

# Replication files for "Liquidity Constraints in the U.S. Housing Market"

Corina Boar, Denis Gorea and Virgiliu Midrigan

## 1. Overview

This replication package is structured as follows:

- the subfolder `model_replication` has all the codes that are necessary to obtain the results of the model.
- the subfolder `data_replication` has all the codes and data that were used to generate the data moments described in the paper. Please note that the PSID data and codes used to produce results in our paper are available in a separate repository at <https://doi.org/10.3886/E145381V1> (more details on this below).

## 2. Model

This section provides details on how to reproduce the model results reported in the figures and tables in the paper.

All model output is generated in MATLAB. The user should first install the CompEcon Toolbox, available at <https://pfackler.wordpress.ncsu.edu/compecon-download/> and use the demo files to verify that the .mex files are compiled correctly. In addition to the CompEcon Toolbox, we made use of the `kronm.m` routine written by Matthias Kredler. That file is included in the folders that use it.

- Table 2, Panel A, first 4 moments describing the income process  
- run `/Income Process/start.m`
- Table 2, Panel A, remaining moments (all other than income)  
Table 3  
Table 4  
- run `/Main nu 3/start.m`
- Table 2, Panel B: see the parameters at the beginning of `start.m`  
- Notice that although the model is quarterly, we report annualized values in the table, whenever appropriate.  
- Note that average annual income in our model is 5.4614 and the average income of the bottom 80% of households in the SCF in 2001 was equal to 70,430 2016 USD, so a unit in our model corresponds to 12,896 2016 USD. The fixed cost of mortgage refinancing (0.102974 in the model) is therefore equal to  $0.102974 * 12,896 = 1,328$ , which we round to 1,330 in the text.
- Table 5  
- Column labeled "nu = 1/3": run `/Main nu 3/start.m` and change the switch indicator to 'table5' on line 248 of `start.m`  
- Column labeled "nu = 1/10": run `/Main nu 10/start.m` and change the switch indicator to 'table5' on line 248 of `start.m`  
- Column labeled "nu = 0": run `/Main nu infinity/start.m` and change the switch indicator to 'table5' on line 248 of `start.m`
- Table 6  
- Column labeled "nu = 1/3": run `/Main nu 3/start.m` and change the switch indicator to 'table6ss' on line 248 of `start.m`  
- Column labeled "nu = 1/10": run `/Main nu 10/start.m` and change the switch indicator to 'table6ss' on line 248 of `start.m`  
- Column labeled "nu = 0": run `/Main nu infinity/start.m` and change the switch indicator to 'table6ss' on line 248 of `start.m`  
- Column labeled "nu = 1/3, rm increase": run `/Main nu 3/start.m` and change the switch indicator to

- 'table6up' on line 248 of start.m
- Column labeled "nu = 1/3, rm decrease": run /Main nu 3/start.m and change the switch indicator to 'table6down' on line 248 of start.m
- Table 7
  - Panel A: run "Main nu 3/start.m and change the switch indicator to 'table7' on line 248 of start.m
  - Panel B: run "Main nu 10/start.m and change the switch indicator to 'table7' on line 248 of start.m
  - Panel C: run "Main nu infinity/start.m and change the switch indicator to 'table7' on line 248 of start.m
- Table 8
  - Column labeled "Only Payment Reduction": run /Main nu 3/start.m and change the switch indicator to 'table8A' on line 248 of start.m
  - Column labeled "Payment and Principal Reduction": run /Main nu 3/start.m and change the switch indicator to 'table8B' on line 248 of start.m [Note that the code produces numbers for the change in welfare/transfer slightly less than 1 due to numerical errors in approximating the derivative of the value function. We report the theoretically correct number (1) in text].
- Table 9
  - Panel A: run /Main nu 3 Forbearance/start.m to report results from the "after" Column. The "Before" column simply reproduces the numbers from Table 2. (lines 31 and 250 in start.m need to be commented out.)
  - Panel B: run /Main nu 3 Forbearance/start.m after uncommenting lines 31 and 250 in start.m
- Table 10
  - Panel A: run /Main nu 3 2016/start.m
  - Panel B: see the parameters at the beginning of start.m
  - Panel C: comment the return statement on line 245 and change the switch indicator to 'steady', 'rup' or 'rdown' to reproduce the first, second and third column of Panel C, respectively.
- Table 11
  - To reproduce Panels A and C, run the following files below. For the parameter values in Panel B, see their values at the beginning of each file.
  - Column labeled "Constant rl": run /Main no rl heterogeneity/start.m
  - Column labeled "Moving shocks": run /Main Moving/start.m
  - Column labeled "No home production": run /Main no home production/start.m
  - Column labeled "House price growth": run /Main nu 3 House Price growth/start.m
  - Column labeled "Nonsep. utility": run /Main nu 3 Non Separable Utility/start.m
  - Column labeled "Higher rm": run /Main nu 3 High rm/start.m
- Figures 1 and 2
  - run /Simple Exponential/start.m
- Figure 3
  - run "Main nu 3/start.m
- Figure 4
  - run /Main nu 3/start.m and change the switch indicator to 'figure4' on line 248 of start.m
- Figure 5
  - run /Main nu 3/start.m and change the switch indicator to 'figure5' on line 248 of start.m

