

Replication files: Subjective Models of the Macroeconomy

This repository contains the replication scripts and data of the analyses of the paper:

Andre, P., Pizzinelli, C., Roth, C., & Wohlfart, J. (2021). **Subjective Models of the Macroeconomy: Evidence From Experts and Representative Samples.** *Review of Economic Studies*.

Please contact Peter Andre (or the other authors) in case you have any questions or encounter problems.

Data and Code Availability Statement

All **data** are made available, except:

1. Data from the ifo World Economic Survey (WES). Access to these data is provided by the EBDC in ifo institute in Munich, Germany.
2. The data from the expert surveys are available only in anonymized form (instead of its original pseudonymized form).

All **code** is made available. Parts of the code only run with access to the WES data (or the pseudonymized expert survey data). This means that some figures and tables can only be generated with access to these data. The respective parts of the code are out-commented in the scripts.

Working with the WES data is only possible locally at the EBDC in Munich or remotely with the support of the EBDC. For security reasons, the computers of the EBDC in Munich do not have access to the internet. A key challenge is to install the required packages and all their dependencies. The `groundhog` approach (see below) will not work. The packages need to be installed manually. We last accessed the WES data in May 2021. See <https://www.ifo.de/EBDC> for contact information.

For more information on the WES data see: Boumans, Dorine and Johanna Garnitz, "Ifo World Economic Survey Database: An International Economic Expert Survey," *Jahrbücher für Nationalökonomie und Statistik*, 2017, 237 (1), 71–80.

WES results All results obtained with the WES data (which cannot be reproduced without access to these data) are reported in the folders `report/Figures_WES` and `report/Tables_WES`.

License

Data and code made available via this replication package are licensed under a [Creative Commons Attribution–NonCommercial 4.0 International \(CC BY-NC 4.0\)](#) license.

Version information

The code was programmed on macOS (11.5.1) in RStudio (2021.09.0), using R 4.4.1.

The package management package `groundhog` is used. Once this package is installed, all other required packages will automatically be installed. The package version at date 2021-10-01 will be installed. This ensures replicability. For more information, please see the documentation of the `groundhog` package.

Your goals? How to start?

- **Replicate the full project.** Read section "Main script".
- **Replicate a specific analysis.** Read section "Main script", compile the preparatory section plus the data management, execute the script of interest or open it (see section "Folder structure") and execute it manually.
- **Run new analyses with the data.** Read section "Main script", compile the preparatory section plus the data management. Read section "Data and codebook". Do not work with the functions in `functions_reg`. They are not well-documented and it is probably easier, safer, and faster to write your own code. Contact Peter Andre (or the other authors) in case anything is unclear.

Overview of all analysis scripts

The `xlsx` file `analysis_directory.xlsx` provides an overview of all analysis scripts and exhibits and documents which scripts produce which exhibit. It also lists exhibits that are not produced by code, and it lists analysis scripts that do not produce an exhibit but whose outputs are mentioned in the main text of the paper.

Main script

The main script `code/main.R` executes the full analysis and generates the figures and tables of the paper.

Note: The analysis was programmed and run on macOS. In order to avoid encoding problems for the text data collected in waves 3-5 on Windows, please replace `en_US.UTF-8` in line 14 with `English`.

First, the data of waves 1 and 2 is cleaned and prepared for analysis. Each table and figure is then derived by an *independent* script. The process is analogous for the data from waves 3 to 5. (See table A.3 of the paper for an overview of all waves of data collection.)

Important: At the beginning of each R session, the PREPARATION section of the main script needs to be run in order to set the correct working directory and load the `file_path` function (more details below). The working directory is set automatically to the project folder. No changes to the script are required.

Folder structure

- `appendix_I/` : Contains a separate replication folder for the analyses of Appendix I.
- `code/` : Contains all scripts.
 - `code/data_analysis/` : Contains the code of all analyses, tables, figures. Analyses for data collection waves 3-5 are in a separate subfolder `associations/`.
 - `code/data_cleaning/` : Contains the code of the data management. Data cleaning for data collection waves 3-5 is in a separate subfolder

- associations/ .
 - code/functions/ : Auxiliary functions.
 - code/misc/ : Miscellaneous scripts.
 - code/specs/ : Specifications or definitions (e.g. the benchmarks derived from the literature, list of file paths) that are used throughout the analysis and data management.
- data/ : Contains the *raw* data (see "Data and codebook").
- out/ : Contains all outputs of the data management and data analysis.
 - out/data/ : Cleaned data, potentially temporary data manipulation output (see "Data and codebook").
 - out/figures/ : Figures.
 - out/models/ : R regression models (rds), temporary storage for later reference.
 - out/rmd : Output for R markdown scripts.
 - out/tables : Latex tables.
- codebook/ : Contains tabular descriptions of the variables and coding of the final data sets.
- qualtrics/ : Contains the *qsf* files of all Qualtrics surveys (except for Wave 2, expert survey, which was programmed with the WES programming survey).
- report/ : Creates a report with all figures and tables of the paper (except for Appendix I, see "Appendix I"). The folder also includes additional tables and figures that cannot be produced without access to the WES data or have been produced manually.

