Internet Appendix for "The Predominance of Real Estate in the Household Portfolio"

Laurent Barras and Sebastien Betermier*

August 9, 2016

This Internet Appendix provides a detailed description of the data and the construction of variables, verifies the robustness of our empirical results to alternative assumptions, and presents a full derivation of the model.

*Desautels Faculty of Management, McGill, Bronfman Bdg, 1001 Sherbrooke St West, Montreal, QC H3A 1G5, Barras: laurent.barras@mcgill.ca, Betermier: sebastien.betermier@mcgill.ca

Contents

I Data Description

This Section provides additional information about the data used in the paper and the construction of variables.

I.A Datasets

We use panel data from the Panel Study of Income Dynamics (PSID), a national survey of U.S. households that is widely used in the household finance literature. The survey, which now tracks more than 7,000 U.S. households, is conducted on an annual basis from 1968 to 1997 and on a bi-annual basis from 1997 to 2013. It contains information on the households' income, demographics (age, household size, marital status, high school or post-high school education), and real estate holdings (value of the house, mortgage, tenure choice). We use an interpolation approach when information about some characteristics is missing. For example, values for total taxable family income are unavailable between 1994 and 1996 and in 2001, so we use linear interpolation from the first available surrounding years. Similarly, civil status is not available between 1994 and 1997.

In some years, the PSID also provides a special wealth survey that contains additional information on the households' wealth (e.g., holdings in stocks, bonds, and businesses). Prior to 1999, the wealth survey was only conducted in 1984, 1989, 1994. Starting in 1999, the wealth survey is part of the regular survey conducted by the PSID. Unless stated otherwise, our main empirical analysis is based on the years when the wealth survey is conducted.

In Sections 2 and 4.5 of the main text, we also use another database in the PSID that indicates the location of each survey participant at the Census Tract level between 1968 and 2009. This location data is available only by special request. Combining this information with data from the U.S. Census Bureau, we obtain a set of characteristics related to the neighborhood of each home. Specifically, we use the Neighborhood Change Database (NCDB) which provides instant access to most of the census variables for years 1980, 1990, and 2000 while adjusting for changes in the boundaries of census tracts over the years. We use linear interpolation to infer values for all the other years. For the years after 2000, we interpolate the data using the census years 1990 and 2000. For the rural areas that were not followed in the 1980s, we interpolate the data using the 1990 and 2000 census values.

Regarding the financial and inflation time series used throughout the main text, we use the value-weighted NYSE/AMEX/NASDAQ index and the 1-month Treasury Bill from CRSP as proxies for stocks and bonds. The parameters of the model in Section **??** are based on the estimated first and second return moments of annual log real returns between 1968 and 2013 when the PSID data is available. For the deflator, we use an inflation index that excludes the price of the housing good (as in the model). Following Piazzesi et al. (2007), we form this index from the NIPA consumption tables that excludes housing services, durable goods, cloth and shoes (Table 2.3.5 Personal Consumption Expenditure by Major Type of Product). It has a correlation of 84% with the standard CPI obtained from Robert Shiller's website over the 1968-2013 period.

I.B **Household Definition**

We impose the following criteria to characterize a household in the PSID:

- A household is represented by a head member. Each year, the PSID automatically selects one member of an interviewed family unit to be the head and it ranks all the other members in terms of their relationship to that person (e.g., partner, child, sibling). For multiple-member units, the PSID convention is that the head must be older than 16 years old and the person with the greatest financial responsibility. If that person is female and she has a partner (husband, boyfriend living in the same unit, or civil partner), then her partner is designated as the head, unless he is incapacitated. We add the restriction that the household head member must be older than 18 years and younger than 100 years. As defined by the PSID, the age, gender, and education variables all refer to the household head. The income and wealth variables are all aggregated at the household level.

- To make sure that a household can be properly compared over time, we impose that any serious change to the household composition results in the creation of new household(s). This allows us to conduct proper statistical inference as standard errors are clustered at the household level.

– A household no longer exists when there is a change in the marital status of the head couple. A change can come from either divorce, separation, death, or a new partnership in the case of a head member who used to be single. In the event any of the other household members keeps being interviewed by the PSID afterwards, we consider him/her as a member of a new household.

– A new household is created when a member of an existing household who is not the head or his partner (e.g. child, sibling) leaves and creates his or her own household.

– A new household is created when the PSID extends the sample to new families and interviews them for the first time.

I.C Portfolio Composition

In the wealth surveys, households are asked to report the market value of their assets via the following type of question: *if you sold all [the amount in asset x that you or anyone in your family own] and paid off anything you owed on it, how much would you have?* From their responses we compute the market value invested in each asset class.

Gross worth includes the net value invested in stocks (including mutual funds and retirement accounts), the net value invested in bonds (including the cash value in life insurance policies), the amount of cash (including checking and savings accounts, CDs, government savings bonds or Treasury bills), the value of the household's primary house, the net value of his other real estate properties, the net value invested in farms and private businesses, and the net value invested in cars. We also compute net worth as gross worth minus the amount remaining on non-collateralized debt (such as credit card charges, student loans, medical bills, or loans from relatives) and the mortgage on the household's primary house. Before 1999, the value invested in stocks included both retirement and non-retirement accounts. Afterwards, households were asked to report the values of both accounts separately. To be consistent with the pre-1999 years, we take the sum of these two accounts.