### 3. Data

The code in this section constructs the data moments used in the paper from two main sources: the Panel Study of Income Dynamics (PSID) and the Survey of Consumer Finances (SCF). The codes were written in Stata and can be run on versions 15/16/17.

*Data Availability Statement:*

- PSID: All PSID data sets used to produce results in this paper are publicly available, but require registration. The raw PSID family files can be accessed at <https://psidonline.isr.umich.edu/GettingStarted.aspx> (check section on "Downloading the Data"). We stored the raw data and the codes that we used to produce results in our paper on the PSID Public Data Extract Repository: <https://doi.org/10.3886/E145381V1>. See the contents inside the zipped folder `data_replication`. The raw data is in the subfolder `data_replication\psid\raw_data`.
- SCF: All SCF data sets used to produce results in this paper are publicly available (see <https://www.federalreserve.gov/econres/scfindex.htm>). We stored the raw data files that were used to produce results for our analysis inside our replication package, under `data_replication\scf\raw_data`.

The replication package contains a subfolder called `data_replication`. This subfolder includes another subfolder (`scf`) that contains all of the codes used to clean the SCF data and generate the output for figures and tables of our main text and appendix. A similar structure is followed in our online repository for the PSID data (<https://doi.org/10.3886/E145381V1>). The `psid` and `scf` subfolders have the following contents (subfolders):

- `\code\` contains all codes necessary to produce the output used in our tables and figures. IMPORTANT: All codes require the users to specify the path to folders that contain the raw and processed data (check codes for comment line `"*Set path"`).
- `\proc_data\` is used as a storage for processed data that our codes produce at interim steps.
- `\output\` contains `.csv` files with our results that are summarized in various tables of the paper.
- `\raw_data\` contains the raw data and some codes that help clean the data.

## A. Data cleaning

*Panel Study of Income Dynamics* (all files available at <https://doi.org/10.3886/E145381V1>)

The folder `psid` contains all the `.do` files necessary to produce the cleaned PSID sample. We enumerate below each of these files and the necessary raw data used as an input.

### 1. `taxsim.do`

- (a) Uses the raw PSID data to compute disposable income for each year by specifying the inputs for the `Taxsim` routine. `Taxsim` is a model designed by the NBER that computes federal and state income taxes. We used the `taxsim9` command in Stata to run the model.

IMPORTANT:

- i. `taxsim9` has been obsoleted by `taxsim27` (more information on this here: <https://users.nber.org/~taxsim/>). It's still available but no longer updated. We provided the `ado` and `help` files for `taxsim9` that were used to process our raw data. These files are stored in `\data_replication\psid\code\taxsim9_ado_files` for both the PSID and the SCF data replication folders. In order to run our codes on your computer you should paste the `ado` and `help` files in the folder that Stata calls when looking for `ado` files on your computer. On Windows computers, these are typically stored in `C:\Users\...\ado\plus\t` (i.e., subfolder containing `ado` files that start with the letter "t").
- ii. Some users of `Taxsim` experience errors when running the `taxsim9` routine due to `ftp` firewall access restrictions on their computers. You can find more information on how to overcome these here: <http://users.nber.org/~taxsim/ftp-problem.html>. If you cannot make this work, please use our processed data files that are stored in the folder `\data_replication\psid\proc_data\our_proc`. These files are the ones you will need in order to run the codes that produce the tables and figures in the paper. Just copy them from `\data_replication\psid\proc_data\our_proc` into `\data_replication\psid\proc_data` and you'll be able to run the rest of codes without any issue.

