

Housing over Time and over the Life Cycle: A Structural Estimation Online Appendix*

Wenli Li Haiyong Liu Fang Yang Rui Yao

ABSTRACT

This online appendix presents model simplifications and numerical solutions, the estimation of labor income process, the construction of comparable state house price indexes, and estimation mechanics in the method of simulated moments (MSM) estimator. It also includes two tables and nine figures that were not included in the main paper due to space limitations.

1. Appendix A: Model Simplifications and Numerical Solutions

The household's optimization problem is scale-independent and can be simplified by normalizing the household's continuous state and choice variables by its permanent income P_t^Y . We transform the vector of endogenous state variables to $x_t = \{D_{t-1}^o, q_t, \bar{h}_t, P_t^H\}$, where $q_t = \frac{Q_t}{P_t^Y}$ is the household's wealth-to-permanent labor income ratio, and $\bar{h}_t = \frac{P_t^H H_{t-1}}{P_t^Y}$ is the beginning-of-period house value to permanent income ratio. We transform choice variables similarly. Let $c_t = \frac{C_t}{P_t^Y}$ be the consumption-to-permanent income ratio, $h_t = \frac{P_t^H H_t}{P_t^Y}$ be the house value-to-permanent income ratio, and $s_t = \frac{S_t}{P_t^Y}$ be the net financial wealth-permanent income ratio. The evolution of normalized endogenous state variables can then be rewritten as:

*Wenli Li: Research Department, Federal Reserve Bank of Philadelphia, wenli.li@phil.frb.org; Haiyong Liu: Department of Economics, East Carolina University, liuh@ecu.edu; Fang Yang: Department of Economics, Louisiana State University, fyang@lsu.edu; Rui Yao: the Zicklin School of Business, Baruch College, rui_yao@baruch.cuny.edu. We thank Andra Ghent, Christopher Carroll, Juan Contreras, Alex Michaelides, Jesus Fernandez-Villaverde, three anonymous referees, and participants at various seminars and conferences for their comments. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, or the Federal Reserve System.

$$\bar{h}_{t+1} = \left[\frac{(\xi_{t+1}^s \xi_{t+1}^i)}{\exp\{f(t+1)\} \nu_{t+1}^s \nu_{t+1}^i} \right] h_t, \quad (1)$$

$$P_{t+1}^H = P_t^H (\xi_{t+1}^s \xi_{t+1}^i), \quad (2)$$

and

$$q_{t+1} = \begin{cases} \max\left\{ \frac{s_t [1+r(s_t)] + D_t^o h_t [(\xi_{t+1}^s \xi_{t+1}^i)^{1-\phi}]}{\exp\{f(t+1)\} \nu_{t+1}^s \nu_{t+1}^i} + \varepsilon_{t+1}, \eta \right\}, & \text{if } t+1 < t_R \\ \max\{s_t [1+r(s_t)] + D_t^o h_t [(\xi_{t+1}^s \xi_{t+1}^i)^{1-\phi}] + \theta, \eta\}, & \text{if } t+1 \geq t_R \end{cases} \quad (3)$$

in which

$$r(s_t) = \begin{cases} r & \text{if } s_t \geq 0, \\ r_m & \text{if } s_t < 0. \end{cases} \quad (4)$$

The intertemporal budget constraint is then governed by:

$$q_t = \begin{cases} c_t + s_t + \alpha h_t, & \text{rent at } t-1 \text{ and } t (D_{t-1}^o = D_t^o = 0); \\ c_t + s_t + \alpha h_t, & \text{own at } t-1, \text{ sell and rent at } t (D_{t-1}^o = D_t^s = 1, D_t^o = 0); \\ c_t + s_t + (1+\psi)h_t, & \text{rent at } t-1, \text{ own at } t (D_{t-1}^o = 0, D_t^o = 1); \\ c_t + s_t + (1+\psi)h_t, & \text{own at } t-1, \text{ sell and own at } t (D_{t-1}^o = D_t^s = D_t^o = 1); \\ c_t + s_t + (1+\psi-\phi)\bar{h}_t, & \text{own the same h at } t-1 \text{ and } t (D_{t-1}^o = D_t^o = 1, D_t^s = 0). \end{cases} \quad (5)$$

