Read Me File for: The Gender Gap in Household Bargaining Power: A Revealed-Preference Approach

Ran Gu, Cameron Peng and Weilong Zhang

Data Availability Statement

The main sources of confidential data used in the paper are:

- **HILDA GENERAL RELEASE 18 (Waves 1-18)**: The majority of the analysis presented in the paper comes from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Survey provides individual-longitudinal level data. Data from the HILDA Survey is available to researchers living in Australia or overseas. Data may only be accessed through the National Centre for Longitudinal Data Dataverse.
  - For more information on data access, please see. https://dataverse.ada.edu.au/dataverse/hilda
- **SOEP-Core V33.1:** The data are provided by the German Institute for Economic Research (DIW Berlin). Because the SOEP data are subject to data protection legislation, researchers have to sign a contract with the SOEP providers first (free of charge). Data access is usually established within three to five days.
  - You can find more information about the SOEP and data access here: https://www.diw.de/en/diw_01.c.601584.en/data_access.html
  - If you have questions, please contact the SOEPhotline (email: soepmail@diw.de , phone: +49 30 89789-292).
- **RAND HRS:** You will be able to see the data descriptions and documentation without an account, but you will need to register with HRS to access the data. You will be able to see the data descriptions and documentation without an account here, but you will need to register with HRS to access the data.
  - You can find more information about the RAND HRS and data access here: https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200005
  - For more information, email RANDHRSHelp@rand.org.

Description of the replication code

Part 1: description of the Stata Codes (do files)

Because of the confidentiality issues discussed above, this replication package does not contain the original data from HILDA, GSOEP, or HRS. However, the following is a summary of the data structure needed to run the replication code we provide, conditional on obtaining access to the data as discussed above:

All analyses were implemented in Stata 17.

1.1 Steps when using HILDA dataset

1. Step 1: Transfer the necessary data files (CNEF_Long_r180c.dta, Rperson_`wave'180c.dta, Household_`wave'180c.dta, Master_r180c) from GENERAL RELEASE 18 (Waves 1-18) to data input directory.
2. Step 2: Configure the file path in "*master_hilda.do*." For our example, it was set to " **~/Dropbox/HILDA_Financial_literacy/Replication file/Stata_part/hilda/**". Update the bold part with the location of your replication files.
3. Step 3: Execute "master_hilda.do." This Stata script generates the figures and tables in the paper and the sample for estimation in Part II.

Input data files (from GENERAL RELEASE 18 (Waves 1-18))

- CNEF_Long_r180c.dta: A long HILDA dataset for the Cross-National Equivalence File, a project started by Cornell University which creates comparable longitudinal social and economic population datasets for other countries, such as the USA, Germany, the UK, Switzerland, Russia, Korea and Canada.
- Rperson_`wave'180c.dta: Responding person (individual interview completed). Data from the New Person Questionnaire, the Continuing Person Questionnaire and the Self Completion Questionnaire and associated derived and history variables.
- Household_ `wave'180c.dta: Information collected or aggregated for the household as a whole. Data from the Household Form and Household Questionnaire and derived variables relating to the household.
- Master_r180c.dta: A longitudinal sample description file which contains identifiers, household status and interview status at each wave.

Output files:

- main_sample2.csv: Data files for structural estimation in the next phase.
- log_summary.log: Provide information for Table 1 and A1.
- log_summary_riska.log: Provide information for Table A1.
- log_variable_table.log: Provide information for Table 2, 12, 13.
- log_variable_second.log: Provide information for Table 12 and A5.
- log_basicinfo.log: Provide information for Table 4.
- reg_skRk.tex: Table 3
- TR_role_m.tex: Table A2
- norm_change.tex: Table A3
- RDN.tex: Table A4
- role_agreed.eps: Figure 1

1.2 Steps when using GSOEP dataset

1. Step 1: Transfer the necessary data files (cognit.dta, pequiv.dta, pl.dta, hl.dta) from SOEP-Core V33.1 to data input directory.
2. Step 2: Configure the file path in "*master_soep.do*." For our example, it was set to " **~/Dropbox/HILDA_Financial_literacy/Replication file/Stata_part/International comparison/Germany_SOEP/**". Update the bold part with the location of your replication files.
3. Step 3: Execute "master_soep.do." This Stata script generates the sample for estimation in Part II and tabulates households by spousal risk preferences.

