Code for replicating the findings of "Housing Consumption and Investment: Evidence from Shared Equity Mortgages" by Matteo Benetton, Philippe Bracke, Joao Cocco and Nic Garbarino

Date: July 2021

Code structure

Most of the code is written in Stata. The Stata script `00_mainscript.do` lists all globals (including customizable folder locations) and all the packages needed. It calls the other scripts and describes what output (Table or Figure in the paper) each script produces.

There are also four R scripts and three Excel spreadsheets. They are listed in `00_mainscript.do` too.

Scripts named with the prefix `01_` can be run first (after the globals and functions in `00_mainscript.do` have been called). Scripts named with `02_` depend on one or more `01_` scripts, and so on.

Data

Proprietary datasets:

1. The confidential mortgage *Product Sales Data* from the UK Financial Conduct Authority (FCA) records all residential mortgages originated in the UK (known as PSD001) and a semi-annual snapshot of all residential outstanding mortgages (known as PSD007). It can be accessed only by researchers affiliated to the FCA or the Bank of England.
2. Information on the full postcode and date of origination of all equity loans issued up to March 2017 were made available to the authors by Ministry of Housing, Communities and Local Government. The authors were also granted access to information on all repayment of equity loans that occurred before September 2017.
3. To investigate whether households have additional wealth beyond that used for the down payment, we have obtained anonymized data on the checking account balances at the end of the month for a 10% random sub-sample of UK residents. These checking account data are collected by a credit reference agency for the FCA. The credit records also 12 contain information on the financial products taken by individuals, including mortgages. The data can be accessed only by researchers affiliated to the FCA.
4. The Bank of England/NMG household survey data is available on the Bank of England website (https://www.bankofengland.co.uk/statistics/research-datasets) but for the paper we use an internal restricted version with information on the local authority of the respondent (which we use to measure local house price dynamics). This version of the dataset can be accessed only by researchers affiliated to the Bank of England.

Please contact Philippe Bracke (philippe.bracke@fca.org.uk) if you would like more information on these data sources.

Publicly available datasets:

5. Several of our analyses of housing transactions are based on the *England and Wales Land Registry*, which can be downloaded from https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads#single-file.
6. Local house price indices were downloaded from https://www.gov.uk/government/collections/uk-house-price-index-reports#archived-reports-and-data.
7. To assign our property postcodes to the correct UK regions and local authorities, we use the *National Statistics Postcode Lookup* (NSPL) from https://geoportal.statistics.gov.uk/.
8. To obtain information on the floor area of properties, we use the UK *Energy Performance Certificates* data, which are downloadable at https://epc.opendatacommunities.org/ after registration as researcher.
9. To draw the map of local authority boundaries we use the shapefile available at https://data.gov.uk/dataset/daaafdcc-f7c7-41ff-80eb-b0b15efd1414/local-authority-districts-december-2017-generalised-clipped-boundaries-in-united-kingdom-wgs84.