We measure the homeowner's *housing share* as the total investment in real estate out of gross worth. We select gross worth as the denominator because the measured net worth can be very small or even negative for the younger households with large educational loans. In these cases, the housing share is either arbitrarily large or undefined. However, as we show later in this Internet Appendix, the results do not change if we replace gross worth with net worth.

I.D House Prices

To build the time-series of home prices among households, we use the entire PSID data available from 1968 to 2013. We require at least ten consecutive annual observations for each house (37% of all houses in the sample). We eliminate observations where the average value and/or the volatility of the house's price growth is greater than 100% per year. For the years after 1997 where the PSID survey is only available every other year, the percentage changes in house prices are annualized. We compute the first and second moments of each individual house price over all the available years.

For the calibration of the model in Section **??** of the main text, we also compute the median of the mean, volatility, and stock correlation of the price growth across individual homes and use these values as inputs for the baseline case.

I.E Data Requirements

We impose a series of filters to guarantee the reliability of the data for our analysis in Sections 2 and 4 of the main text. To begin, we exclude households that belong to the Survey of Economic Opportunities, which is a special subsample of the PSID drawn from lower income levels that allows researchers to study poverty. Households must also appear in the survey more than once, have gross worth and real total taxable family income greater than $1,000 and strictly positive financial wealth (cash, stocks, and bonds). Any missing value for a given asset results in a missing value for gross worth. To eliminate outliers, we remove the top 0.1% values of gross worth and real estate wealth. For landlords, we remove the observations for which the values of rental income earned are in the top 0.1%.

In Section 4 of the main text, we restrict the sample to landlords and non-landlord homeowners which both have a strictly positive housing share. The resulting panel is unbalanced and includes 23,566 observations from 1984 to 2013. In Section 2 of the main text, the summary statistics are based on a larger sample that also includes both renters and homeowners.

II Additional Empirical Results

We now present the robustness tests discussed in Section 4.5 of the main text.

II.A Linear Specification

One potential concern with the baseline specification is that the true relationship between the set of characteristics and the housing share is not linear. If the characteristic values for landlords and non-landlords span different intervals, the local slope s_U estimated from landlord data could differ from the slope of the demand curve for non-landlords. This concern is most relevant for gross wealth because (i) it has the highest explanatory power among all characteristics, and (ii) landlords tend to be wealthier than non-landlords.

To examine the potential non-linearity of the wealth effect, we first perform a simple non-parametric regression of the housing share on gross wealth for both landlords and non-landlords. Each year, we sort both types of households into deciles based on their gross worth, and compute for each decile (i) the equally-weighted average share of real estate out of gross worth and (ii) the equally-weighted average log gross worth. We then compute the time-series averages for the housing share and gross worth over the years when the survey is conducted. The results are shown in Figure IA.1, where landlords and non-landlords are depicted in grey and black respectively.

Our analysis reveals several insights. First, 80% of the wealth observations for non-landlords fall in the interval computed with landlords' wealth data. We can therefore rule out the concern that the group of landlords is a non-representative sample clustered at the top of the wealth distribution. Second, we document a mild degree of non-linearity as the slope coefficient becomes flatter for low levels of wealth. Third, the regression lines are remarkably stable for the bottom tercile of the wealth distribution. This stability suggests that the regression results are not driven by potential outliers in the extreme tails of the wealth distribution for landlords and non-landlords.

We then explicitly control for the non-linearity in wealth by adding additional interaction terms for the less fortunate households. For each group (landlords, non-landlords), we introduce a variable equal to the wealth of landlords if their wealth is below the group's median wealth in year t, and zero otherwise. The results in Table IA.1 show that both interaction terms have statistically insignificant coefficients and a minor impact on the main results. The adjusted R^2 only increases by 0.30% and the unconstrained demand remains the primary driving factor of the variance decomposition (78% vs 7% for the constrained demand).

II.B Infrequent Rebalancing

Another potential concern is that the negative wealth effect is due to high trading costs that prevent households from rebalancing their housing shares frequently. A short-term increase in the level of financial wealth that is unrelated to housing could mechanically lead to a low housing share and vice-versa. We estimate two alternate specifications that eliminate this mechanical link. First, we replace gross wealth with its lagged value in the set of characteristics. Second, we only retain the observations following a home purchase. In both cases, Table IA.2 and IA.3 show that the variance decomposition remains unchanged. The unconditional demand explains either 81% or 75% of the cross-sectional variation in the total demand, versus 11% and 13% for the constrained demand. We also see that the coefficients associated with lagged wealth remain negative and statistically significant for both the unconstrained and total demands. In Table IA.3, the fit of the regression even reaches above 30%.

