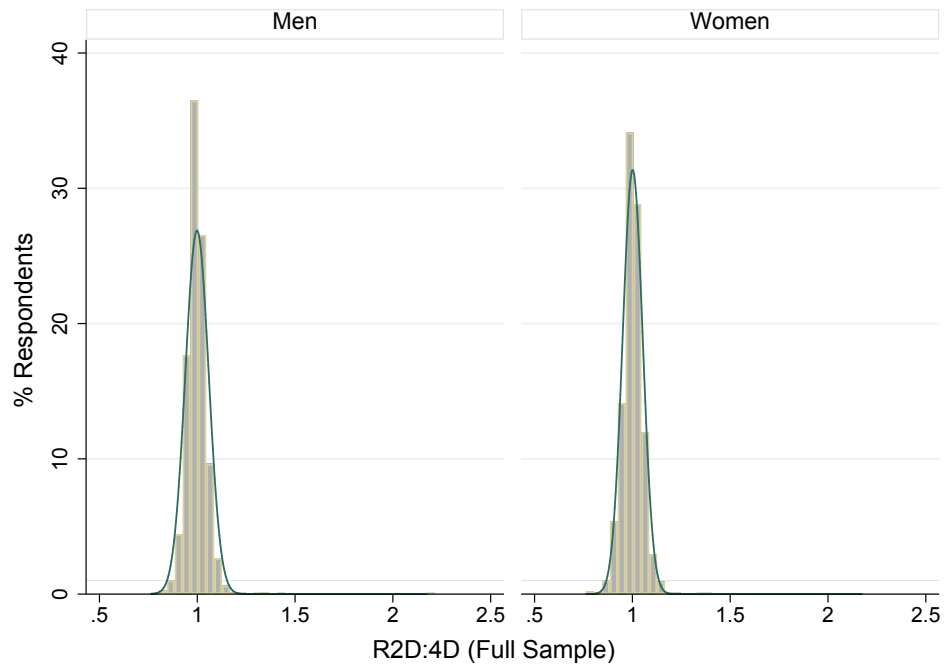


Fig. S3. Distribution of right hand 2D:4D across gender based on full sample, including outliers.

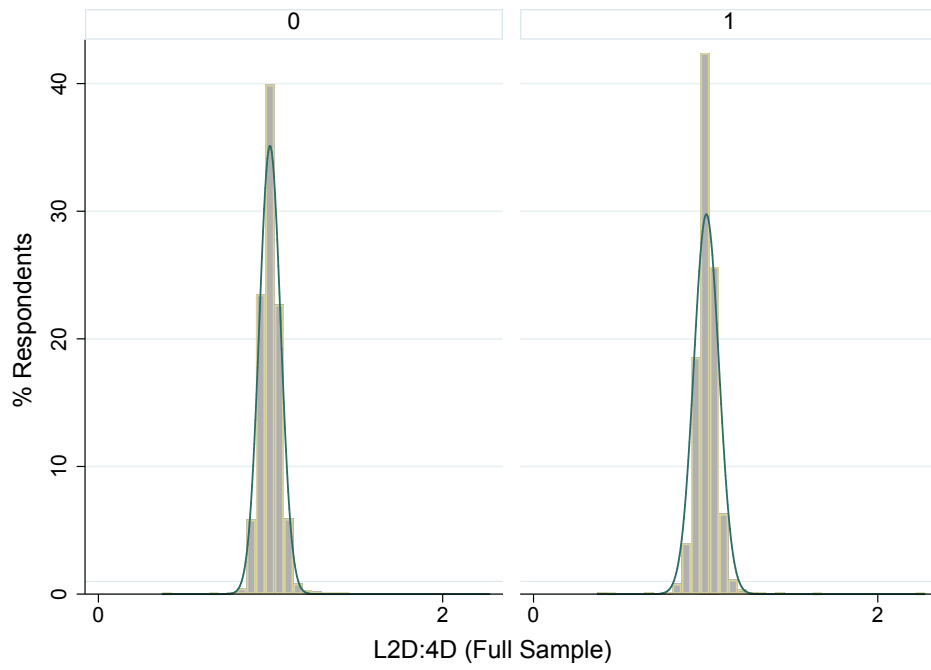


Fig. S4. Distribution of left hand 2D:4D across gender based on full sample, including outliers.

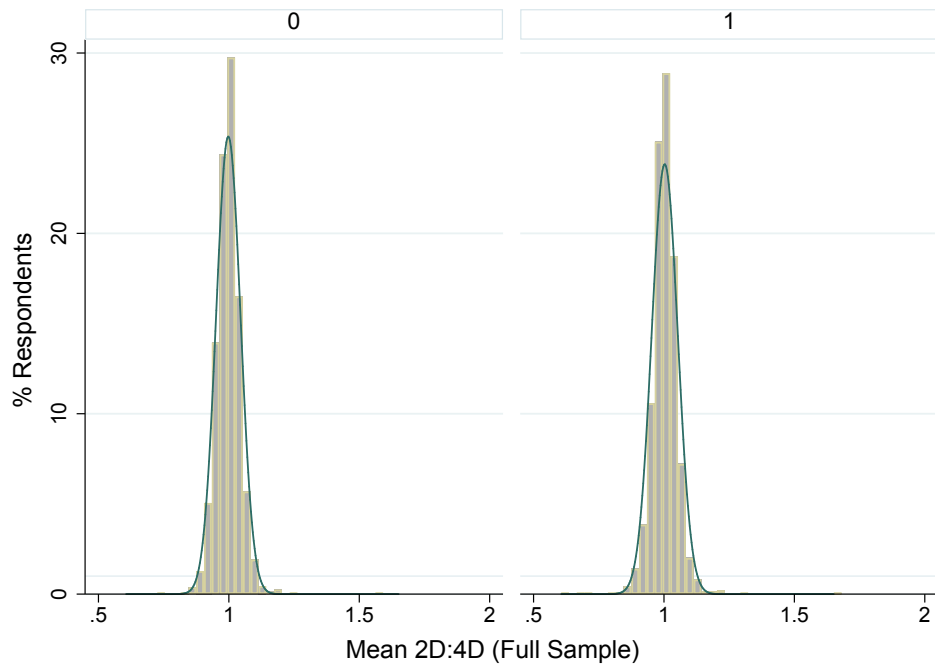


Fig. S5. Distribution of 2D:4D mean of both hands across gender based on full sample, including outliers.