Define $b(q_t)$ to be the normalized bequest function:

$$b(q_t) = L^\gamma q_t^{1-\gamma} \frac{\left[(1-\omega) \left(\frac{(1-\omega)^\zeta}{(1-\omega)^\zeta + \omega^\zeta (\alpha P_t^H)^{1-\zeta}} \right)^{1-\frac{1}{\zeta}} + \omega \left(\frac{\omega^\zeta (\alpha P_t^H)^{-\zeta}}{(1-\omega)^\zeta + \omega^\zeta (\alpha P_t^H)^{1-\zeta}} \right)^{1-\frac{1}{\zeta}} \right]^{\frac{1-\gamma}{1-\frac{1}{\zeta}}}}{1-\gamma}. \quad (6)$$

We denote $v_t(x_t) = \frac{V_t(X_t)}{(P_t^Y)^{1-\gamma}}$ to be the normalized value function and $a_t = \{c_t, h_t, s_t, D_t^o, D_t^s\}$ to be the normalized vector of choice variables. The normalized recursive optimization problem can be rewritten as:

$$v_t(x_t) = \max_{a_t} \left\{ \lambda_t \left[\frac{N_t^\gamma}{1-\gamma} \left((1-\omega) c_t^{1-\frac{1}{\zeta}} + \omega (h_t/P_t^H)^{1-\frac{1}{\zeta}} \right)^{\frac{1-\gamma}{1-\frac{1}{\zeta}}} \right. \right. \\ \left. \left. + \beta E_t \left(v_{t+1}(x_{t+1}) (\exp\{f(t+1)\} \nu_{t+1}^s \nu_{t+1}^i)^{1-\gamma} \right) \right] + (1-\lambda_t) b(q_t) \right\},$$

subject to

$$c_t > 0, \quad h_t > 0, \quad s_t \geq -(1 - \delta)D_t^o h_t,$$

and equations (1) to (5). With P_t^Y no longer serving as a state variable, the normalization reduces the number of continuous state variables to three.

We solve the normalized recursive optimization problem using backward induction. At the terminal date T , since $\lambda_T = 0$, the household's value function coincides with the bequest function, i.e., $v_T(x_T) = b(q_T)$. The value function at date T is then used to solve for the optimal decision rules on the state space at date $T - 1$.

We discretize the wealth-to-labor income ratio (q_t) into 80 grids equally-spaced in the logarithm of the ratio, the house value-to-labor income ratio (\bar{h}_t) into equally spaced grids of 80, and the house price (P_t^H) into 80 grids equally-spaced in the logarithm of the price. The boundaries for the grids are chosen to be large enough so that our simulated sample observations always fall within the defined state space.

At each period t , for a household coming into the period as a renter ($D_{t-1} = 0$), we perform two separate optimizations conditional on house tenure choices – renting or owning – for the current period. We then compare the contingent value functions of renting and owning to determine a renter's optimal house tenure choice for the current period. To calculate the expected value function in the next period, we use discrete states to approximate the realizations of continuous exogenous shock variables by Gaussian quadrature (Tauchen and Hussey 1991). For points that lie between grid points in the state space, we use a two-dimensional cubic spline interpolation to approximate the value function of renting and a three-dimensional cubic spline interpolation to approximate the value function of owning.

For a household coming into period t as a homeowner, we perform an optimization conditional on staying in the existing house for the current period. The value function contingent on moving – either endogenously or exogenously – is the same as the value function of a renter at period t who is endowed with the same wealth-to-income ratio (q_t) and house price (P_t^H). We then compare the value function conditional on moving with that on staying to determine the optimal house liquidation decision for the current period.

This procedure is repeated recursively for each period until period $t = 0$.

2. Appendix B: Constructing Labor Income Process

Using data from the Panel Study of Income Dynamics (PSID) from 1984 to 2005, we eliminate the Survey of Economic Opportunities subsample and households living in public housing projects owned by a local housing authority or public agency. We further exclude households that neither own nor rent, or whose head is female, a farmer, or a rancher. We use only households whose heads are married and are between 20 and 70 years of age. As described in Section 3 of the main paper, the federal and state income tax liabilities are obtained from the TAXSIM program. We regress the logarithm of after-tax household labor income on dummy variables for age, education, and household composition, using a household fixed effect model. A sixth-order polynomial is used to fit the age dummies in order to obtain the labor income profile.

Following the variance decomposition procedure described by Carroll and Samwick (1997), we first define a d -year income difference as follows:

$$r_d = \log(Y_{t+d}) - \log(Y_t).$$

Thus,

$$\text{Var}(r_d) = 2 * \sigma_\varepsilon^2 + d * \sigma_\nu^2.$$

We then regress $\text{Var}(r_d)$'s calculated from the data on d 's to obtain estimates on σ_ε^2 and σ_ν^2 , where σ_ν^2 is the sum of $(\sigma_\nu^s)^2$ and $(\sigma_\nu^i)^2$. We choose d to be 1, 2, ..., 22.