II.C Omitted Variables

We control for potential omitted variables by including in the main regression specification a new set of neighborhood characteristics (at the census tract level) that provide additional information regarding the characteristics of households and their homes. We control for the fractions of homes that are owner-occupied, vacant, recreational, and for sale, the local unemployment rate, the average education rate, and the average number of years spent by the household in the neighborhood. We also include dummies regarding the ethnicity of each household (White/Caucasian, Asian, African-American, Hispanic), as well as the local fraction of households of each ethnic background. We see in Table IA.4 that the introduction of these variables leaves our main results unchanged and slightly reduces the fit of the regression (i.e., the adjusted R^2 declines from 27.43% to 25.26%).

II.D Net Worth

In Section 4.1 of the main text, we motivate the use of gross worth for our baseline analysis. We now re-estimate the coefficients of the panel regression where the dependent variable is the share of housing out of net worth. We impose a maximum share of housing out of net worth of 500% to eliminate outlier observations. The variance decomposition reported in Table IA.5 confirms that the unconstrained demand remains the main determinant of the housing share (48% versus 14% for the constrained demand). The main difference comes from the covariance term which explains 38% of the variance of the housing share (versus 14% in the baseline case). We also find that in 85% of the cases, the signs of the coefficients remain the same regardless of whether net or gross worth is used to compute the housing share.

One difference in Table IA.5 is that the relation between the household income and the total housing share turns positive. This result is consistent with intuition as household with higher income have an easier access to mortgage debt and thus exhibit a higher housing share. For these households, the housing share based on net worth potentially overestimates the importance of housing in the portfolio because it does not account for the household's human capital wealth.

III Derivation of the Portfolio Choice Model

III.A Setup

We consider a dynamic model in which the household consumes a basket of two goods: a perishable non-housing good (C_t) and a durable housing good (K_t). Here, a unit of housing is a one-dimensional summary of the quality of the house which accounts for size, location, and specific characteristics. At each time t, the utility derived by the household is the Cobb-Douglas function

$$U(C_t, K_t) = \frac{1}{1-\gamma} \left(C_t^{\beta} K_t^{1-\beta} \right)^{1-\gamma},$$

where γ is the coefficient of relative risk aversion over the entire consumption basket and β is the degree to which the household values non-housing consumption relative to housing consumption. If we further assume that the utility stream is additively separable and the time horizon infinite, the household's lifetime expected utility is given by

$$E\left[\int_0^{\infty} e^{-\delta s} U(C_s, K_s) ds \right],$$

where δ is the time discount factor.

The household can invest in a short-term risk-free bond and a risky stock fund without any short-sales restrictions. Using the non-housing good as numeraire, we denote by r the constant risk-free rate of the bond and write the dynamics of the stock price $P_{S,t}$ as

$$\frac{dP_{S,t}}{P_{S,t}} = \mu_S \, dt + \sigma_S \, dZ_{1,t},$$

where μ_S and σ_S are constants, and $Z_{1,t}$ is a Wiener process. The household can also invest in H_t units of housing whose price $P_{H,t}$ is stochastic,

$$\frac{dP_{H,t}}{P_{H,t}} = \mu_H \, dt + \sigma_{H,1} \, dZ_{1,t} + \sigma_{H,2} \, dZ_{2,t},$$

where μ_H, $\sigma_{H,1}$, and $\sigma_{H,2}$ are constants, and $Z_{2,t}$ is a Wiener process that is uncorrelated with $Z_{1,t}$. We assume that, as a durable good, housing does not depreciate over time so that $dH_t = 0$ if the agent does not trade. Therefore, the expected return on housing can be viewed as net of maintenance costs.

The household can bridge the difference between H_t and K_t via the rental market. We assume a constant rent-price ratio ρ, which means that the price of renting one unit of housing is $\rho \, P_{H,t}$. Consequently, the net amount spent on the rent is $(K_t - H_t) \, \rho \, P_{H,t}$. We define ϕ_t as the ratio of housing investment to housing consumption, i.e. $\phi_t = H_t / K_t$. The household is a partial homeowner if $0 < \phi_t < 1$ and a full homeowner if $\phi_t = 1$.

Combining this information, we can write the dynamics of the wealth process as

$$dW_t = \left[rW_t + \Theta_t' (\mu - r\mathbf{1}) - C_t - (K_t - H_t) \, \rho P_{H,t} \right] dt + W_t \, \Theta_t' \, \sigma \begin{pmatrix} dZ_{1,t} \\ dZ_{2,t} \end{pmatrix}, \qquad \text{(IA-1)}$$

where $\Theta_t = [\Theta_{S,t}, P_H H_t]'$ is the vector of wealth invested in stocks and housing at time t, and the parameters μ are σ are given by

$$\mu = \begin{pmatrix} \mu_S \\ \mu_H \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_{S,1} & 0 \\ \sigma_{H,1} & \sigma_{H,2} \end{pmatrix}. \tag{IA-2}$$

III.B Format of the Solution

The infinite-horizon setting is convenient because portfolio decisions depend on preferences and state variables but not on time. For notational convenience, we therefore eliminate the time subscript from now on. At time t, the household's value function V is given as,

$$V(W, P_H) = \sup_{\Gamma} \int_0^{\infty} e^{-\delta s} U(C_\tau, K_\tau) ds,$$

where $\Gamma = \{\Theta, K, C\}$ is the set of admissible controls.