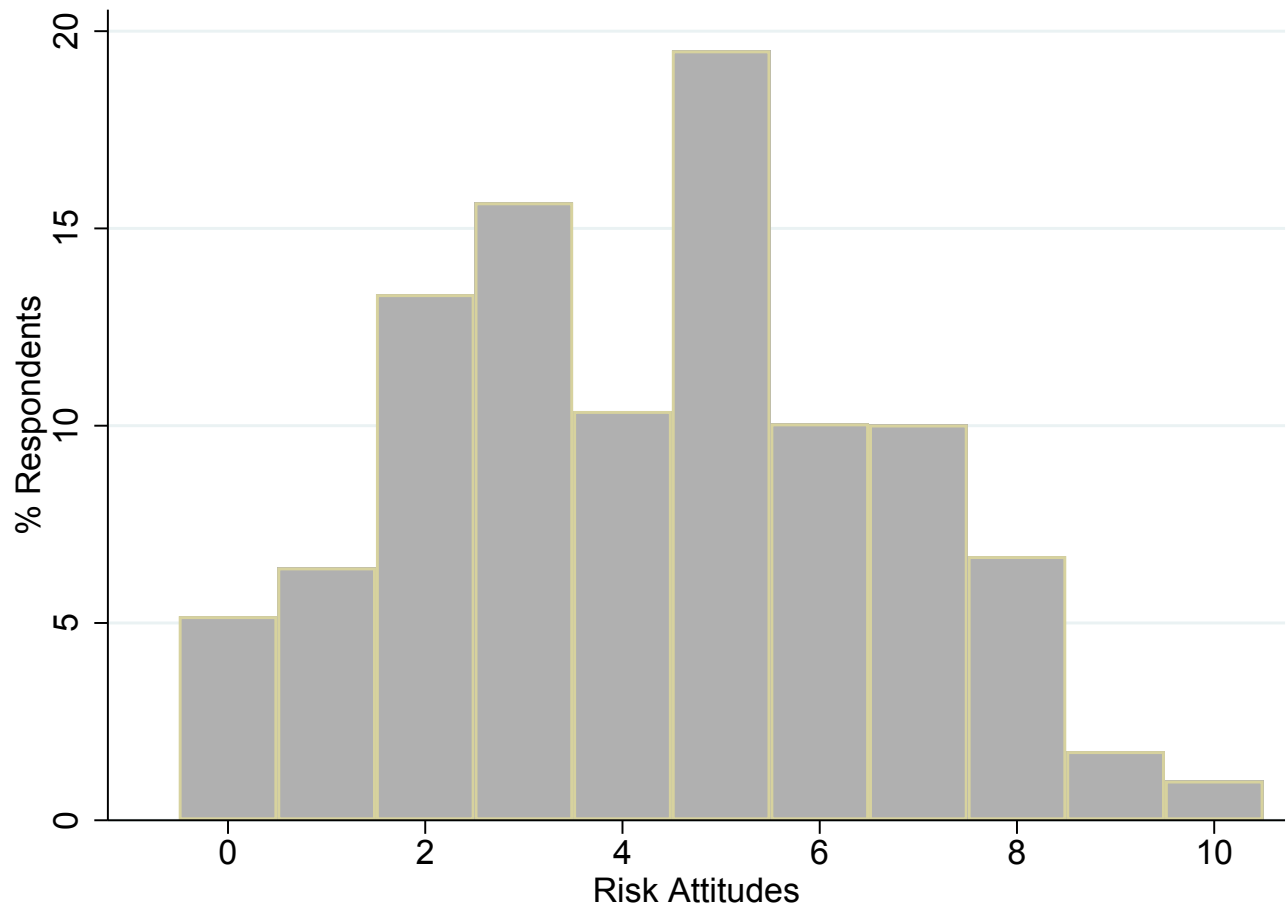


Fig. S6. Distribution of responses to *risk attitudes* item [0,10]. 0 refers to "Do not agree at all" and 10 refers to "Fully Agree". *y*-axis presents the percentage of respondents in the working sample. This variable has been standardized in the regression analyses.