3. Appendix C: Constructing House Price Series at State Level

Our annual state-level house price index (HPI) comes from the Office of Federal Housing Enterprise Oversight (OFHEO, now part of the new Federal Housing Finance Agency). The HPI is a time series price index that is set to 100 for every state for the base year 1980. This price index is thus not comparable cross-sectionally. To create a state-level price index series that is also cross-sectionally comparable, we use the housing price information from the PSID. In particular, we define house prices as prices per square foot of living space. Unfortunately, the PSID does not provide information on living space, and we have to impute the square

footage of homes for our data. Following Flavin and Nakagawa (2008), we first use data from the American Housing Survey (AHS) (1987-2005) to estimate a model of square footage as a function of the number of rooms and other housing characteristics common to both the AHS and the PSID, such as dummy variables representing whether the household was (1) located in a suburb, (2) located in a non-MSA region, (3) living in a mobile home, and a third-order polynomial in the number of rooms. Separate models were estimated for each of the four regions (Northeast, Midwest, South, and West). The regional models estimated from the AHS data, reported in Table A2, were then used to generate estimated square footage data for each PSID household. Using these estimates, we predict house sizes for all homeowners in our PSID sample. The nominal house prices per square foot are then obtained by dividing the house value reported from the PSID by the predicted house size. For each state, we can use the imputed nominal price in any year, along with the HPI from the FHFA to calculate the nominal house price for a benchmark year, 1993, which is the midpoint of the time frame of our data. Given the fact that the FHFA and the PSID surveyed different random samples of American households, we anticipate that the nominal prices for 1993 converted from different years might vary. We therefore choose to use the median of these converted values. Once the median nominal price is determined for each state in the benchmark year, we can scale the HPI from the FHFA so that the new HPI for each state i in year t as follows, $HPI_{i,t}^{new} = HPI_{i,t}^{FHFA} * NominalPrice_{i,1993} / HPI_{i,1993}^{FHFA}$.

4. Appendix D: Estimation Mechanics in the MSM Estimator

We assume that the “true” parameter vector $\theta^* = \{\beta, \gamma, L, \omega, \zeta, \alpha, \psi\}$ lies in the interior of the compact set $\Theta \subset \mathbb{R}^7$. Our estimator, $\hat{\theta}$, is the value of θ that minimizes the weighted distance between the estimated life-cycle profiles for wealth, mobility rate, homeownership rate, house value, and rent observed from the data and the simulated profiles generated by the model. We choose to match these five variables, which we interact with age cohort (C) and calendar year (T). Additional interactions are used for the last three house-related variables, which we further interact with two house price levels in the state where a household resides. This interaction results in six additional moments. The moment count per year and cohort is therefore equal to $11(= 5 + 6)$. The overall count of moments is $11 \times C \times T = 33T$ because

we have three age cohorts. We combine all these moment conditions by stacking them and solving the optimal problems jointly.

Theoretically, the most efficient weighting matrix is the inverse of the sample variance-covariance matrix. However, according to Altonji and Segal (1996), the optimal weighting matrix, though asymptotically efficient, can be severely biased in small samples. We use a diagonal matrix for weighting given our small sample size. Our weighting matrix takes the diagonal terms of the optimal weighting matrix for scaling, while setting the off-diagonal term to be zero. A similar approach is adopted by De Nardi, French, and Jones (2010).

References

- Altonji, J. G., and Lewis M. Segal, “Small-Sample Bias in GMM Estimation of Covariance Structures,” *Journal of Business and Economic Statistics* 14(3) (1996), 353-366.
- Carroll, C. D., and A. A. Samwick, “The Nature of Precautionary Wealth,” *Journal of Monetary Economics* 40(1) (1997), 41-71.
- De Nardi, M., E. French, and J. B. Jones, “Why Do the Elderly Save? The Role of Medical Expenses,” *Journal of Political Economy* 118 (2010), 39-75.
- Flavin, M., and S. Nakagawa, “A Model of Housing in the Presence of Adjustment Costs: A Structural Interpretation of Habit Persistence,” *American Economic Review* 98(1) (2008), 474-495.