The Hamilton-Jacobi-Bellman (HJB) equation of this problem can be written as

$$\delta V(W, P_H) = \sup_{\Gamma} \Big[U(C, K) + E\left[dV(W, P_H) \right] \Big]. \tag{IA-3}$$

We begin by expanding Equation (IA-3),

$$\begin{aligned} \delta V = \sup_{\Gamma} \Big[& U(C, K) + V_W E(dW) + \frac{1}{2} V_{WW} E(dW)^2 \\ & + V_{P_H} E(dP_H) + \frac{1}{2} V_{P_H P_H} E(dP_H)^2 + V_{W P_H} E(dW dP_H) \Big]. \end{aligned} \tag{IA-4}$$

We then plug in the processes for dW and dP_H,

$$\begin{aligned} \delta V = \sup_{\Gamma} \Big[& U(C, K) + V_W \Big(rW + \Theta'(\mu - r\mathbf{1}) - C - (K - H)\rho P_H \Big) + \frac{1}{2} V_{WW} \Theta' \sigma \sigma' \Theta \\ & + V_{P_H} \mu_H P_H + \frac{1}{2} V_{P_H P_H} \left\| \sigma_H \right\|^2 P_H^2 + V_{W P_H} \Theta' \sigma \sigma_H P_H \Big], \end{aligned} \tag{IA-5}$$

where σ_H is the vector of coefficients $\sigma_{H,1}$ and $\sigma_{H,2}$.

Building on the work of Damgaard, Fuglsbjerg and Munk (2003), we guess that the value function has the form

$$V(W, P_H) = \frac{1}{1-\gamma} \kappa P_H^{-(1-\beta)(1-\gamma)} W^{1-\gamma}. \tag{IA-6}$$

where κ is a constant. Since V is homogeneous, we can set $V(W, P_H) = P_H^{\beta(1-\gamma)} v(\tilde{W})$, where $v(\tilde{W}) = \frac{1}{1-\gamma}\kappa \tilde{W}^{1-\gamma}$, and $\tilde{W} = W/P_H$. Therefore, all the derivatives of $V(W, P_H)$ can be re-expressed in terms of $v(\tilde{W})$,

$$\begin{aligned} V_W &= P_H^{\beta(1-\gamma)-1} v_{\tilde{W}}, & V_{WW} &= P_H^{\beta(1-\gamma)-2} v_{\tilde{W}\tilde{W}}, \\ V_{P_H} &= P_H^{\beta(1-\gamma)-1} v_{P_H}, & V_{P_H P_H} &= P_H^{\beta(1-\gamma)-2} v_{P_H P_H}, \\ V_{W P_H} &= P_H^{\beta(1-\gamma)-1} v_{\tilde{W} P_H}, \end{aligned}$$

where, with a slight abuse of notation, the derivatives of $v(\tilde{W})$ with respect to P_H are really just functions of $\tilde{W}$,

$$\begin{aligned} v_{P_H} &= \beta(1-\gamma)v - \tilde{W} v_{\tilde{W}}, \\ v_{P_H P_H} &= (\beta(1-\gamma)-1)\beta(1-\gamma)v - 2(\beta(1-\gamma)-1)\tilde{W} v_{\tilde{W}} + \tilde{W}^2 v_{\tilde{W}\tilde{W}}, \\ v_{\tilde{W} P_H} &= (\beta(1-\gamma)-1)v_{\tilde{W}} - \tilde{W} v_{\tilde{W}\tilde{W}}. \end{aligned}$$

Let $\tilde{C} = C/P_H$, $\tilde{\Theta}_S = \Theta_S/P_H$, and $\tilde{\Theta} = [\tilde{\Theta}_S, H_t]'$. Once we apply the change of variables to Equation (IA-5), the term $P_H^{\beta(1-\gamma)}$ cancels out and we are left with

$$\begin{aligned} \delta v = \sup_{\Gamma} \Bigg[& U(\tilde{C}, K) + v_{\tilde{W}} \left(r\tilde{W} + \tilde{\Theta}'(\mu - r\mathbf{1}) - \tilde{C} - (K-H)\rho \right) + \frac{1}{2} v_{\tilde{W}\tilde{W}} \tilde{\Theta}'\sigma\sigma'\tilde{\Theta} \\ & + v_{P_H}\mu_H + \frac{1}{2} v_{P_H P_H} \left\| \sigma_H \right\|^2 + v_{\tilde{W} P_H} \tilde{\Theta}'\sigma\sigma_H \Bigg]. \end{aligned} \tag{IA-7}$$

The modified HJB equation (IA-5) is now independent of P_H.