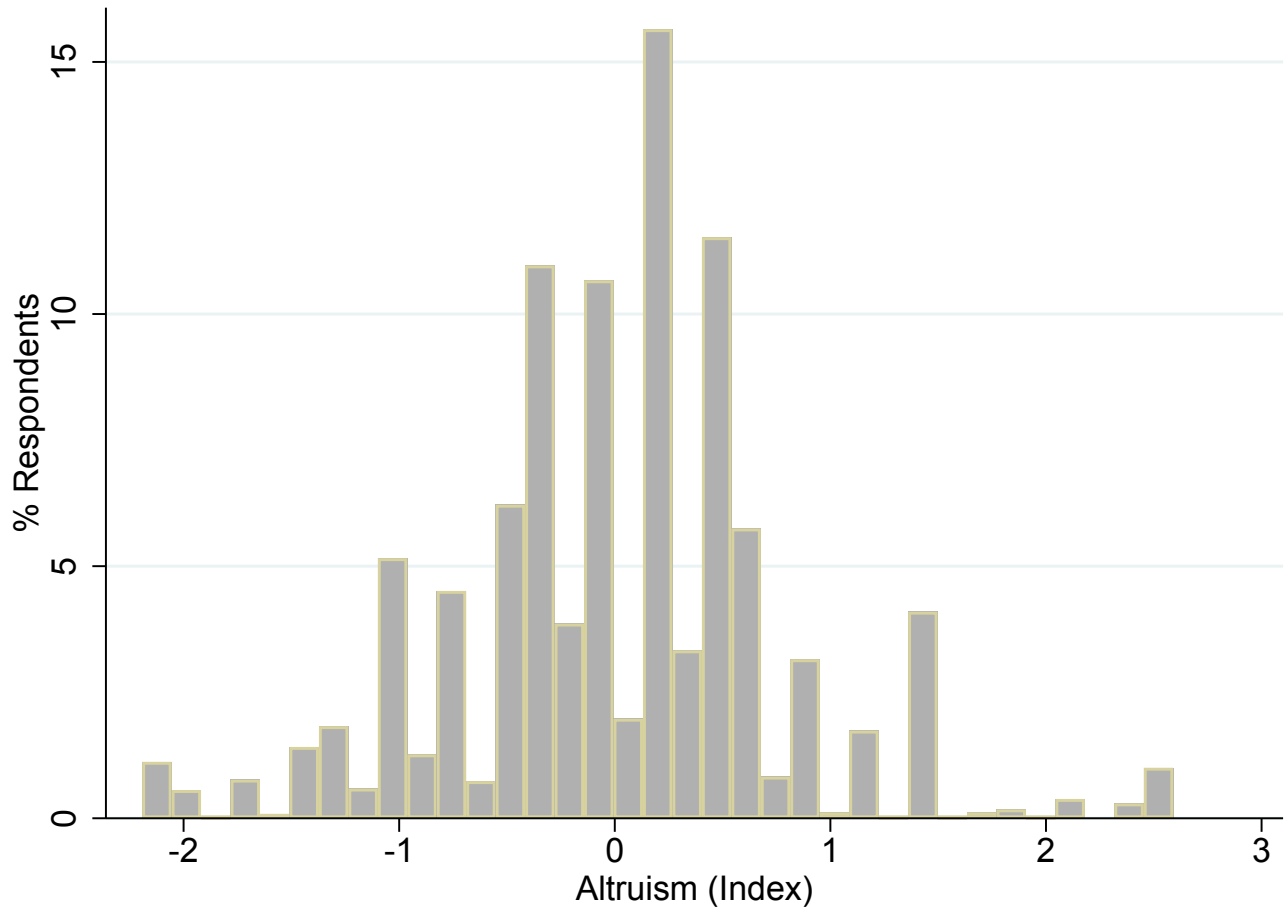


Fig. S7. Distribution of *altruism* index variable generated as in [Falk et al. \(2018\)](#). See Supporting Information Text for weights and calculation. *y*-axis presents the percentage of respondents in the working sample. This variable has been standardized in the regression analyses.

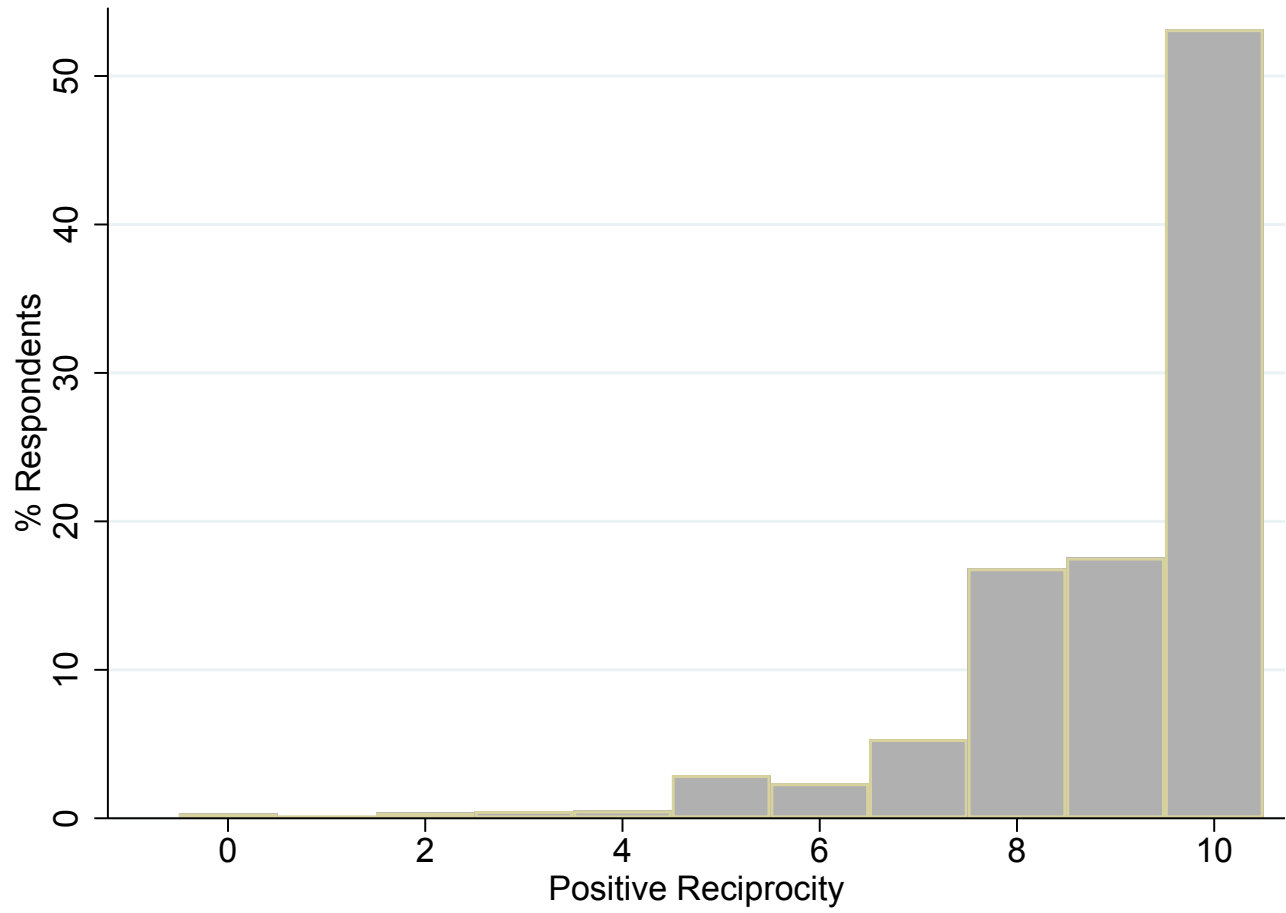


Fig. S8. Distribution of responses to *positive reciprocity* item [0,10]. 0 refers to "Do not agree at all" and 10 refers to "Fully Agree". *y*-axis presents the percentage of respondents in the working sample. This variable has been standardized in the regression analyses.