Input data files (from SOEP-Core V33.1)

- *cognit.dta*: the file provided variables related cognitive achievement potentials.
- *pequiv.dta*: this file contains individual-level and household-level data to ensure a harmonized cross-country comparison. It is keyed on PID (Person ID) and SYEAR (Survey Year)
- *pl.dta*: this file contains individual-level data for respondents by answering the annual individual questionnaire. It is keyed on PID (Person ID) and SYEAR (Survey Year)
- *hl.dta*: this file contains household-level data from the annual household questionnaire. It is keyed on HID (Household ID) and SYEAR (Survey Year)

Output files:

- TR_risk_soep.tex: Table A7 in the appendix.
- sample_GSOEP.csv: Data files for structural estimation in the next phase.

1.3 Steps when using HRS dataset

4. Step 1: Transfer the necessary data files (randhrs1992_2018v1.dta, h06f3a.dta, h08f3a.dta, h14f2b.dta, h16f2c.dta) from RAND HRS to data input directory.
5. Step 2: Configure the file path in "*master_hrs.do*." For our example, it was set to " **~/Dropbox/HILDA_Financial_literacy/Replication file/Stata_part/International comparison/hrs/**". Update the bold part with the location of your replication files.
6. Step 3: Execute "master_hrs.do." This Stata script generates the sample for estimation in Part II and tabulates households by spousal risk preferences.

Input data files (from RAND HRS)

- *randhrs1992_2018v1.dta*: the RAND HRS Longitudinal File includes fourteen waves of Core Interview data across sixteen survey years (1992, 1993, 1994, 1995, and biennially 1996-2018) as well as Exit Interview data collected across thirteen survey years (1994, 1995, and biennially 1996-2016). All variables are named and derived consistently across survey years, and any cross-wave differences are documented.
- *h06f3a.dta*: the RAND HRS 2006 Fat File is based on HRS 2006 Core (Final V3.0, December 2014) data, and the Tracker 2006 (V1.0, October 2007) file.
- *h08f3a.dta*: the RAND HRS 2008 Fat File is based on HRS 2008 Core (Final V3.0, December 2014) data, and the Tracker 2008 (Final V1.0, October 2010) file.
- *h14f2b.dta*: the RAND HRS 2014 Fat File is based on HRS 2014 Core (Final V2.0, December 2017) data, and the Tracker 2014 (Final V1.0, July 2017) file.
- *h16f2c.dta*: the RAND HRS 2016 Fat File is based on HRS 2016 Core (Final V2.0, December 2019) data, and the Tracker 2020 (Early V2.0, December 2022) file.

Output files:

- TR_risk_hrs.tex: Table A11 in the appendix.
- log_analysis_hrs_head.log: Provide information for Table A12 in the appendix.
- sample_HRS.csv: Data files for structural estimation in the next phase.

Part 2: description of the matlab Codes (m files)

The code requires Matlab R2020b or a later version, and the Optimization Toolbox must be installed to utilize the "particleswarm" global optimizer.
We run the Matlab component on a high-performance cluster equipped with 18 CPUs (36 threads).
Note that the structural estimation uses the data file generated by the Stata part above (main_sample2.csv, sample_GSOEP.csv, sample_HRS.csv).

2.1 Steps when using HILDA dataset:

1. Execute "main_program.m" located in "\Replication file\Matlab_part\HILDA_Baseline". This file estimates the baseline model using HILDA data.
2. Execute "display_result.m" in "\Replication file\Matlab_part\HILDA_Baseline" to generate counterfactual results based on the baseline model.
3. Execute "Main_server.m" in "\Replication file\Matlab_part\HILDA_Intermediate" to estimate the intermediate model using HILDA data.
4. Execute "Main_server.m" in "\Replication file\Matlab_part\HILDA_Full" to estimate the full model using HILDA data.

Additional functional files called by the main files:

- "counterfactual_experiment.m": Simulates stock market participation choices for a given parameter set.
- "individual_likelihood.m": Calculates individual likelihood value based on observed stock market participation decisions and risky asset holdings.
- "individual_simulation.m": Simulates individuals' stock market participation decisions and equity holding choices for a given parameter set.
- "likelihood.m": Aggregates individual likelihood values.
- "simulation_fun.m": Displays distributions for simulated stock market participation decisions and equity holdings.
- "simulation_fun3.m": Calculates simulated distributions (equity holdings, stock market participation) in counterfactual cases.
- "fliplegend.m": Reorders legend on a figure.