Numbers in text and additional scripts (not part of the main script)

The following analyses scripts are not executed in the main script but can be executed manually and separately.

- ``data_analysis/calc_numbers_in_text'`: This script re-calculates all numbers reported in the paper that cannot be found in and read from the exhibits of the paper.
- `data_analysis/calc_alpha` : Derives the estimates for `mu_alpha` and `sigma_alpha` used in Appendix I.
- `data_analysis/calc_sample_size` : Loads all waves of data collection after the cleaning in order to determine the sample size for Table A.3.
- `data_analysis/associations/cal_coding_correlation` : Derives different measures of inter-coder-reliability for the manual coding of open-text responses.
- `data_analysis/reg_tab_politics_delta_SUR.do` : Runs the F-test reported in script `reg_tab_politics` .

Data and codebook

Codebooks for all used (or otherwise important) variables of all cleaned data sets are available in the folder `codebook/` .

A list of cleaned data sets: All cleaned data are in the `out/data/` folder. Typically, each file is available as `csv` or `rds` (an R-internal storage format) and as Stata `dta` file. The variables are not labeled. But a short description of the most important variables is provided in the `codebook/` folder.

- `associations_text` : Wave 3. The `associations` data + indicators for the manually-coded mechanism associations. The `rds` file contains a list of lists. One list for each vignette. Each vignette has a sublist with the cleaned survey data and objects of quantative text analysis (`quanteda` package: DFMs). The `stata` version contains the pooled cleaned survey data of all vignettes.
- `associations` : Wave 3.
- `data_delta` : Waves 1 and 2. Each observation in this data set is a response to a vignette. Respondents received two vignettes. Thus, for each respondent, there are two observations (exception: a few experts who received four vignettes).
- `data_wide` : Waves 1 and 2. Each observation in this data set is a respondent. This data set is used only for a few auxiliary respondent-level analyses.
- `experiences_text` : Wave 5. Analogous to `associations_text` .
- `experiences` : Wave 5.
- `prime_text` : Wave 4. Analogous to `associations_text` .
- `prime` : Wave 4.

All other files are auxiliary or intermediate files that can be ignored.

A list of raw data sets: All raw data collected by the authors are in the `data/` folder.

- `associations_coding/` : manually-coded open-text data of waves 3-4.
- `associations_distribution/launch.csv` : scholars invited to wave 3 of the expert survey.
- `associations_expert.csv` : wave 3, expert survey, exported data from Qualtrics.
- `associations_household.csv` : wave 3, household survey, exported data from Qualtrics.
- `experiences_coding/` : manually-coded open-text data of wave 5.
- `experiences.csv` : wave 5, household survey, exported data from Qualtrics.
- `expert_data_demo.csv/` : characteristics of experts who participated in wave 1 of the expert survey.
- `expert_data` : wave 1, expert survey, exported data from Qualtrics.
- `gate.csv` : wave 1, expert survey, exported data from Qualtrics, (experts started the survey here and were then forwarded to another Qualtrics survey)
- `hh_data_2.csv` : wave 2, household survey, exported data from Qualtrics.
- `hh_data.csv` : wave 1, household survey, exported data from Qualtrics.
- `imf_data.csv` : wave 1, expert survey, exported data from Qualtrics (there was a separate Qualtrics survey for participants from the IMF).
- `prime.csv` : wave 4, household survey, exported data from Qualtrics.

Special features

path_to All relative file paths are stored in `misc/specs/file_paths.json` . The function `path_to()` allows to access these paths with a keyword (see the json file). It is used throughout the code.

Regressions and stargazer wrapper All regressions are estimated with the `lm` method. Robust standard errors or clustered standard errors are derived to yield the same output as in Stata. Regression tables are created with the `stargazer` package and are further manipulated and polished by adjusting the latex code with the help R text manipulations. Both the estimation and the generation of the latex code take place in wrapper functions (`code/functions/functions_reg`).

Appendix I

The analyses of Appendix I can be replicated with the separate replication package that is available in the `appendix_I/` folder.