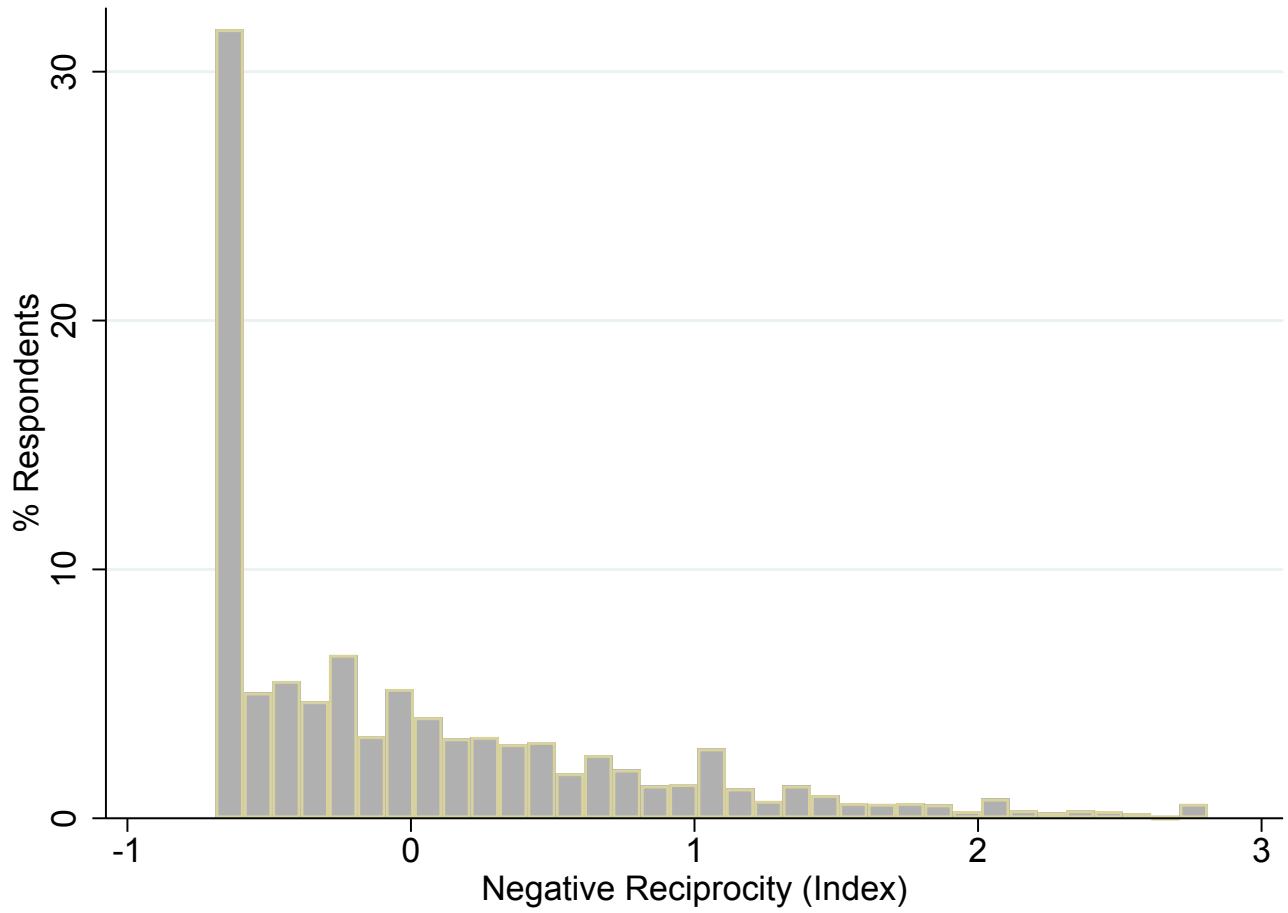


Fig. S9. Distribution of *negative reciprocity* index variable which was generated as in [Falk et al. \(2018\)](#). See Supporting Information Text for weights and calculation. *y*-axis presents the percentage of respondents in the working sample. This variable has been standardized in the regression analyses.

