README

The associated folder contains the data and programs used to make the empirical analysis in Epple, Quintero, and Sieg (2019).

****
**************1. Folders****************
****

1. The folder includes the following folders and subfolders:

- Data: Includes the datasets used to run the final models.
- Programs: Includes the programs used to run the final models, as well as some robustness tests.
- Programs/Structural Model Estimation: Includes all Matlab code needed to run the estimation of the structural model for the 1 metro area and the 2 metro area versions.
- Programs/Calculate Data Correlations: Includes Stata code needed to calculate and export to Matlab the data rent and income correlations that will be matched by some of the structural model's moments
- Programs/UserCostwithEquity: Includes Stata code needed to calculate the equity to be added to the income of owners
- Programs/WagePremiumMoversACS: Includes Stata code needed to calculate the wage premium of movers from one metro area to another.

****
**************2. Programs**************
****

2. Before running, note the following:

A. The estimation of the structural model is run in Matlab. Besides all .m code files included in the folder Programs/Structural Model Estimation, the calculations require Matlab's statistics toolbox. Run the 4 sections of the file a0Master to run both the 1 metro area and the 2 metro areas model estimation. Each of the main programs and subroutines include a description in the first line. Documentation about each sub routine 'file.m' can be obtained by typing 'help file.m' in Matlab's command line.

B. To calculate rent and income correlations in general and by observed type run 01_CalculateCorrs_IncomeRent_Chi_NYC.do in Stata. The output of this is included in the structure array used in A. in the field Corr.

C. To calculate income that is due to owners' equity, first run 01_CalculateEquity_Chicago1999.do and 01_CalculateEquity_Chicago2003.do in Stata. These programs are transformations of the code used in the National Table Specs. pdf downloaded from http://www2.census.gov/programs-surveys/ahs/2003/AHS_2003_Table_Specs.zip. This code was previously unavailable in Stata, so we expect this adaptation to widen its use. Second, run 02_CalculateUserCostwithEquity_Chicago1999.do and

02_CalculateUserCostwithEquity_Chicago2003.do. These follow an iterative procedure that takes this equity and transforms it to the implied rental income and adds it to income. This is used to generate a new iteration of the user cost, and iterate until convergence. The modified resulting income can be used to change the owners' income used in the estimation of the structural model in A.

D. To calculate the income premium of movers and stayers, run 02_RunRegressionACS_Measurewagepremiuminmovers.do in Stata. This uses regressions on data obtained from the harmonized samples in Integrated Public Use Microdata Series (IPUMS) in 2000 to test the compensating variation premiums implied in the model. The calculation is made only using employed heads of household.

****
**************3. Data files**************
****

3. This is the description of the data files included in the Data folder:

3.1. 'MetroAreasData.mat' is a structure array type matlab dataset used to estimate the structural model. The structure array has fields for New York (NYC), Chicago (Chi), and Aggregate (Agr) data. The subfields Micro1999 and Micro2003 in fields NYC and Chi include the micro data from the AHS. The YIk and VIk fields include empirical cdfs based on the micro data for the general population and for the 3 observed types determined by k-means clustering, as well as parameters of the fitted cdfs.

3.2. 'corrrentyChi99_k.dta', 'corrrentyChi03_k.dta', 'corrrentyNYC99_k.dta', 'corrrentyNYCi03_k.dta' are American Housing Survey micro data files used to calculate the data correlation between rent and income for all households and for sub types of households.

3.3. 'hud1999.xpt' and 'hud2003.xpt' provide American Housing Survey micro data with mortgage information used to calculate the outstanding balance and equity of households on their residence. 'Chicago_equity_1999.dta' and 'Chicago_equity_2003.dta' has the resulting equity data. Chicago_ucequity_income_iter_1999.dta and 'Chicago_ucequity_income_iter_2003.dta' includes the user costs for every iteration of the process in D. above.

3.4. 'ACSReg.dta' provides data obtained from the harmonized samples in Integrated Public Use Microdata Series (IPUMS) in 2000. The data was cleaned and organized using the variable labels and categories generated by IPUMS Fixed-width Text Files system (https://usa.ipums.org/usa-action/extract_requests/download) for Stata.

Last revision: Feb 2, 2019.