Before we solve for the optimal controls for the unconstrained and constrained households, we apply one additional change of variables. Let $\alpha_C = C/W$, $\alpha_K = KP_H/W$ denote the fractions of wealth spent on non-housing and housing consumption, and $\alpha_\Theta = [\alpha_S, \alpha_H]' = \Theta/W$ denote the shares invested in stocks and housing. It follows that $\alpha_C = \tilde{C}/\tilde{W}$, $\alpha_K = K/\tilde{W}$, and $\alpha_\Theta = \tilde{\Theta}/\tilde{W}$. The HJB equation (IA-5) becomes

$$\begin{aligned} \delta v = \sup_{\Gamma} \Bigg[& \tilde{W}^{1-\gamma} U(\alpha_C, \alpha_K) - v_{\tilde{W}} \tilde{W} \left(\alpha_C + \alpha_K \rho + r + \alpha_\Theta'(\bar{\mu} - r\mathbf{1}) \right) \\ & + \frac{1}{2} v_{\tilde{W}\tilde{W}} \tilde{W}^2 \alpha_\Theta' \sigma\sigma' \alpha_\Theta + v_{P_H}\mu_H + \frac{1}{2} v_{P_H P_H} \left\| \sigma_H \right\|^2 + v_{\tilde{W} P_H} \tilde{W} \alpha_\Theta' \sigma\sigma_H \Bigg], \end{aligned} \tag{IA-8}$$

where

$$\bar{\mu} = \begin{pmatrix} \mu_S \\ \mu_H + \rho \end{pmatrix}.$$

III.C Solution for the Unconstrained Household

We begin by deriving the optimal decisions of the household that has the full choice set $\Gamma = \{\Theta, K, C\}$. The optimal shares and the value function v are denoted by the superscript U (for unconstrained). The first order conditions with respect to α_C, α_K and α_Θ are

$$U_{\alpha_C^U} = v_{\tilde{W}}^U \tilde{W}^\gamma, \tag{IA-9}$$

$$U_{\alpha_K^U} = \rho v_{\tilde{W}}^U \tilde{W}^\gamma, \tag{IA-10}$$

$$\alpha_\Theta^U = -\frac{v_{\tilde{W}}^U}{v_{\tilde{W}\tilde{W}}^U \tilde{W}} \left(\sigma\sigma'\right)^{-1} \left(\bar{\mu} - r\mathbf{1}\right) - \frac{v_{\tilde{W}P_H}^U}{v_{\tilde{W}\tilde{W}}^U \tilde{W}} \left(\sigma'\right)^{-1} \sigma_H. \tag{IA-11}$$

where U_{α_C} and U_{α_K} are the marginal utilities of $U(\alpha_C, \alpha_K)$ with respect to α_C and α_K respectively. The optimal shares in stocks, bonds, and housing in Proposition 1 in the main text directly follow from Equation (IA-11).

Next, we solve for the optimal level of consumption for the housing and non-housing goods. After merging Equations (IA-9) and (IA-10), we obtain

$$\alpha_Q^U = \frac{\alpha_C^U}{\alpha_K^U} = \frac{\beta}{1-\beta}\rho,$$

which implies that the ratio of non-housing consumption to housing consumption only depends on the rental price of housing. We can then re-express α_C^U and α_K^U as

$$\alpha_C^U = \beta \left(v_{\tilde{W}}^U\right)^{-\frac{1}{\gamma}} \tilde{W}^{-1}\epsilon,$$
$$\alpha_K^U = \frac{1}{\rho}(1-\beta) \left(v_{\tilde{W}}^U\right)^{-\frac{1}{\gamma}} \tilde{W}^{-1}\epsilon.$$

where $\epsilon = \beta^{\frac{1-\gamma}{\gamma}} \left(\alpha_Q^U\right)^{-\frac{(1-\beta)(1-\gamma)}{\gamma}}$. We make one final change of variables:

$$\bar{\beta} = \beta(1-\gamma),$$
$$\bar{r} = r - \mu_H(1-\beta) + \frac{1}{2}\left\|\sigma_H\right\|^2 \left((\bar{\beta}-1)(\beta-2)-\gamma\right),$$
$$\bar{\lambda} = \lambda - (1-\gamma)(1-\beta)\sigma_H,$$

where $\lambda = \sigma^{-1}(\mu - r\mathbf{1})$ is the vector of the market prices of risk for the two shocks Z_1 and

Z_2. Inserting the optimal controls into the HJB Equation (IA-8), we get

$$\begin{aligned} 0 = &-v^U \left(\delta - r - \mu_H \bar{\beta} - \frac{1}{2} \|\sigma_H\|^2 \bar{\beta} \left(\bar{\beta} - 1\right)\right) \\ &+ \frac{\gamma}{1-\gamma} \left(v^U_{\tilde{W}}\right)^{\frac{\gamma-1}{\gamma}} \epsilon + v^U_{\tilde{W}} \tilde{W}(r - \mu_H) - \frac{1}{2} \frac{\left(v^U_{\tilde{W}}\right)^2}{v^U_{\tilde{W}\tilde{W}}} \|\lambda\|^2 , \end{aligned} \qquad \text{(IA-12)}$$

where $v^U = \frac{1}{1-\gamma} \kappa^U \tilde{W}^{1-\gamma}$. If we divide Equation (IA-12) by $\tilde{W}^{1-\gamma}$ and $\frac{\gamma}{1-\gamma}$, the $\tilde{W}$ term cancels out and the value of κ^U can be written as

$$\kappa^U = \left(\frac{\epsilon}{\frac{\delta}{\gamma} - \frac{1-\gamma}{\gamma}\bar{r} - \frac{1-\gamma}{2\gamma^2}\left\|\bar{\lambda}\right\|^2}\right)^{\gamma}.$$