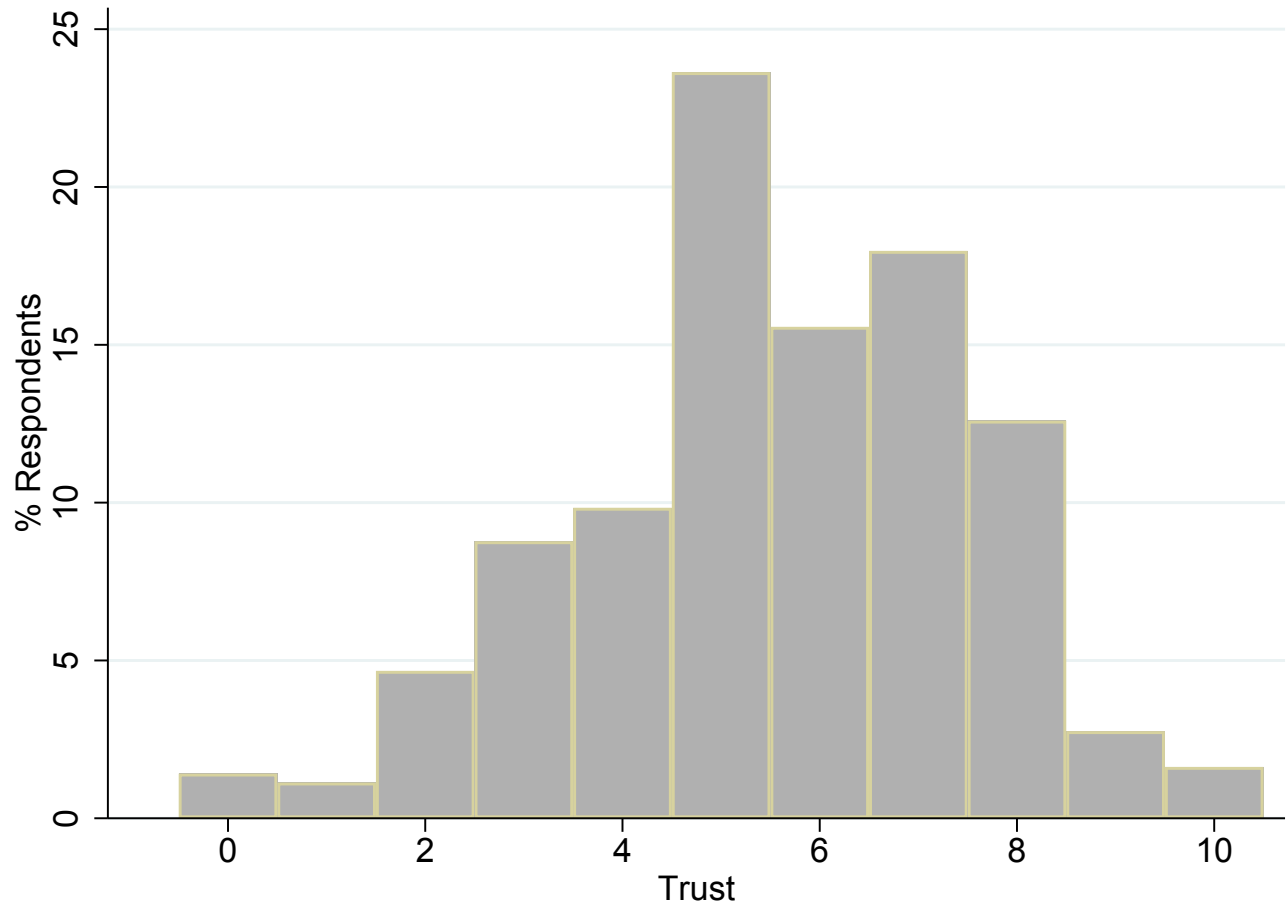


Fig. S10. Distribution of responses to *trust* item [0,10]. 0 refers to "Do not agree at all" and 10 refers to "Fully Agree". *y*-axis presents the percentage of respondents in the working sample. This variable has been standardized in the regression analyses.

Table S1. 2D:4D Descriptives by Gender

	All	Men		Women	p-value	e-size
R2D:4D (FS)	1.000	0.999	<	1.002	0.151	0.049
(n=3433)	(0.055)	(0.060)		(0.051)		
R2D:4D (CS)	0.999	0.996	<	1.000	0.021	0.049
(n=3433)	(0.051)	(0.054)		(0.048)		
R2D:4D (RS)	0.999	0.997	<	1.001	0.013	0.086
(n=3419)	(0.048)	(0.048)		(0.049)		
L2D:4D (FS)	1.001	0.997	<	1.004	0.001	0.112
(n=3454)	(0.068)	(0.062)		(0.073)		
L2D:4D (CS)	0.998	0.994	<	1.001	<0.001	0.131
(n=3454)	(0.054)	(0.053)		(0.055)		
L2D:4D (RS)	1.000	0.996	<	1.003	<0.001	0.137
(n=3422)	(0.055)	(0.055)		(0.054)		
Mean 2D:4D (FS)	1.001	0.998	<	1.003	0.003	0.102
(n=3405)	(0.049)	(0.047)		(0.050)		
Mean 2D:4D (CS)	0.998	0.995	<	1.001	<0.001	0.127
(n=3405)	(0.043)	(0.043)		(0.042)		
Mean 2D:4D (RS)	1.000	0.998	<	1.002	0.001	0.112
(n=3393)	(0.044)	(0.044)		(0.044)		

Note: Standard deviations are in parentheses. Reported tests are based on t-tests; effect sizes (*cohen's d*) are reported in e-size column. Full Sample (FS) includes participants with injured hands and outliers. In the Corrected Sample (CS) 1st measures were replaced with available 2nd measures. The Restricted Sample (RS) involves 2D:4D measurements in the range between 0.8 and 1.2.

Table S2. Regression Analysis based on R2D:4D: Primary Hypotheses (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
R2D:4D	-0.011 (0.345)	0.215 (0.300)	0.184 (0.306)	0.006 (0.340)	-0.238 (0.557)
female	-0.361** (0.034)	0.222** (0.035)	0.066 (0.034)	-0.247** (0.034)	-0.084 (0.054)
constant	0.208 (0.345)	-0.336 (0.300)	-0.219 (0.306)	0.126 (0.341)	0.276 (0.555)
<i>N</i>	3431	3350	3422	3378	1391
adj. <i>R</i> ²	0.032	0.012	0.001	0.015	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S3. Regression Analysis based on R2D:4D: Exploratory Analysis 1 (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
R2D:4D	-0.563 (0.578)	0.215 (0.398)	-0.103 (0.449)	-0.524 (0.427)	0.635 (0.864)
female	-1.501 (0.771)	0.221 (0.604)	-0.528 (0.637)	-1.337* (0.636)	1.467 (1.126)
R2D:4D X female	1.142 (0.772)	0.001 (0.603)	0.594 (0.636)	1.091 (0.636)	-1.557 (1.129)
constant	0.759 (0.575)	-0.336 (0.398)	0.067 (0.448)	0.654 (0.429)	-0.590 (0.860)
R2D:4D (F)	0.578 (0.512)	0.216 (0.453)	0.491 (0.450)	0.567 (0.470)	-0.922 (0.726)
<i>N</i>	3431	3350	3422	3378	1391
adj. <i>R</i> ²	0.032	0.012	0.001	0.015	0.001

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,

** $p < 0.005$. R2D:4D (F) presents the R2D:4D effect for women which were calculated by summing the coefficients of R2D:4D and R2D:4D X female variables.

