

Replication instructions for:

Homo cooperans: Understanding the nature of human cooperation

Peter Andre, Teodora Boneva, Felix Chopra, Armin Falk

May 2026

Section I describes the steps required to replicate our analysis, including a description of how to obtain secondary data.

Section II contains a description of variables. We refer to the published paper for a detailed description of the data and variables: science.org/doi/10.1126/science.aec9483

Section III contains a description of methods. We refer to the published paper for a detailed description of the methods and analyzes.

I. REPLICATION INSTRUCTIONS

STEP 1: External data

This section contains instructions on how to obtain the external data that is not included in the replication package.

We cannot share the Gallup World Poll data on donated money, volunteered time, and helped a stranger because they are part of the proprietary Gallup World Poll (items WP108, WP109, WP110). Users with access to the Gallup World Poll can match these data to our sample using the identifier `wpid`. These data are only relevant for the Supplementary Table S.7, column (7). They play only a peripheral role in the paper. We have the right to share Gallup World Poll data on core demographics such as income, education, or age.

External data should be placed in the corresponding folders of the data folder. The folder structure is replicated below:

```
replication
- code
- out
- data
  - world_raw.dta      (file included)
  - gcd_raw.dta        (file included)
  - currency_calibration.csv      (file included)
  - ancestral_distance
    - Spolaore and Wacziarg 2017
      - bilateral.dta
    - Spolaore and Wacziarg 2009
      - newgendist.dta
  - cultural_distance
    - matrix_all_beliefs.csv
    - matrix_all_group_memberships.csv
    - matrix_all_law.csv
    - matrix_all_misc.csv
    - matrix_all_political.csv
    - matrix_all_sexuality.csv
    - matrix_all_social_relations.csv
    - matrix_all.csv
  - ethnographic_atlas_distance
    - Ancestral_Characteristics_Database_Extended_Eastern_Europe_Siberia_and_WES_Version_1_1.dta
  - external_country
    - Cohn2019
      - behavioral_data (stata data file).dta
    - covid_who
      - vaccination-data.csv
    - QOG
      - qogcs.dta
```

- Romano2021
 - 41467_2021_24787_MOESM4_ESM.xlsx
- Schulz2019
 - alldata.dta
- WDI
 - API_NY.GDP.PCAP.PP.KD_DS2_en_csv_v2_4772808.csv
 - wdi_pop_share_14.csv
 - wdi_pop_total.csv

For each external file, we describe the process of how to obtain the data that is to be placed in this folder below:

bilateral.dta, newgendist.dta

- Download the data from Enrico Spolaore's website:
http://www.anderson.ucla.edu/faculty_pages/romain.wacziarg/downloads/2009_genetic_distance.zip
- Download the data from Enrico Spolaore's website:
http://www.anderson.ucla.edu/faculty_pages/romain.wacziarg/downloads/2017_newgendist.zip

matrix_all_beliefs.csv

matrix_all_group_memberships.csv

matrix_all_law.csv

matrix_all_misc.csv

matrix_all_political.csv

matrix_all_sexuality.csv

matrix_all_social_relations.csv

matrix_all.csv

- All of the above files can be obtained here: <http://culturaldistance.com/>
- Specification: years 1994-2014, years group: combine, countries: all countries

Ancestral_Characteristics_Database_Extended_Eastern_Europe_Siberia_and_WES_Version_1_1.dta

- The data can be obtained from Nathan Nunn's website:
<https://nathannunn.arts.ubc.ca/data/>
- Select the file "Country-level dataset, EE, Siberia, WES extension (version 1.1)"

behavioral data (stata data file).dta

- The dataset is included in the replication package of Cohn et al. (2019):
<https://dataverse.harvard.edu/dataverse/honesty/>

vaccination-data.csv

- The data was obtained from the WHO on October 24, 2024, using this link:
<https://data.who.int/dashboards/covid19/vaccines>
- The file is no longer accessible from this website, but can be obtained from the following link:
<https://srhdpeuwpubsa.blob.core.windows.net/whdh/COVID/vaccination-data.csv>

qogcs.dta

- The QOG standard data set is available from the QoG Institute at the University of Gothenburg.
- We used the 2021 Standard dataset – Cross-Section, which can be downloaded in the data archive:
<https://www.gu.se/en/quality-government/qog-data/data-downloads/data-archive>
- Download link: https://www.qogdata.pol.gu.se/dataarchive/qog_std_cs_jan21.dta
- Rename the dta file to “qogcs.dta”

41467_2021_24787_MOESM4_ESM.xlsx

- Download the “Source Data” here:
<https://www.nature.com/articles/s41467-021-24787-1#Sec20>

alldata.dta

- The file is part of the replication data of Schulz et al. (2019):
<https://datadryad.org/dataset/doi:10.5061/dryad.2rbnzs7hs>

wdi_pop_share_14.csv

- Go to: <https://databank.worldbank.org/source/population-estimates-and-projections>
- Select: “All countries”, all years since 1960, and the time series “Population ages 0-14 (% of total population)” (SP.POP.0014.TO.ZS)
- Download the file and rename it.

wdi_pop_total.csv

- Go to: <https://databank.worldbank.org/source/population-estimates-and-projections>
- Select: “All countries”, all years since 1960, and the time series “Population, total” (SP.POP.TOTL)
- Download the file and rename it.