III.D Solution for the Constrained Household

We now incorporate the homeownership constraint in the optimization. Specifically, we impose the household to own what he chooses to consume of housing, i.e., $K_t = H_t$, or equivalently $\alpha_K = \alpha_H$. In other words, the constrained household only has access to the reduced choice set $\Gamma = \{\Theta, C\}$. We denote the household's optimal shares and its value function by the superscript T (for total) and any difference between the constrained and unconstrained weights by superscript C (for constrained), e.g., $\alpha^C_S = \alpha^T_S - \alpha^U_S$. Using Equation (IA-7), we can write the first order conditions with respect to α_C, α_K and α_S as

$$U_{\alpha^T_C} = v^T_{\tilde{W}} \tilde{W}^{\gamma}, \qquad \text{(IA-13)}$$

$$\alpha^T_K = \alpha^U_H - \frac{1}{v^T_{\tilde{W}\tilde{W}} \tilde{W}^2 \sigma^2_{H_2}} \left(U_{\alpha^T_K} - v^T_{\tilde{W}} \tilde{W}\right), \qquad \text{(IA-14)}$$

$$\begin{aligned} \alpha^T_S &= \alpha^U_S + \frac{\sigma_{H_1}}{\sigma_{S_1}} \frac{1}{v^T_{\tilde{W}\tilde{W}} \tilde{W}^2 \sigma^2_{H_2}} \left(U_{\alpha^T_K} - v^T_{\tilde{W}} \tilde{W}\right), \\ &= \alpha^U_S - \frac{\sigma_{H_1}}{\sigma_{S_1}} \alpha^C_H. \end{aligned} \qquad \text{(IA-15)}$$

The degree to which the consumption decisions differ from those of the unconstrained household also depends on the impact of the homeownership constraint on housing investment as measured by α^C_H. Merging Equations (IA-13) and (IA-14), we obtain

$$\alpha^T_Q = \frac{\alpha^T_C}{\alpha^T_K} = \alpha^U_Q - \frac{\beta}{1-\beta} \frac{v^T_{\tilde{W}\tilde{W}} \tilde{W}}{v^T_{\tilde{W}}} \sigma^2_{H_2} \alpha^C_H. \qquad \text{(IA-16)}$$

The optimization of the constrained household's problem is identical to that in Damgaard, Fuglsbjerg and Munk (2003). We could therefore proceed with their solution, which consists of inserting the optimal controls above back into Equation (IA-8) and deriving an ODE. Here, we derive an alternative formulation for the solution that allows us to decompose the total demand for real estate $\alpha_K^T = \alpha_H^T$ into its unconstrained and constrained components, α_H^U and α_H^C. To begin, we re-express Equation (IA-8) in terms of $\alpha_\Theta^U = [\alpha_S^U, \alpha_H^U]'$ and $\alpha_\Theta^C = [\alpha_S^C, \alpha_H^C]'$,

$$\begin{aligned}\delta v^T &= \tilde{W}^{1-\gamma} U(\alpha_C^T, \alpha_K^T) - v_{\tilde{W}}^T \hat{W} \left(\alpha_C^T + \alpha_K^T\right) \\ &+ v_{\tilde{W}}^T \tilde{W} r - \frac{1}{2} v_{\tilde{W}\tilde{W}}^T \tilde{W}^2 \alpha_\Theta^{U\prime} \sigma\sigma' \alpha_\Theta^U + \frac{1}{2} v_{\tilde{W}\tilde{W}}^T \tilde{W}^2 \alpha_\Theta^{C\prime} \sigma\sigma' \alpha_\Theta^C \\ &+ v_{P_H}^T \mu_H + \frac{1}{2} v_{P_H P_H}^T \left\|\sigma_H\right\|^2 . \end{aligned} \tag{IA-17}$$

We then split the right-hand side of Equation (IA-17) into two terms: one term T^U that only includes the optimal weights of the unconstrained household, and another term T^C that includes all the remaining elements:

$$\delta v^T = T^U + T^C. \tag{IA-18}$$

Both terms are given by

$$\begin{aligned}T^U &= \tilde{W}^{1-\gamma} \xi U\left(\alpha_C^U, \alpha_K^U\right) - v_{\tilde{W}}^T \tilde{W}\left(\alpha_C^U + \alpha_K^U \rho\right) + v_{\tilde{W}}^T \tilde{W} r - \frac{1}{2} v_{\tilde{W}\tilde{W}}^T \tilde{W}^2 \alpha_\Theta^{U\prime} \sigma\sigma' \alpha_\Theta^U \\ &+ v_{P_H}^T \mu_H + \frac{1}{2} v_{P_H P_H}^T \left\|\sigma_H\right\|^2 , \end{aligned} \tag{IA-19}$$

and

$$\begin{aligned}T^C &= \tilde{W}^{1-\gamma} \left[U\left(\alpha_C^T, \alpha_K^T\right) - \xi U\left(\alpha_C^U, \alpha_K^U\right)\right] \\ &- v_{\tilde{W}}^T \tilde{W}\left[\left(\alpha_C^T + \alpha_K^T \rho\right) - \left(\alpha_C^U + \alpha_K^U \rho\right)\right] + \frac{1}{2} v_{\tilde{W}\tilde{W}}^T \tilde{W}^2 \alpha_\Theta^{C\prime} \sigma\sigma' \alpha_\Theta^C, \end{aligned} \tag{IA-20}$$