Table A1
Descriptive Statistics: Data Sample

Independent Variables	Mean	Standard Deviation
Age of family head	45.81	9.45
Number of children	1.20	1.17
Head high school graduate and above (%)	93	25
Head college graduate and above (%)	28	45
Head white (%)	80	40
Own house (%)	92	27
Annual family income (\$000)	58.04	37.91
Monthly rent	618	2525
House value (\$000)	167	149
Net worth (\$000)	122	262
Number of observations	17,392	

Notes: Data are from 1985 to 2004 Panel Study of Income Dynamics. The values are in 2005 dollars.

Table A2
Relationship Between House Size and Housing Characteristics
(Dependent Variable: House Size in Square Feet)

Independent Variables	Northeast	Midwest	South	West
Constant	-69.40 (51.47)	89.45 (45.40)	456.46 (34.71)	221.22 (32.85)
Urban	-75.44 (11.55)	-94.50 (8.05)	-91.32 (5.70)	-113.10 (8.47)
MSA	27.62 (14.07)	67.48 (8.09)	41.41 (5.84)	9.76 (8.88)
Mobile home	-492.63 (25.44)	-467.63 (15.46)	-299.46 (8.87)	-236.33 (12.53)
# rooms	282.68 (21.92)	204.28 (19.98)	-40.10 (15.01)	107.60 (13.86)
(# rooms) ²	20.88 (3.12)	27.39 (2.87)	55.90 (2.12)	34.55 (1.97)
(# rooms) ³	-1.55 (0.14)	-1.71 (0.13)	-2.50 (0.09)	-1.70 (0.86)
Adjusted R ²	0.23	0.25	0.30	0.33
Number of observations	77,126	108,727	159,671	94,800

Notes: Data are from 1987 to 2005 biennial American Housing Survey. Robust standard errors are reported in parentheses. We do not report estimates of survey year dummies.

MSA=metropolitan statistical area.

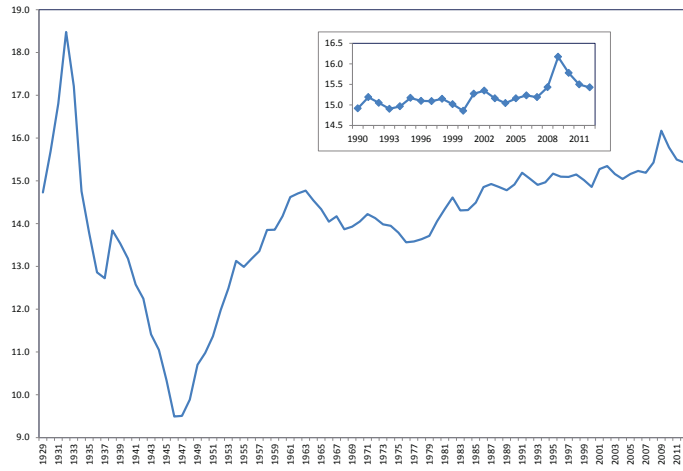


Figure A1. Aggregate housing expenditure share (data source: U.S. Department of Commerce, National Income and Product Account)

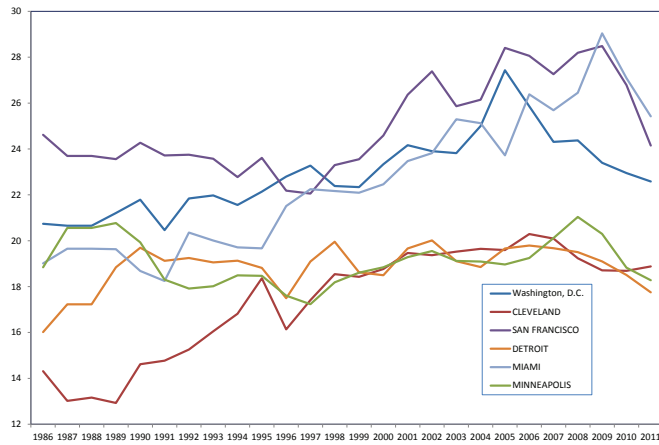


Figure A2. Housing expenditure shares in selected MSAs (data source: U.S. Department of Labor, Consumer Expenditure Survey)