Table S4. Regression Analysis based on R2D:4D: Exploratory Analysis 2 (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
R2D:4D	-0.847 (1.285)	0.291 (1.022)	-0.895 (0.971)	2.090* (1.021)	-6.361 (6.052)
female	-0.360** (0.034)	0.222** (0.035)	0.066 (0.034)	-0.249** (0.034)	-0.085 (0.054)
R2D:4D ²	0.384 (0.517)	-0.0348 (0.343)	0.495 (0.319)	-0.954** (0.338)	3.051 (2.926)
constant	0.660 (0.790)	-0.378 (0.682)	0.363 (0.655)	-1.000 (0.687)	3.342 (3.141)
<i>N</i>	3431	3350	3422	3378	1391
adj. <i>R</i> ²	0.032	0.012	0.000	0.015	0.000

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S5. Regression Analysis based on R2D:4D: Primary Hypotheses (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
R2D:4D	-0.043 (0.360)	0.204 (0.357)	-0.059 (0.355)	0.314 (0.361)	-0.116 (0.574)
female	-0.365** (0.034)	0.219** (0.035)	0.066 (0.034)	-0.249** (0.035)	-0.085 (0.054)
constant	0.244 (0.359)	-0.325 (0.357)	0.022 (0.355)	-0.179 (0.363)	0.156 (0.573)
<i>N</i>	3417	3336	3408	3364	1386
adj. R^2	0.033	0.011	0.000	0.015	0.000

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S6. Regression Analysis based on R2D:4D: Exploratory Analysis 1 (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
R2D:4D	-0.841 (0.547)	0.233 (0.544)	-0.310 (0.535)	-0.448 (0.543)	0.820 (0.826)
female	-1.807* (0.725)	0.271 (0.720)	-0.388 (0.716)	-1.618* (0.727)	1.729 (1.134)
R2D:4D X female	1.443* (0.725)	-0.052 (0.721)	0.455 (0.715)	1.370 (0.726)	-1.819 (1.136)
constant	1.039 (0.545)	-0.354 (0.543)	0.273 (0.534)	0.581 (0.544)	-0.775 (0.823)
R2D:4D (F)	0.603 (0.476)	0.181 (0.473)	0.145 (0.475)	0.922 (0.482)	-0.999 (0.780)
<i>N</i>	3417	3336	3408	3364	1386
adj. <i>R</i> ²	0.034	0.011	0.000	0.016	0.001

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,

** $p < 0.005$. R2D:4D (F) presents the R2D:4D effect for women, which were calculated by summing the coefficients of R2D:4D and R2D:4D X female variables.

Table S7. Regression Analysis based on R2D:4D: Exploratory Analysis 2 (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
R2D:4D	12.250 (9.391)	7.501 (8.682)	2.273 (8.538)	-1.258 (8.634)	-17.050 (14.430)
female	-0.365** (0.034)	0.219** (0.035)	0.066 (0.034)	-0.249** (0.034)	-0.084 (0.054)
R2D:4D ²	-6.152 (4.694)	-3.649 (4.345)	-1.167 (4.297)	0.786 (4.324)	8.511 (7.254)
constant	-5.887 (4.693)	-3.964 (4.335)	-1.141 (4.239)	0.605 (4.307)	8.561 (7.170)
<i>N</i>	3417	3336	3408	3364	1386
adj. <i>R</i> ²	0.033	0.011	0.000	0.015	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S8. Regression Analysis Based on L2D:4D: Primary Hypotheses (Full Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	0.048 (0.250)	0.015 (0.243)	0.143 (0.275)	0.124 (0.250)	0.456 (0.345)
female	-0.363** (0.034)	0.223** (0.035)	0.064 (0.034)	-0.256** (0.034)	-0.085 (0.054)
constant	0.151 (0.250)	-0.133 (0.244)	-0.176 (0.275)	0.013 (0.250)	-0.411 (0.345)
<i>N</i>	3452	3368	3441	3397	1396
adj. <i>R</i> ²	0.032	0.012	0.001	0.016	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S9. Regression Analysis Based on L2D:4D: Exploratory Hypotheses 1 (Full Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	-0.412 (0.437)	-0.178 (0.425)	0.061 (0.432)	0.025 (0.408)	0.468 (0.589)
female	-1.106* (0.527)	-0.089 (0.517)	-0.068 (0.560)	-0.415 (0.518)	-0.066 (0.729)
L2D:4D X female	0.743 (0.527)	0.311 (0.516)	0.132 (0.558)	0.159 (0.518)	-0.019 (0.727)
constant	0.610 (0.436)	0.059 (0.424)	-0.094 (0.431)	0.111 (0.407)	-0.424 (0.587)
L2D:4D (F)	0.331 (0.295)	0.133 (0.293)	0.193 (0.354)	0.184 (0.318)	0.449 (0.427)
<i>N</i>	3452	3368	3441	3397	1396
adj. <i>R</i> ²	0.033	0.012	0.000	0.015	0.001

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,

** $p < 0.005$. L2D:4D (F) presents the L2D:4D effect for women which were calculated by summing the coefficients of L2D:4D and L2D:4D X female variables.

Table S10. Regression Analysis Based on L2D:4D: Exploratory Hypotheses 2 (Full Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	1.312 (1.347)	0.337 (0.910)	0.515 (1.205)	-0.199 (1.108)	-0.476 (1.329)
female	-0.363** (0.034)	0.223** (0.035)	0.065 (0.034)	-0.256** (0.034)	-0.086 (0.054)
L2D:4D ²	-0.597 (0.642)	-0.152 (0.370)	-0.176 (0.549)	0.153 (0.451)	0.413 (0.477)
constant	-0.514 (0.725)	-0.303 (0.557)	-0.371 (0.681)	0.183 (0.668)	0.105 (0.864)
<i>N</i>	3452	3368	3441	3397	1396
adj. <i>R</i> ²	0.032	0.011	0.000	0.015	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S11. Regression Analysis Based on L2D:4D: Primary Hypotheses (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	-0.170 (0.314)	-0.275 (0.327)	0.237 (0.309)	0.286 (0.309)	0.181 (0.535)
female	-0.362** (0.034)	0.225** (0.035)	0.064 (0.034)	-0.257** (0.034)	-0.083 (0.054)
constant	0.367 (0.313)	0.154 (0.326)	-0.269 (0.309)	-0.148 (0.307)	-0.137 (0.534)
<i>N</i>	3452	3368	3441	3397	1396
adj. <i>R</i> ²	0.032	0.012	0.001	0.016	0.000