- (b) We stored the raw PSID data that we used to produce the data moments in our paper in `data_replication\psid\raw_data`. The files have the following format: `FAM1999ER.txt` where 1999 stands for survey wave year.
  - (c) In order to run `taxsim.do`, you'll need to specify the path to the following folder `...\data_replication\psid\` on your computer (check comment line for path: `"*Set path"` in the `.do` file).
  - (d) The `taxsim.do` file calls a `.do` file `FAM`x'ER.do` where ``x'` is the year of survey (1999-2011). These `.do` files read in the `.txt` files containing raw PSID data and label them appropriately. These `.do` files are stored in `...\data_replication\psid\raw_data`
  - (e) The `taxsim.do` file also includes an additional sub-routine `save_data`x'.do`, where ``x'` is the year of survey (1999-2011). The sub-routine generates intelligible names for PSID variables. Note that the original variable names are changing across survey waves without significant changes in the contents of each variable. These `.do` files are also stored in `...\data_replication\psid\raw_data`
  - (f) The output of `taxsim.do` is saved in `\data_replication\psid\proc_data`
2. `merge_panel.do`
- (a) Generates a panel from individual waves of the PSID data that were created in the previous step by `taxsim.do`.
  - (b) It uses the *Cross-year Individual File* of the PSID in order to create a personal ID and a household head status (raw data file that we used can be found under the name `IND2011ER.txt` in subfolder `...\data_replication\psid\raw_data`).
  - (c) The file `merge_panel.do` calls another `.do` file (`IND2011ER.do`) that reads in the raw data. This `.do` file is in `...\data_replication\psid\raw_data`.
  - (d) It then joins the data on disposable income for each PSID wave (output of `taxsim.do`) by exploiting the unique household ID and tracking the household heads across time.
3. `consumption.do`
- (a) Uses raw PSID data on consumption to generate `.dta` files with variables that have intelligible names
  - (b) It relies on the panel generated by `merge_panel.do` to single out the households in our sample for which we keep data on consumption.
  - (c) The file `consumption.do` requires raw data that comes under the names `CON`x'.txt`, where ``x'` is the year of survey (1999-2011). The raw files that we used can be found in subfolder `...\data_replication\psid\raw_data`. To read in the raw data, `consumption.do` calls a sub-routine `CON`x'.do`, which is also stored in subfolder `...\data_replication\psid\raw_data`.

### *Survey of Consumer Finances*

The folder `data_replication\scf\code` contains all the `.do` files necessary to produce the cleaned SCF sample. We store the cleaned sample in `data_replication\scf\proc_data` and use it as input in our codes producing the tables and charts in the paper. We enumerate below each of these `.do` files and the necessary raw data used as an input.

#### 1. `taxsim_2001.do` and `taxsim_2016.do`

- (a) The two files use raw SCF data from 2001 and 2016 to compute disposable income by specifying the inputs for the Taxsim model. As, in the case of the PSID data, we used the `taxsim9` command in Stata to run the model.

IMPORTANT:

- i. `taxsim9` has been obsoleted by `taxsim27` (more information on this here: <https://users.nber.org/~taxsim/>). It's still available but no longer updated. We provided the ado and help files for `taxsim9` that were used to process our raw data. These files are stored in `|.\code\taxsim9_ado_files|` for both the PSID and the SCF data replication folders. In order to run our codes on your computer you should paste the ado and help files in the folder that Stata calls when looking for ado files on your computer. On Windows computers, these are typically stored in `C:\Users\...\ado\plus\t` (i.e., subfolder containing ado files that start with the letter "t").
  - ii. Some users of `Taxsim` experience errors when running the `taxsim9` routine due to ftp firewall access restrictions on their computers. You can find more information on how to overcome these here: <http://users.nber.org/~taxsim/ftp-problem.html>. If you cannot make this work, please use our processed data files that are stored in the folder `\data_replication\scf\proc_data\our_proc`. These files are the ones you will need in order to run the codes that produce the tables and figures in the paper. Just copy them from `\data_replication\scf\proc_data\our_proc` into `\data_replication\scf\proc_data` and you'll be able to run the rest of codes without any issue.
- (b) `taxsim_2001.do` and `taxsim_2016.do` use as inputs the raw and summary SCF files from <https://www.federalreserve.gov/econres/scf-previous-surveys.htm> (check the pages for the appropriate survey wave). We stored the raw SCF data that we used to produce the data moments in our paper in `data_replication\scf\raw_data`.
  - (c) In order to run `taxsim_2001.do` and `taxsim_2016.do`, you'll need to specify in each of them the path to the following folder `...\data_replication\scf\` on your computer (check comment line for path: `"*Set path"`).

## B. Tables and Figures

The list below describes the codes that need to be run in order to compute the data moments used in the paper. The first 4-5 characters in the file names stand for the table or figure count used in the paper and appendix (e.g., `tab1` stands for table 1 in the paper).