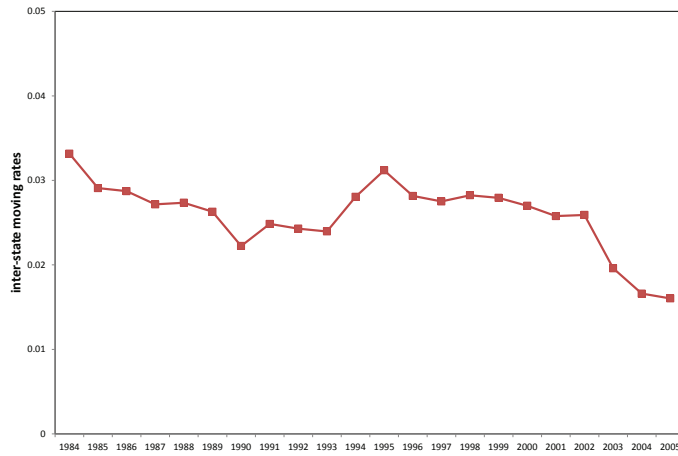


Figure A3. Average moving rates across states (data source: U.S. Census Bureau, Current Population Survey)

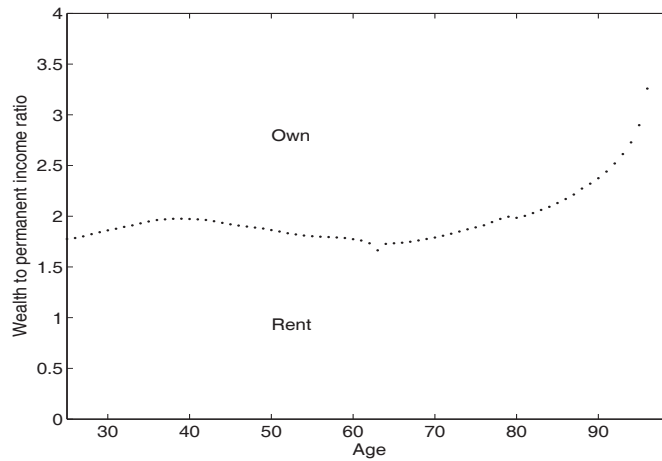


Figure A4. A renter's house tenure decision as a function of his wealth-to-permanent labor income ratio

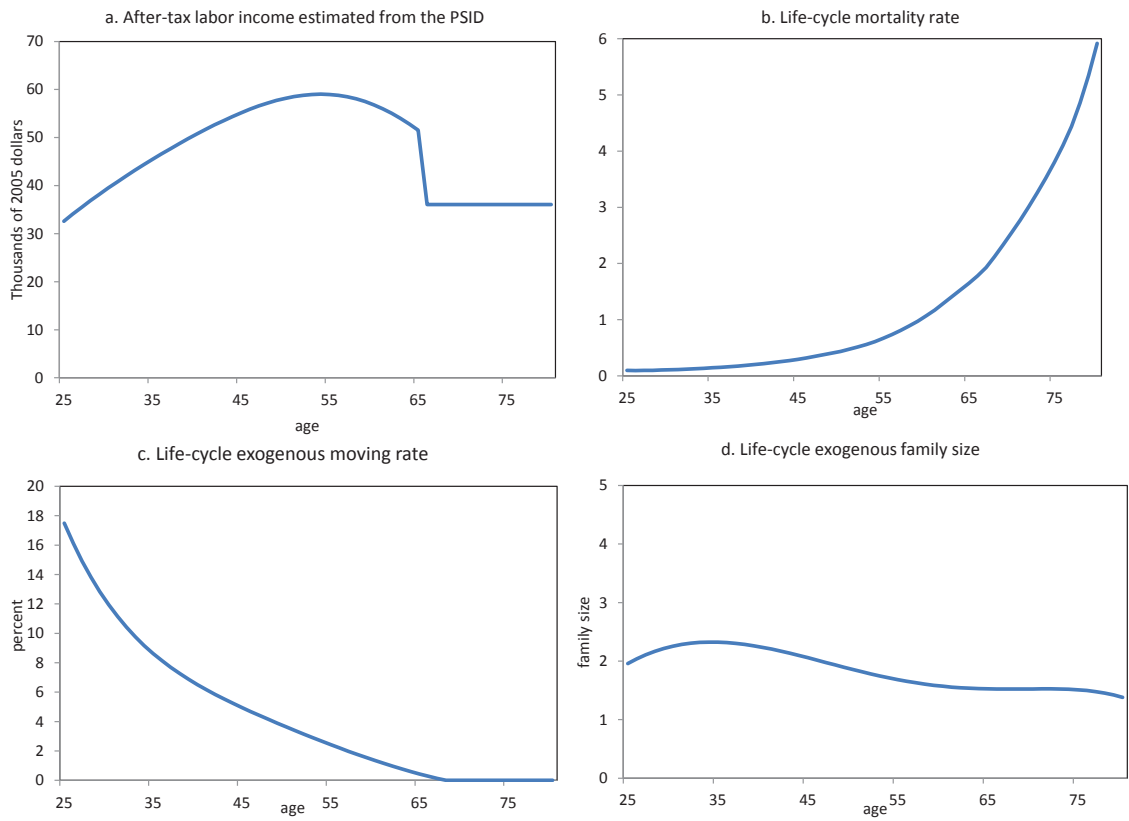


Figure A5. Exogenous processes in the model (data sources: Panel Study of Income Dynamics; the National Center for Health Statistics)