API_NY.GDP.PCAP.PP.KD_DS2_en_csv_v2_4772808.csv

- Go to: <https://databank.worldbank.org/source/world-development-indicators>
- Select: “All countries”, all years since 1960, and the time series “GDP per capita, PPP (constant 2017 international \$)” (NY.GDP.PCAP.PP.KD)

STEP 2: R

Version: R 4.5.2 and RStudio 2025.09.2. The code was last executed on a 32-GB laptop running macOS Sequoia 15.6.1 with an Apple M2 Pro processor.

Instructions:

- Navigate to the replication folder.
- Open R project `homo_cooperans.Rproj`.
- Three packages are required: `rjson`, `groundhog`, and `Rtools`. Install them if necessary.

- Open and run `code/main_R.R`. If you use Windows, you may need to manually change line 7 (see also comment in line 4).

The following things should happen:

- The function `path_to` is loaded. This function returns relative file paths, drawing on the list in `code/specs/file_paths.json`.
- The package management package `groundhog` is used. All other required packages will automatically be installed. The package version at date 2025-12-01 will be installed. This ensures replicability. For more information, please see the documentation of the `groundhog` package.
- A number of data cleaning scripts are executed. The folder `out/data_out/` contains the cleaned data.
- The following figures are produced.
 - `fig_2.pdf`
 - `fig_3.pdf`
 - `fig_S1.pdf`

STEP 3: Python

Version: Python 3.13.11. The code was last executed on a 32-GB laptop running macOS 15.5 with an Apple M1 Pro processor.

The data cleaning for the dyadic regressions is performed in Python.

Instructions:

- Prerequisite is a working installation of Python 3 and miniconda or anaconda.
- Open a terminal and navigate to `code/cleaning_python/`
- Open `environment.yml` and modify line 23 to match your repository structure.
- Run `conda env create -f environmen.yml` in your terminal. This will setup Python and install all necessary packages. On Windows, this may require using a Powershell.
- Run `conda activate cooperation_replication` in your terminal to activate the environment.
- Run `python main.py` in the terminal to perform all cleaning steps.

STEP 4: STATA

Version: Stata 19.5. The code was last executed on a 32-GB laptop running macOS 15.5 with an Apple M1 Pro processor.

The next step is to run the part of the data analysis in Stata.

Instructions:

- First open the file `setup.do` located in the `code/` folder and adjust the path to the root of the replication package on lines 13-15 to match the replicator's directory structure. You may need to uncomment lines 26-34 to install necessary packages.
- Run the `setup.do` file in Stata.

- Run the `main.do` file in Stata.

The output from the do-files is located in `out/tables/` and `out/figures/`. The following files are produced:

- `fig_1.pdf`
- `fig_S2.pdf`
- `tab_1.tex`
- `tab_S1.tex`
- `tab_S2.tex`
- `tab_S3.tex`
- `tab_S4.tex`
- `tab_S6.tex`
- `tab_S7.tex`

II. DESCRIPTION OF VARIABLES AND METHODS

The raw data for the Global Cooperation Dataset was collected as part of the Gallup World Poll 2021/2022 in 125 countries and includes 129,902 respondents. For each respondent, we observe basic demographic characteristics (e.g., age, gender, education, urban/rural, household income). We also collect information about preferences, beliefs and norms. The primary variable of interest is the decision in a two-person cooperation game. The survey also included an information provision treatment that was administered to half of the sample before the cooperation decision.

We provide a description of the variables contained in the raw data files contained in this replication package below. In addition, we provide a description of the variables derived from our primary survey data.

world_raw

The unit of observation in this dataset is a country.

- `iso2` - the ISO-2 code of the country
- `iso3` - the ISO-3 code
- `country` - the country name
- `continent` - the continent to which the country belongs (5 categories)
- `world_region` - indicator for different world regions (8 categories)
- `world_region_names` - the name of the world region encoded in `world_region`
- `western_europe` - dummy for the indicated level of the `world_region` variable
- `eastern_europe` - dummy for the indicated level of the `world_region` variable
- `south_and_east_asia` - dummy for the indicated level of the `world_region` variable
- `south_america` - dummy for the indicated level of the `world_region` variable
- `north_africa_and_middle_east` - dummy for the indicated level of the `world_region` variable
- `sub_saharan_africa` - dummy for the indicated level of the `world_region` variable
- `north_america` - dummy for the indicated level of the `world_region` variable
- `australia_and_new_zealand` - dummy for the indicated level of the `world_region` variable

gcd_raw

The unit of observation in this dataset is a World Poll respondent. The dataset covers respondents from 125 different countries.