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S12. Regression Analysis Based on L2D:4D: Exploratory Hypotheses 1 (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	-0.554 (0.487)	-0.467 (0.499)	0.288 (0.461)	0.065 (0.506)	0.302 (0.731)
female	-1.049 (0.635)	-0.121 (0.658)	0.155 (0.622)	-0.652 (0.631)	0.140 (1.060)
L2D:4D X female	0.689 (0.637)	0.347 (0.659)	-0.091 (0.621)	0.396 (0.634)	-0.224 (1.061)
constant	0.749 (0.485)	0.346 (0.497)	-0.319 (0.461)	0.072 (0.502)	-0.258 (0.729)
L2D:4D (F)	0.136 (0.410)	-0.120 (0.430)	0.197 (0.417)	0.461 (0.383)	0.079 (0.770)
<i>N</i>	3452	3368	3441	3397	1396
adj. <i>R</i> ²	0.032	0.012	0.000	0.016	-0.000

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,

** $p < 0.005$. L2D:4D (F) presents the L2D:4D effect for women which were calculated by summing the coefficients of L2D:4D and L2D:4D X female variables.

Table S13. Regression Analysis Based on L2D:4D: Exploratory Hypotheses 2 (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	2.542 (3.863)	3.622 (3.401)	-1.176 (2.277)	0.447 (2.435)	-13.090 (6.990)
female	-0.361** (0.034)	0.226** (0.035)	0.064 (0.034)	-0.257** (0.034)	-0.083 (0.054)
L2D:4D ²	-1.370 (1.968)	-1.969 (1.749)	0.714 (1.137)	-0.082 (1.237)	6.609 (3.417)
constant	-0.971 (1.904)	-1.768 (1.664)	0.428 (1.157)	-0.227 (1.216)	6.503 (3.580)
<i>N</i>	3452	3368	3441	3397	1396
adj. <i>R</i> ²	0.032	0.012	0.000	0.016	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S14. Regression Analysis Based on L2D:4D: Primary Hypotheses (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	-0.391 (0.316)	-0.274 (0.325)	0.228 (0.321)	0.375 (0.319)	-0.079 (0.516)
female	-0.360** (0.034)	0.226** (0.035)	0.063 (0.034)	-0.255** (0.035)	-0.086 (0.054)
constant	0.589 (0.316)	0.153 (0.325)	-0.260 (0.322)	-0.239 (0.318)	0.118 (0.516)
<i>N</i>	3420	3336	3409	3365	1383
adj. <i>R</i> ²	0.033	0.012	0.001	0.016	0.000

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S15. Regression Analysis Based on L2D:4D: Exploratory Hypotheses 1 (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	-0.674 (0.482)	-0.547 (0.475)	0.172 (0.470)	0.191 (0.509)	0.078 (0.736)
female	-0.879 (0.638)	-0.278 (0.652)	-0.039 (0.646)	-0.590 (0.647)	0.204 (1.031)
L2D:4D X female	0.519 (0.638)	0.503 (0.651)	0.103 (0.643)	0.336 (0.649)	-0.290 (1.030)
constant	0.871 (0.481)	0.425 (0.475)	-0.204 (0.471)	-0.056 (0.507)	-0.038 (0.735)
L2D:4D (F)	-0.155 (0.418)	-0.044 (0.445)	0.274 (0.439)	0.526 (0.402)	-0.211 (0.721)
<i>N</i>	3420	3336	3409	3365	1383
adj. <i>R</i> ²	0.033	0.012	0.000	0.015	-0.000

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,
 ** $p < 0.005$. L2D:4D (F) presents the L2D:4D effect for women which
 were calculated by summing the coefficients of L2D:4D and L2D:4D X female
 variables.

Table S16. Regression Analysis Based on L2D:4D: Exploratory Hypotheses 2 (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
L2D:4D	20.830** (6.935)	9.421 (7.688)	-0.565 (7.530)	2.689 (7.039)	-15.250 (11.580)
female	-0.359** (0.034)	0.226** (0.035)	0.063 (0.034)	-0.255** (0.035)	-0.087 (0.054)
L2D:4D ²	-10.600** (3.471)	-4.840 (3.827)	0.396 (3.736)	-1.156 (3.524)	7.584 (5.763)
constant	-10.000** (3.461)	-4.688 (3.856)	0.136 (3.790)	-1.394 (3.513)	7.682 (5.816)
<i>N</i>	3420	3336	3409	3365	1383
adj. <i>R</i> ²	0.035	0.012	0.000	0.015	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S17. Regressions Analysis Based on the Mean of Both Hands: Primary Hypotheses (Full Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	0.085 (0.353)	0.180 (0.338)	0.127 (0.363)	0.203 (0.351)	0.406 (0.546)
female	-0.359** (0.034)	0.217** (0.035)	0.063 (0.034)	-0.250** (0.035)	-0.089 (0.055)
constant	0.115 (0.353)	-0.298 (0.338)	-0.159 (0.363)	-0.070 (0.352)	-0.365 (0.545)
<i>N</i>	3403	3325	3394	3350	1374
adj. R^2	0.032	0.011	0.000	0.015	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S18. Regressions Analysis Based on the Mean of Both Hands: Exploratory Hypotheses 1 (Full Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	-0.733 (0.588)	-0.034 (0.527)	-0.099 (0.559)	-0.407 (0.523)	1.116 (0.874)
female	-1.789* (0.735)	-0.158 (0.687)	-0.332 (0.737)	-1.313 (0.703)	1.057 (1.142)
m2D:4D X female	1.430 (0.735)	0.375 (0.686)	0.395 (0.735)	1.063 (0.702)	-1.148 (1.142)
constant	0.931 (0.587)	-0.084 (0.527)	0.067 (0.558)	0.538 (0.523)	-1.072 (0.871)
m2D:4D (F)	0.697 (0.441)	0.341 (0.439)	0.296 (0.478)	0.656 (0.468)	-0.032 (0.735)
<i>N</i>	3403	3325	3394	3350	1374
adj. <i>R</i> ²	0.033	0.011	0.000	0.015	0.001

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,

** $p < 0.005$. m2D:4D (F) presents the m2D:4D effect for women which were calculated by summing the coefficients of m2D:4D and m2D:4D X female variables.