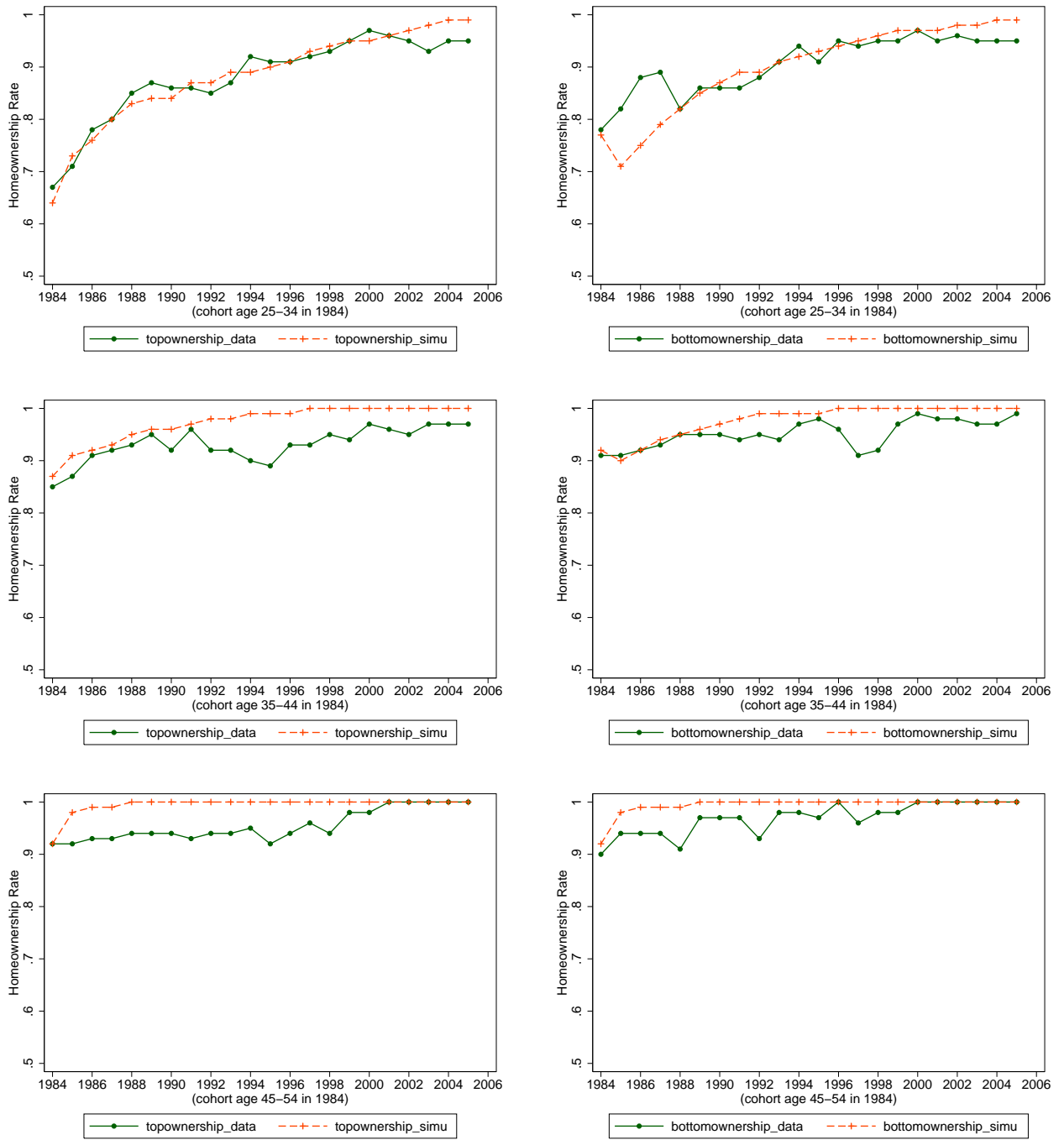


Figure A6. Homeownership by cohorts in high and low house price states