2.2 Steps when using GSOEP dataset

- Execute "main_program.m" located in "\Replication file\Matlab_part\GSOEP". This file

estimates the baseline model using GSOEP data.

2.3 Steps when using HRS dataset

- Execute "main_program.m" located in "\Replication file\Matlab_part\HRS". This file estimates the baseline model using HRS data.

List of Tables and the Code used to produce them:

Table #	Program	Output file
Table 1	Stata_part\hilda\variable_generator_summary.do	log_summary.log
Table 2	Stata_part\hilda\variable_generator_table.do	log_variable_table.log
Table 3	Stata_part\hilda\variable_generator_table_master.do	reg_skRk.tex
Table 4	Stata_part\hilda\variable_generator_basicinfo.do	log_basicinfo
Table 5	Matlab_part\HILDA_Baseline\main_program.m	Command window
Table 6	Matlab_part\HILDA_Baseline\main_program.m	Command window
Table 7	Matlab_part\HILDA_Baseline\main_program.m	Command window
Table 8	Matlab_part\HILDA_Baseline\simulation_fun.m	Command window
Table 9	Matlab_part\HILDA_Baseline\display_result.m	Command window
Table 10	Matlab_part\HILDA_Baseline\ main_program.m Matlab_part\HILDA_Intermediate\ main_program.m Matlab_part\HILDA_Full\ main_program.m	Command window
Table 11	Matlab_part\HILDA_Baseline\simulation_fun.m Matlab_part\HILDA_Intermediate\simulation_fun.m Matlab_part\HILDA_Full\simulation_fun.m	Command window
Table 12	Stata_part\hilda\variable_generator_table.do Stata_part\hilda\variable_generator_second.do	log_variable_table.log log_variable_second.log
Table 13	Stata_part\hilda\variable_generator_table.do	log_variable_table.log
Table 14	Matlab_part\HILDA_Baseline\display_result.m Matlab_part\GSOEP\main_program.m Matlab_part\HRS\ main_program.m	Command window
Table A1	Stata_part\hilda\variable_generator_summary.do Stata_part\hilda\variable_generator_summary_riska.do	log_summary.log log_summary_riska.log
Table A2	Stata_part\hilda\variable_generator_table.do	TR_role_m.tex
Table A3	Stata_part\hilda\variable_generator_table.do	norm_change.tex
Table A4	Stata_part\hilda\variable_generator_table.do	RDN.tex
Table A5	Stata_part\hilda\variable_generator_second.do	log_variable_second.log
Table A6	Matlab_part\HILDA_Baseline\main_program.m	Command window
Table A7	Stata_part\International comparison\Germany_SOEP\analysis_soep.do	TR_risk_soep.tex
Table A8	Matlab_part\GSOEP\main_program.m	Command window
Table A9	Matlab_part\GSOEP\main_program.m	Command window
Table A10	Matlab_part\GSOEP\ simulation_fun.m	Command window
Table A11	Stata_part\International comparison\hrs\analysis_hrs.do	TR_risk_hrs.tex
Table A12	Stata_part\International comparison\hrs\analysis_hrs_head.do	log_analysis_hrs_head.log
Table A13	Matlab_part\HRS\main_program.m	Command window
Table A14	Matlab_part\HRS\main_program.m	Command window
Table A15	Matlab_part\HRS\ simulation_fun.m	Command window

List of Figures and the Code used to produce them:

Figure #	Program	Output file
Figure 1	Stata_part\hilda\variable_generator_table.do	role_agreed.eps
Figure 2	Matlab_part\HILDA_Baseline\main_program.m	part_cost.eps
Figure 3	Matlab_part\HILDA_Baseline\simulation_fun.m	husbands_gamma.eps wives_gamma.eps
Figure 4	Matlab_part\HILDA_Baseline\simulation_fun.m	asset_fit.eps
Figure 5	Matlab_part\HILDA_Baseline\simulation_fun.m Matlab_part\HILDA_Baseline\simulation_fun3.m	bargaining_weight.eps bargaining_weight_3.eps bargaining_weight_5.eps