where ξ is the constant that links the value functions of the unconstrained and constrained households, i.e., $v^T = \xi v^U$. We see that Equation (IA-19) is the HJB Equation (IA-8) for the unconstrained household, which implies that $\delta v^T = T^U$ or $T^C = 0$. We then insert Equations (IA-13) and (IA-15) into Equation (IA-20), and divide all the terms by $\tilde{W}^{1-\gamma}$ and κ^T to obtain the following quadratic equation in terms of α_H^C,

$$A_q \left(\alpha_H^C\right)^2 + B_q \alpha_H^C + C_q = 0, \tag{IA-21}$$

where the parameters A_q, B_q, and C_q are defined as

$$A_q = \frac{\gamma\sigma_{H_2}^2}{\rho}\left(1 + \frac{(1-\beta)(1-\gamma)}{2\gamma}\right), \quad B_q = 1 + \frac{\alpha_H^U\sigma_{H_2}^2}{\rho}\left(1-\beta(1-\gamma)\right), \quad C_q = \alpha_H^U\left(1-\frac{1}{\phi^U}\right).$$

The existence of a solution requires that $B_q^2 - 4A_q\,C_q \geq 0$. As shown by Damgaard, Fuglsbjerg and Munk (2003), only one root leads to positive values for α_C^T and α_K^T. The proportionality constant ξ is equal to the ratio of the marginal utilities over non-housing consumption in the unconstrained and constrained cases,

$$\xi = \frac{v_{\tilde{W}}^T}{v_{\tilde{W}}^U} = \left(\frac{\alpha_C^T}{\alpha_C^U}\right)^{\beta(1-\gamma)-1}\left(\frac{\alpha_K^T}{\alpha_K^U}\right)^{(1-\beta)(1-\gamma)}.$$

REFERENCES

Damgaard, A., Fuglsbjerg, B. and Munk, C. (2003), 'Optimal consumption and investment strategies with a perishable and an indivisible durable good', *Journal of Economic Dynamics and Control* **28**(2), 209–253.

Piazzesi, M., Schneider, M. and Tuzel, S. (2007), 'Housing, consumption, and asset pricing', *Journal of Financial Economics* **83**, 531–569.

Table IA.1
Panel Regression of Housing Shares
Non-Linear Wealth Effects

The table reports the panel regression of the housing share $\alpha_{H,i,t}$ on the vector $X_{i,t}$, which includes (i) the baseline characteristics, and (ii) an interacted variable that is equal to log gross worth if the household is below the sample median of its group (landlord/non-landlord) in year t and zero otherwise. The regression is estimated between 1984 and 2013 using micro-level data from the PSID,

$$\alpha_{H,i,t} = a_t + s'_U X_{i,t} + (b_t + s'_C X_{i,t}) D_{i,t}^{NL} + \varepsilon_{i,t},$$

where $D_{i,t}^{NL}$ is a dummy variable that equals 1 if the household is a non-landlord homeowner and 0 otherwise. We refer to s_U and s_C as the slope vectors of the unconstrained and constrained components of the housing demand. We refer to $s_T = s_U + s_C$ as the slope vector of the total housing demand. Panel A reports, for each demand component, (i) the variance ratio of its predicted values relative to those of the total demand (e.g., $var(s'_U X_{i,t}) / var(s'_T X_{i,t})$), (ii) the standard deviation of its predicted values (e.g., $stdev(s'_U X_i)$), and (iii) the statistical significance of the hypothesis test that its slope coefficients are jointly equal to zero, (e.g., H_o: $s_U = 0$). Panel B reports the individual slope coefficients for s_T, s_U, and s_C and their t-statistics in separate columns. We cluster observations for each household to compute the variance of the coefficients, and standardize each coefficient by the standard deviation of its associated characteristic among non-landlords.

Panel A: Variance Decomposition and Joint Tests

	Ratio of Variances	Std. Dev.	Joint Test	
			F-stat	p-value
Total demand	100%	0.123	107.360	<.0001
Unconstrained demand	78%	0.109	21.870	<.0001
Constrained demand	7%	0.032	1.860	0.047

Panel B: Regression Slope Coefficients of the Housing Share

	s_T		s_U		s_C	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Financial Characteristics						
Household income (log)	-0.001	-0.29	0.008	0.84	-0.010	-0.98
Gross worth (log)	-0.101	-20.25	-0.103	-8.57	0.002	0.15
Gross Worth - low wealth (log)	0.002	5.38	0.001	1.19	0.001	0.86
Demographic Characteristics						
Age	-0.019	-5.89	-0.008	-0.71	-0.011	-1.02
Household size	0.035	13.28	0.019	2.18	0.016	1.78
Married dummy	-0.015	-4.76	-0.019	-1.89	0.004	0.39
High school dummy	-0.012	-4.35	-0.024	-2.66	0.013	1.42
Post-high school dummy	0.000	-0.10	-0.008	-1.05	0.008	1.07
House Price Characteristics						
Average growth	0.023	5.91	0.012	1.15	0.012	1.13
Volatility	-0.031	-7.62	-0.018	-1.87	-0.014	-1.36
Adjusted R-square	27.70%					
Number of observations	23,566					