- wpid - unique identifier for a respondent
- wgt - weight
- COUNTRY_ISO3 - country code
- countrynew - country name
- INCOME_4 - per capita annual income in international dollars
- INCOME_5 - per capita income quintiles
- WP1219 - gender of the respondent
- WP1220 - age
- WP14 - urban/rural indicators
- WP3117 - education level
- WP22198 - willingness to take risks
- WP22199 - measure of altruism
- WP22200 - measure of patience
- WP22203 - dummy for whether people are willing to give up 1% of their income to fight climate change
- WP22204 - dummy for whether people are willing to give less than 1% of their income to fight climate change
- WP22206 - measure of social norms regarding climate action
- WP22208 - consent to participate in the cooperation module
- RNDM_CLIMATE - treatment dummy (1 = received information)
- RNDM_CLIMATE_ALT - treatment dummy (1 = received information) for countries in which the cooperation decision was hypothetical
- WP22209 - decision in the cooperation game (2 = cooperate)
- WP22209_ALT - hypothetical version of WP22209
- WP22210 - belief about how many people in one's country decided to cooperate
- WP22210_ALT - hypothetical version of WP22210

currency_calibration

The unit of observation in this dataset is a country. For each country and each survey question that involved values denominated in local currency units, we report the numerical value of the currency unit that was used in the survey.

- region - region that the country belongs to
- country - country name
- currencycode - abbreviation of the local currency unit (e.g., USD in the USA).
- countrycodeiso3 - ISO-3 country code
- countrycodeworldbank - country code used by the World Bank
- q2 - the numerical value used in question 2
- q3_1today - analogous to q2
- q3_2in12months - analogous to q2
- q15 - analogous to q2
- q16_1option1 - analogous to q2
- q16_2option2 - analogous to q2

- q16_3donation - analogous to q2
- q17 - analogous to q2

For each of the main variables derived from our survey data, we provide the survey question, the response options, the coding and how we standardized responses. For a description of all variables used in the paper, we refer the interested reader to the Materials and Methods section in the published paper and its Supplementary Materials file.

Altruism

Imagine the following situation: Today you unexpectedly receive [\$1,600]. How much of this amount would you donate to a good cause?

Responses: 0-[\$1,600], (Refused)

Coding: Continuous, rescaled to 0-1. (Refused) is coded as missing data.

Standardization: We standardize the altruism measure, i.e., deduct the (weighted) global mean and divide by the (weighted) global standard deviation.

Patience

For the next question, please assume that future prices are the same as today's prices.

Would you rather receive [\$160] today or [\$246] in 12 months?

Responses: [\$160] today, [\$246] in 12 months, (Refused)

Coding: Binary dummy for [\$246] in 12 months. (Refused) is coded as missing data.

Standardization: We standardize the patience measure, i.e., deduct the (weighted) global mean and divide by the (weighted) global standard deviation.

Risk-taking

Please tell me, in general, how willing or unwilling are you to take risks? Please use a scale from 0 to 10, where 0 means "completely unwilling to take risks" and 10 means you are "very willing to take risks." You can use any number between 0 and 10.

Responses: 0-10, (Refused)

Coding: Continuous, 0-10. (Refused) is coded as missing data.

Standardization: We standardize the risk-taking measure, i.e., deduct the (weighted) global mean and divide by the (weighted) global standard deviation.

Preference for common good

Would you be willing to contribute 1% of your household income every month to fight global warming? This would mean that you would contribute [\$1] for every [\$100] of this income.

Responses: Yes, No, (DK), (Refused)

Coding: Binary dummy for Yes. (DK) and (Refused) are coded as missing data.

Injunctive social norms

Do you think that people in [the United States] should try to fight global warming?

Responses: Yes, No, (DK), (Refused)

Coding: Binary dummy for Yes. (DK) and (Refused) are coded as missing data.

Cooperation

Here is the decision you need to make. If you are randomly selected, you will be matched with another selected respondent from [the United States]. Both you and the other respondent can choose between two options.

- Option 1: You receive [100 US dollars] for yourself.
- Option 2: You receive [70 US dollars] for yourself. If BOTH you and the other respondent choose the smaller amount of [70 US dollars], a donation of [400 US dollars] will be made to fight global warming.

If you are randomly selected, this is a real decision. What would you choose?

- Option 1 – to receive [100 US dollars] for yourself, and no donation will be made.
- Option 2 – to receive [70 US dollars] for yourself, and if the other respondent also chooses this option, a donation of [400 US dollars] will be made to fight global warming.

Coding: Binary dummy for Option 1. (DK) and (Refused) are coded as missing data.

Beliefs

We are asking this question to 100 other respondents in [the United States]. How many do you think choose to receive only [70 US dollars]?

Response: 0-100, (DK), (Refused)

Coding: 0-100, (DK) and (Refused) are coded as missing data.

III. DESCRIPTION OF METHODS

This section provides details about the main analysis. In a first step, we regress the decision to cooperate on different respondent characteristics (Figure 1). In a second step, we estimate bivariate regressions of the cooperation decision on different respondent characteristics separately for each country and plot the distribution of beta coefficient (Figure 2). We then report the global (weighted) share of respondents that cooperate (Figure 3a) and plot national cooperation rates against the expected cooperation rate (Figure 3b). Finally, we estimate the treatment effects of our information experiment on cooperation decisions and beliefs (Table 1). For details about the exact empirical specifications, we refer to the documented code in this replication package.