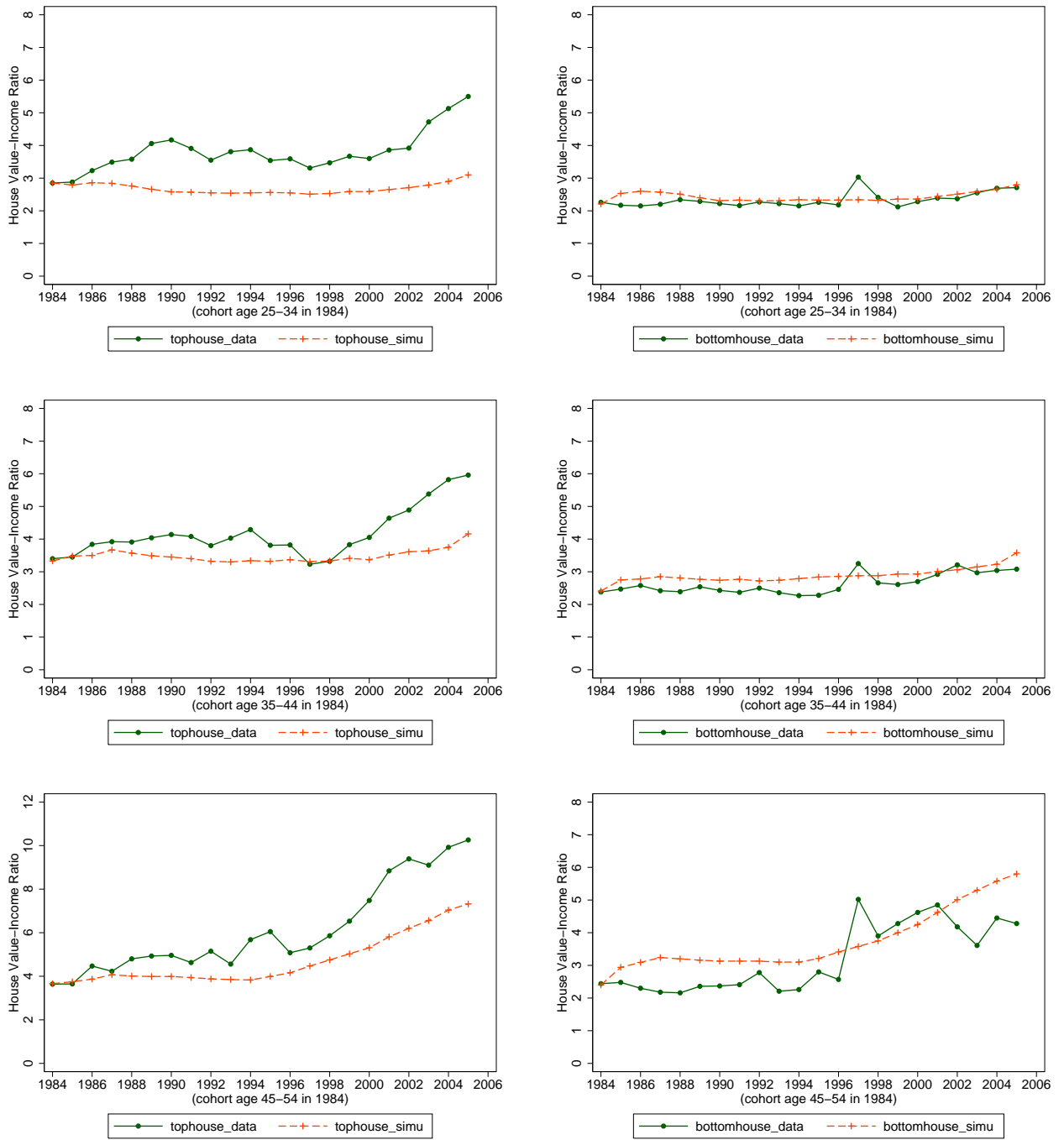


Figure A7. House value-to-income ratio by cohorts in high and low house price states

