**Readme: Replication Files**

*Authors: Marc Francke & Matthijs Korevaar*

This folder contains the code and data to replicate the findings of “Baby Booms and Asset Booms: Demographic Change and the Housing Market”, published in the Journal of Finance. This includes all Figures and Tables produced in the main paper, as well as any additional results published in the Internet Appendix. This document provides guidance on how to replicate the results.

For the results of this paper, most of the analysis was run using R but some supplementary analyses were conducted using Stata and Python. The folders “R”, “Stata” and “Python” contain the codes and data to conduct these analyses, and are self-contained.

*R code and corresponding data*

The main code is in the folder “R” and named “MainAnalysisCode.R”. This code produces the results for Figure 2, Figure 4 and Figure 6 of the main paper, as well as for all Tables (Tables 1 to 5). Additionally, it produces the results for Figures 3, 4, 5 and 7 of the Internet Appendix and Table VI of the Internet Appendix. It uses “AmsterdamMainSeries.csv” and “ParisMainSeries.csv” as input, which contain the main time series used for the analysis in the paper for Amsterdam and Paris. All code is click and run.

To replicate Figure 3 of the Internet Appendix one has to change “cohortsize” in line 74 from 5 to 10, and run the analysis up to line 246 with cohortsize set to 10. Subsequently, the code in line 312 to 319 generates Figure 3. For all other analysis, cohort size should be set to 5.

The folder “R” also contains the code “Figure1.R” and “Figure3.R”. These construct Figures 1 and 3 in the paper and are based on the secondary data in “Figure1.xlsx” and “Figure3.csv”

Finally, the folder “R” contains the code “AppendixSupplementaryAnalysis.R”, which contain the code to produce various supplementary analyses published in the Internet Appendix. The relevant Data is found in the Data folder. More specifically, the code reproduces Table II of the Internet Appendix (data input: “MarriageswithWealthAgeDeath.csv”), Figure 1 of the Internet Appendix (data input: “Match1832Rent1851Residents.csv”), Figure 2 of the Internet Appendix (data input: “BirthRateMultiCityData.xlsx”) and Table III of the Internet Appendix, which use the baseline series (“AmsterdamMainSeries.csv” and “ParisMainSeries.csv”) as input.

*Python code and corresponding data*

The code for section “V.A Mechanisms | Probability of Sale” and “V.B Mechanisms | Segmentation of markets” is in the folder “Python”. Required packages and used versions can be found in “JFBabyBoom Python environment.txt”.

The code for “Probability of Sale” is in the Jupyter notebook “ProbSale.ipynb”.

* It estimates Eq. (2).
* Processing of the data is handled in DataProcess.py and uses the data files: “Data/matches.csv” and “Data/ownrentperstreet.csv”.
* Explanatory variables for the probability of sale model are taken from “R/Data/ AmsterdamMainSeries.csv”.
* The code produces Figure 5 (“Output/ProbSaleBiLags\_5\_ClusteredSE\_MultID\_Year.eps”) and Internet Appendix Table IA.VII.

The code for “Segmentation of markets” is in the Jupyter notebook “RepeatSales.ipynb”.

* It estimates Eq. (IA.1).
* Processing of the data is handled in DataProcess.py and uses the data files: “Data/matches.csv” and “Data/ownrentperstreet.csv”.
* The model specification is provided in the Stan file “Stan/RS\_Cluster.stan”.
* Estimation output is written in “Output/RS\_Cluster\_1yr\_ext.csv”.
* The estimated indices are written in “Output/hpi\_results.csv”.
* The code produces Internet Appendix Figure IA.6 (“Output/AmsHPI.eps” and “Output/AmsHPILowHigh.eps”.
* Panel A of this Figure also displays the repeat sales index, based on the full dataset as provided in the data file “R/Data/ AmsterdamMainSeries.csv”.

*Stata code and corresponding data*

The code for Internet Appendix Section “IA.F. Modern evidence” is in the folder “Stata”. The summary statistics in Tables IA.VIII and IA.IX are based on the data file “OECD\_merge.dta” and the do file “Summary.do”. The do file “OECD\_Panel.do” estimates Eq. (IA.2), based on the same data, and provides the results in Table IA.IV.