Readme.pdf for replication archive

"The Impact of New Housing Supply on the Distribution of Rents"

by Andreas Mense

Instructions

This replication archive reproduces all tables and figures in "The Impact of New Housing Supply on the Distribution of Rents".

To replicate the results, please run the R code in folder "code". You may need to adjust the working directory in the "HEADER" section of the file. Please set the working directory to the directory of the replication package.

- For results in Section 2 and Online Appendices A-C: run *reduced_form_replication.R*
- For results in Section 3 and Online Appendix D: run *dynamic_model_replication.R*

Please note that results in Section 3 and Online Appendix D cannot be reproduced without adding the scientific use files of SOEP v34 including region codes [*Raumordnungsregionen*/Planning Regions] and the *movedist* auxiliary data set. The author does not have permission to share these data sets.

In particular, the restricted-access data not included in replication package is the German Socio-Economic Panel, v34 (as of December 2019) including

- SOEP-core v34 (Sozio-oekonomisches Panel (SOEP), Daten der Jahre 1984-2017. DOI: 10.5684/soep.v34)
- SOEP-core v34 addon *geocodes*
    - geo-identifiers for planning regions
    - move indicators and move distances (*movedist*)

Access can be obtained via the DIW Berlin. For further information, please visit https://www.diw.de/en/diw_01.c.601584.en/data_access.html.

Version information

The archive was tested on: R version 4.3.2 (2023-10-31 ucrt), Platform: x86_64-w64-mingw32/x64 (64-bit), Running under: Windows 10 x64 (build 19045)

Archive structure and content

– **folders** and *files* –

- **code** – contains R code to reproduce all tables and figures
    - *dynamic_model_replication.R* – top-level R code to recreate results in Section 2 and related results from Online Appendices
    - *reduced_form_replication.R* – top-level R code to recreate results in Section 3 and related results from Online Appendices. This file requires restricted-access SOEP files.
    - **inc** – folder containing self-written function to create tex tables
    - **dynamic model** – contains further R scripts and C++ scripts used to estimate the model and run simulations
        - *EM_algorithm_while_loop.R* – estimation of dynamic model via EM algorithm

        - *R_functions.R* – user-written R functions used for model estimation and simulation
        - *functions.cpp* – user-written C++ code called from within R via Rcpp
- **data** – contains data underlying results from Section 2 of the paper and part of the data for Section 3; restricted-access data not included (see below)
    - **orig –** contains data underlying results from Section 2 of the paper and sub-folders with data used in the model estimation
        - *buildings_weather_shocks_panel_municip.RDS* – data file with municipality-level data on new construction and weather instruments
        - *control_variables_panel_pr.RDS* – data file with control variables at PR level
        - *price_data_MFH.RDS* – data file; listings data of multi-family housing units for sale, including housing characteristics and listing prices
        - *price_data_SFH.RDS* – data file; listings data of single-family housing units for sale, including housing characteristics and listing prices
        - *rent_data.RDS* – data file; housing units offered for rent, including housing characteristics and rental prices
        - *Zeitreihe_Fachkraeftemangel_Handwerk-Dezember.xlsx* – data underlying Figure 2 (scarcity of skilled labor and waiting times in the construction sector)
        - **model** – location to store restricted-access data (SOEP v34); see above for details on how to access the data
            - **auxiliary** – folder with auxiliary files
                - **interpolated_ror_rent_psqm_quantiles_MZ.RDS** – quantiles of the local rent distribution, computed from the German Mikrozensus and interpolated between years
                - **landlord_renovation_coefs.RDS** – used to compute landlords' renovation probabilities
                - **matched_rents.RDS** – data set of repeated rentals used to estimate depreciation of rental units
                - **share_new_supply.RDS** – aggregate shares of new supply used in simulations
            - **restricted** – folder for restricted-access SOEP data
                - **Required files:** *hgen.dta, hwealth.dta, movedist.dta, pequiv.dta, ror_l.dta, savings.dta* – store these files here before running *dynamic_model_replication.R*.
        - **shapes** – shapefiles for municipalities (MUN), planning regions (PR), commuting zones (CZ), and raster data
            - **MUN**
                - *VG250_GEM.shp* – shapefile with municipality borders
            - **PR**
                - *PR250.shp* – shapefile with borders of planning regions
            - **CZ**
                - *AMR250.shp* – shapefile with borders of commuting zones
            - **raster** – raster data underlying rainfall instrument
    - **tmp** – empty container for temporary files created by the R scripts
- **figures** – empty container for figures
- **model_output** – container for model output; also includes part of the code used for the estimation of the model
    - **iter** – container to store iterations of the EM algorithm used to estimate the dynamic discrete choice model with unobserved heterogeneity

        - *parallel_psi.R* – R code to update the control function in the EM algorithm, executed in parallel to make use of multiple cores of the user's computer
    - **VCOV** – container to store bootstrap results for parameters of the dynamic discrete choice models
- **tables** – empty container for tables

Variable descriptions

The rents and house price data are labeled in German.

AGS – municipality identifier
Jahr - year observed
PLZ – postcode
Baujahr – Year of construction
Alter – building age (year observed minus year of construction)
Flaeche – living area
Zimmer – number of rooms
Typ – housing type
EtagenGes – total number of floors
Etage – floor
Kaltmiete – net rent
StellplMiete – rent for parking
Fussbodenheizung – floor heating
Etagenheizung – separate heaters on each floor
Ofenheizung – stove heating, typically wood- or coal-fired
Zentralheizung – central heating
Energiekennwert – energy consumption class (certificate)
Ausweistyp – type of energy consumption certificate
Zustand_renoviert – unit is renovated
Zustand_saniert – unit is refurbished
Zustand_Erstbezug – unit is first-use
Zustand_renovbed – unit needs renovation
Zustand_baufaellig – unit is not habitable
Ausstattung – indicator for overall quality of unit
Keller – cellar
GaesteWC – second bathroom
EBK – fitted kitchen
Aufzug – elevator
Sauna – sauna
Pool – swimming pool
Parkett – parquet flooring
Klimaanlage – air conditioning
Kamin – fireplace
Gartenmit – use of garden
BalkonTerrasse – balcony or terrace
Stellplatz – parking lot available
Monat – month of observation
SN_AMR – identifier for commuting zone
SN_ROR – identifier for planning region
nID – non-systematic ID