*Panel Study of Income Dynamics* (psid folder, available as above at <https://doi.org/10.3886/E145381V1>)

1. `tab1_refi_moments.do` and `tab1_refi_liquidity.do`
  - (a) Generate the PSID data output reported in Table 1, Panel B.
  - (b) Use as inputs the panel created by `merge_panel.do`, and the subroutines `FAM`x'ER.do` and `wealth_comp`x'.do` which clean the wealth data. These subroutines are stored in `data_replication\psid\raw_data`.
  - (c) Output: `tab1_refi_moments.csv` and `tab1_refi_liquidity.csv`
2. `tab2_income.do`
  - (a) Generates the moments for the income process reported in Table 2, Panel A.
  - (b) Use as inputs the panel created by `merge_panel.do`, and output files of `consumption.do`.
  - (c) Output: `tab2_income.csv`
3. `tabA3_refi.do`
  - (a) Generates the fraction of households that extracts reported in Table A3 of the Appendix.
  - (b) This file requires you to run first `taxsim_2020.do` and `merge_panel_2020.do`. These two `.do` files are copies of `taxsim.do` and `merge_panel.do` adapted to work on raw data from the PSID surveys of 2013, 2015 and 2017. For more details on how to run these `.do` files, please see the descriptions for `taxsim.do` and `merge_panel.do` in the Data Cleaning section above.

- (c) Also uses the subroutines `FAM`x`ER.do` and `wealth_comp`x`.do` to clean the wealth data.
- (d) Output: `tabA3_refi.csv`

### *Survey of Consumer Finances*

1. `tab1_all_households.do`, `tab1_richest20%.do` and `tab1_poorest80%.do`
  - (a) Generate the motivating evidence from SCF data reported in Table 1, Panel A.
  - (b) Use as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `tab1_all_households.csv`, `tab1_richest20%.csv` and `tab1_poorest80%.csv`
2. `tab2_wealth.do`
  - (a) Generates the moments used in calibration reported in Table 2, Panel B.
  - (b) Uses as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `tab2_wealth.csv`.
3. `tab3_most_panels.do` and `tab3_panelF.do`
  - (a) Generate the moments reported in Table 3.
  - (b) Use as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `tab3_most_panels.csv` and `tab3_panelF.csv`
4. `tab4_ahead_on_payments.do`
  - (a) Generates the moment on "fraction ahead on payments" reported in Table 4, Panel A. Note that the rest of the moments in Table 4 are based on moments that were reported in Table 1, Panel B. Check `tab1_refi_moments.do` and `tab1_refi_liquidity.do` in the `psid` folder for code generating these moments.
  - (b) Uses as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `tab4_ahead_on_payments.csv`.
5. `tab10_wealth.do`
  - (a) Generates the moments reported in Table 10, Panel A. Note that the moment on "fraction borrowers who extract" is based on PSID data and created by the file `tabA3_refi.do` located in the folder `psid`.
  - (b) Uses as inputs the sample created by `taxsim_2016.do`.
  - (c) Output: `tab10_wealth.csv`.
6. `fig3.do`
  - (a) Generates the moments used to plot the life-cycle statistics in Figure 3.
  - (b) Uses as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `fig3.csv`.
7. `tabA3_wealth.do`
  - (a) Generates the moments reported in Table A3, Panel A, of the Appendix.
  - (b) Uses as inputs the sample created by `taxsim_2016.do`.
  - (c) Output: `tabA3_wealth.csv`.
8. `tabA10_wealth.do`

- (a) Generates the moments reported in Table A10, Panel A, of the Appendix.
  - (b) Uses as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `tabA10_wealth.csv`.
9. `tabA12_itemize.do`
- (a) Generates the moments reported in Table A12 of the Appendix.
  - (b) Uses as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `tabA12_itemize.csv`.
10. `tabA13_no_other_prop.do`
- (a) Generates the moments reported in Table A13 of the Appendix.
  - (b) Uses as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `tabA13_no_other_prop.csv`.
11. `fig1&2_appendix.do`
- (a) Generates the homeownership rates used to plot Figures 1 and 2 of the Appendix.
  - (b) Uses as inputs the sample created by `taxsim_2001.do`.
  - (c) Output: `fig1&2_appendix.csv`.

## References

Board of Governors of the Federal Reserve System. 2001 & 2016 Survey of Consumer Finances (SCF) Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI (2014).

PSID replication data (1999-2017) and codes for paper on "Liquidity Constraints in the U.S. Housing Market" by Corina Boar, Denis Gorea and Virgiliu Midrigan. <https://doi.org/10.3886/E145381V1>

## PSID Acknowledgement

The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and R01 AG040213, and the National Science Foundation under award numbers SES 1157698 and 1623684.