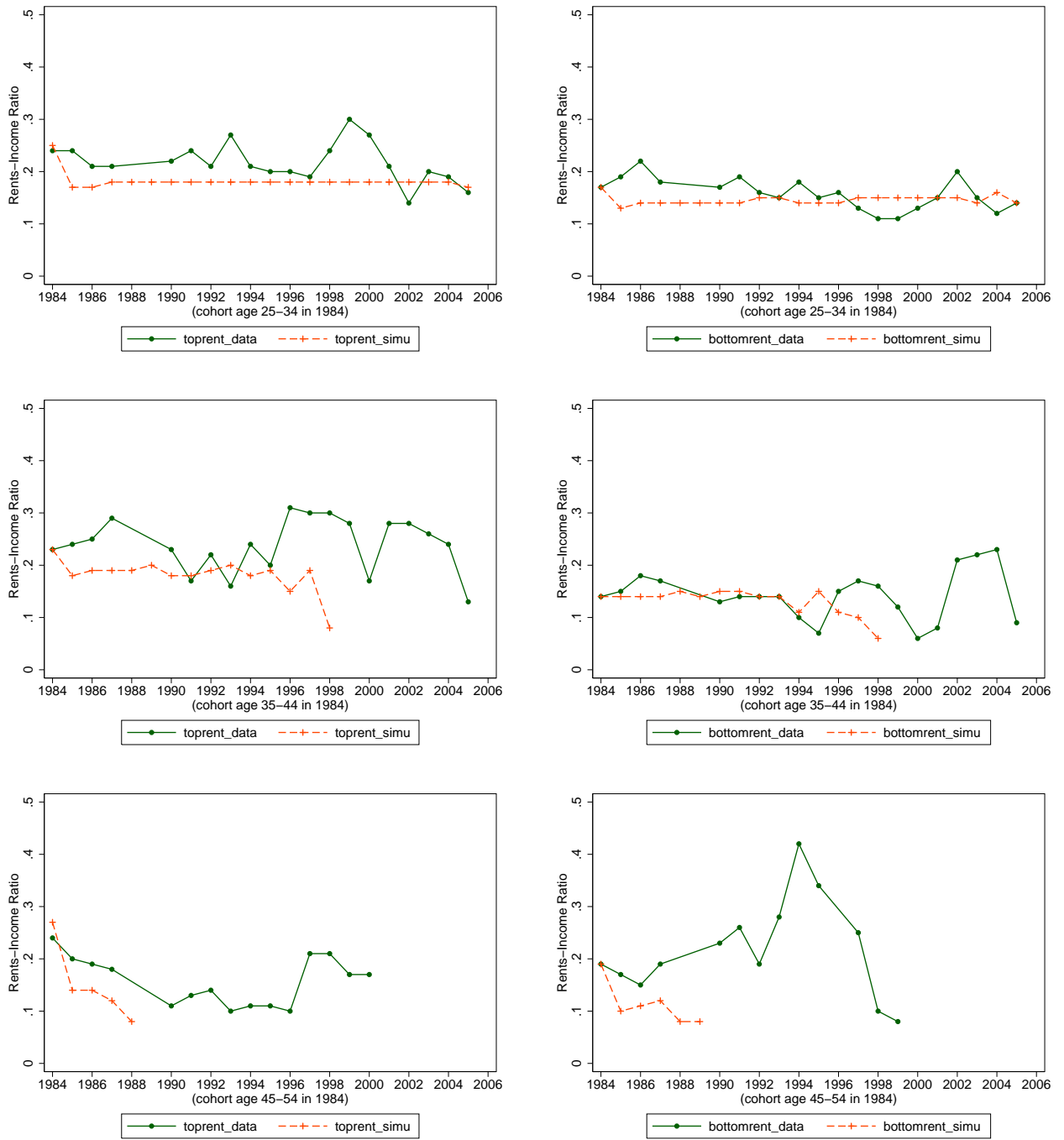


Figure A8. Rent-to-income ratio by cohorts in high and low house price states

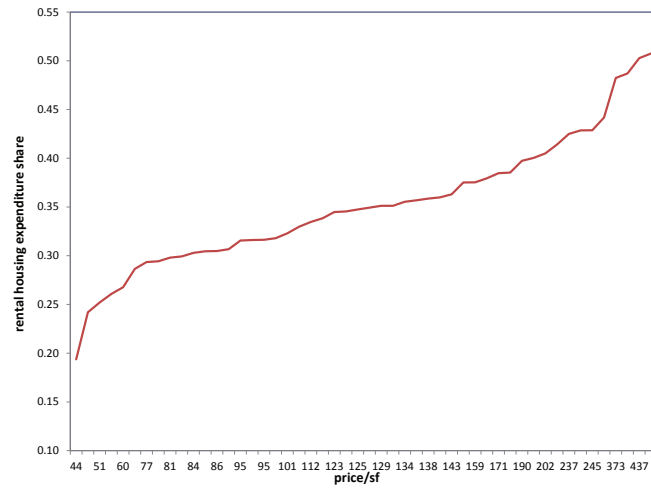


Figure A9. Simulated housing expenditure shares for renters