Table S19. Regressions Analysis Based on the Mean of Both Hands: Exploratory Hypotheses 2 (Full Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	2.858 (3.621)	1.242 (2.285)	-0.559 (2.944)	2.529 (3.143)	-3.461 (3.687)
female	-0.359** (0.034)	0.217** (0.035)	0.063 (0.034)	-0.250** (0.035)	-0.090 (0.055)
m2D:4D ²	-1.350 (1.766)	-0.516 (1.039)	0.334 (1.387)	-1.132 (1.460)	1.863 (1.610)
constant	-1.305 (1.868)	-0.843 (1.261)	0.192 (1.573)	-1.262 (1.694)	1.636 (2.092)
<i>N</i>	3403	3325	3394	3350	1374
adj. <i>R</i> ²	0.032	0.011	0.000	0.015	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S20. Regressions Analysis Based on the Mean of Both Hands: Primary Hypotheses (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	-0.173 (0.406)	-0.106 (0.397)	0.268 (0.384)	0.346 (0.404)	-0.010 (0.663)
female	-0.358** (0.034)	0.219** (0.035)	0.063 (0.034)	-0.251** (0.035)	-0.086 (0.055)
constant	0.372 (0.405)	-0.013 (0.396)	-0.299 (0.385)	-0.212 (0.403)	0.049 (0.661)
<i>N</i>	3403	3325	3394	3350	1374
adj. <i>R</i> ²	0.032	0.011	0.001	0.015	0.000

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S21. Regressions Analysis Based on the Mean of Both Hands: Exploratory Hypotheses 1 (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	-0.782 (0.633)	-0.215 (0.596)	0.056 (0.571)	-0.259 (0.610)	0.844 (1.011)
female	-1.496 (0.838)	0.015 (0.798)	-0.334 (0.776)	-1.377 (0.797)	1.423 (1.333)
m2D:4D X female	1.141 (0.839)	0.204 (0.799)	0.397 (0.775)	1.129 (0.799)	-1.515 (1.335)
constant	0.978 (0.630)	0.0954 (0.594)	-0.088 (0.571)	0.390 (0.608)	-0.800 (1.008)
m2D:4D (F)	0.358 (0.552)	-0.0107 (0.531)	0.453 (0.523)	0.869 (0.516)	-0.670 (0.872)
<i>N</i>	3403	3325	3394	3350	1374
adj. <i>R</i> ²	0.032	0.011	0.000	0.015	0.001

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,

** $p < 0.005$. m2D:4D (F) presents the m2D:4D effect for women which were calculated by summing the coefficients of m2D:4D and m2D:4D X female variables.

Table S22. Regressions Analysis Based on the Mean of Both Hands: Exploratory Hypotheses 2 (Corrected Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	2.512 (7.175)	2.310 (4.472)	-4.371 (2.883)	8.925** (2.566)	-15.740 (12.630)
female	-0.358** (0.034)	0.219** (0.035)	0.063 (0.034)	-0.251** (0.035)	-0.0879 (0.055)
m2D:4D ²	-1.327 (3.588)	-1.193 (2.191)	2.293 (1.323)	-4.239** (1.177)	7.869 (6.292)
constant	-0.983 (3.592)	-1.233 (2.293)	2.042 (1.572)	-4.544** (1.405)	7.900 (6.343)
<i>N</i>	3403	3325	3394	3350	1374
adj. <i>R</i> ²	0.031	0.011	0.000	0.016	0.000

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S23. Regressions Analysis Based on the Mean of Both Hands: Primary Hypotheses (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	-0.343 (0.407)	-0.067 (0.419)	0.088 (0.411)	0.594 (0.414)	-0.061 (0.661)
female	-0.360** (0.034)	0.217** (0.035)	0.063 (0.035)	-0.251** (0.035)	-0.088 (0.055)
constant	0.546 (0.406)	-0.052 (0.419)	-0.121 (0.411)	-0.459 (0.414)	0.098 (0.660)
<i>N</i>	3360	3282	3351	3307	1357
adj. R^2	0.033	0.011	0.000	0.015	0.000

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table S24. Regressions Analysis Based on the Mean of Both Hands: Exploratory Hypotheses 1 (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	-1.084 (0.616)	-0.252 (0.633)	-0.092 (0.620)	-0.217 (0.643)	0.953 (0.974)
female	-1.702* (0.818)	-0.118 (0.845)	-0.263 (0.829)	-1.712* (0.836)	1.778 (1.315)
m2D:4D X female	1.343 (0.818)	0.336 (0.845)	0.326 (0.827)	1.462 (0.837)	-1.869 (1.316)
constant	1.284* (0.614)	0.132 (0.632)	0.059 (0.620)	0.349 (0.642)	-0.912 (0.972)
m2D:4D (F)	0.259 (0.539)	0.084 (0.559)	0.234 (0.548)	1.246* (0.536)	-0.916 (0.884)
<i>N</i>	3360	3282	3351	3307	1357
adj. <i>R</i> ²	0.033	0.011	0.000	0.016	0.001

OLS regressions; robust standard errors in parentheses; * $p < 0.05$,

** $p < 0.005$. m2D:4D (F) presents the m2D:4D effect for women which were calculated by summing the coefficients of m2D:4D and m2D:4D X female variables.

Table S25. Regressions Analysis Based on the Mean of Both Hands: Exploratory Hypotheses 2 (Restricted Sample)

	(1) Risk	(2) Altruism	(3) Positive Recip.	(4) Negative Recip.	(5) Trust
m2D:4D	39.870** (11.68)	17.530 (12.960)	-3.784 (12.401)	7.847 (11.320)	-30.800 (19.57)
female	-0.359** (0.034)	0.218** (0.035)	0.063 (0.035)	-0.250** (0.035)	-0.089 (0.055)
m2D:4D ²	-20.150** (5.854)	-8.819 (6.496)	1.940 (6.188)	-3.635 (5.690)	15.410 (9.829)
constant	-19.480** (5.827)	-8.820 (6.464)	1.807 (6.209)	-4.071 (5.627)	15.400 (9.733)
<i>N</i>	3360	3282	3351	3307	1357
adj. <i>R</i> ²	0.036	0.011	0.001	0.015	0.001

OLS regressions; robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

References

- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4):1645–1692.