Table IA.2
Panel Regression of Housing Shares
Lagged Wealth

The table reports the panel regression of the housing share $\alpha_{H,i,t}$ on the vector $X_{i,t}$, which includes (i) log gross worth in year *t*-1, and (ii) the remaining baseline characteristics in year *t.* The regression is estimated between 1984 and 2013 using micro-level data from the PSID,

$$\alpha_{H,i,t} = a_t + s'_U X_{i,t} + (b_t + s'_C X_{i,t}) D_{i,t}^{NL} + \varepsilon_{i,t},$$

where $D_{i,t}^{NL}$ is a dummy variable that equals 1 if the household is a non-landlord homeowner and 0 otherwise. We refer to s_U and s_C as the slope vectors of the unconstrained and constrained components of the housing demand. We refer to $s_T = s_U + s_C$ as the slope vector of the total housing demand. Panel A reports, for each demand component, (i) the variance ratio of its predicted values relative to those of the total demand (e.g., $var(s'_U X_{i,t}) / var(s'_T X_{i,t})$), (ii) the standard deviation of its predicted values (e.g., $stdev(s'_U X_i)$), and (iii) the statistical significance of the hypothesis test that its slope coefficients are jointly equal to zero, (e.g., H_o: $s_U = 0$). Panel B reports the individual slope coefficients for s_T, s_U, and s_C and their t-statistics in separate columns. We cluster observations for each household to compute the variance of the coefficients, and standardize each coefficient by the standard deviation of its associated characteristic among non-landlords.

Panel A: Variance Decomposition and Joint Tests

	Ratio of Variances	Std. Dev.	Joint Test	
			F-stat	p-value
Total demand	100%	0.099	64.970	<.0001
Unconstrained demand	81%	0.089	12.180	<.0001
Constrained demand	11%	0.033	1.220	0.277

Panel B: Regression Slope Coefficients of the Housing Share

	s_T		s_U		s_C	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Financial Characteristics						
Household income (log)	-0.046	-10.25	-0.030	-2.01	-0.018	-1.18
Lagged Gross worth (log)	-0.062	-15.72	-0.069	-6.05	0.007	0.62
Demographic Characteristics						
Age	-0.002	-8.40	-0.001	-0.57	-0.002	-1.86
Household size	0.029	13.51	0.025	3.16	0.004	0.55
Married dummy	-0.036	-4.16	-0.057	-2.09	0.022	0.78
High school dummy	-0.012	-1.82	-0.050	-2.28	0.041	1.83
Post-high school dummy	-0.013	-1.70	-0.028	-1.26	0.016	0.71
House Price Characteristics						
Average growth	0.326	4.04	0.300	1.39	0.028	0.12
Volatility	-0.174	-4.83	-0.217	-2.43	0.046	0.48
Adjusted R-square	17.94%					
Number of observations	19,324					

Table IA.3
Panel Regression of Housing Shares
Subsample of Households that Recently Moved

The table reports the panel regression of the housing share $\alpha_{H,i,t}$ on the baseline vector of characteristics $X_{i,t}$ estimated between 1984 and 2013 on the subsample of households in the PSID that just moved to a new home,

$$\alpha_{H,i,t} = a_t + s'_U X_{i,t} + (b_t + s'_C X_{i,t}) D_{i,t}^{NL} + \varepsilon_{i,t,}$$

where $D_{i,t}^{NL}$ is a dummy variable that equals 1 if the household is a non-landlord homeowner and 0 otherwise. We refer to s_U and s_C as the slope vectors of the unconstrained and constrained components of the housing demand. We refer to $s_T = s_U + s_C$ as the slope vector of the total housing demand. Panel A reports, for each demand component, (i) the variance ratio of its predicted values relative to those of the total demand (e.g., $var(s'_U X_{i,t}) / var(s'_T X_{i,t})$), (ii) the standard deviation of its predicted values (e.g., $stdev(s'_U X_i)$), and (iii) the statistical significance of the hypothesis test that its slope coefficients are jointly equal to zero, (e.g., H_o: $s_U = 0$). Panel B reports the individual slope coefficients for s_T, s_U, and s_C and their t-statistics in separate columns. We cluster observations for each household to compute the variance of the coefficients, and standardize each coefficient by the standard deviation of its associated characteristic among non-landlords.

Panel A: Variance Decomposition and Joint Tests

	Ratio of Variances	Std. Dev.	Joint Test	
			F-stat	p-value
Total demand	100%	0.116	31.510	<.0001
Unconstrained demand	75%	0.100	6.590	<.0001
Constrained demand	13%	0.042	1.050	0.395

Panel B: Regression Slope Coefficients of the Housing Share

	s_T		s_U		s_C	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Financial Characteristics						
Household income (log)	0.030	3.27	0.042	1.32	-0.013	-0.38
Gross worth (log)	-0.120	-15.95	-0.122	-4.75	0.002	0.06
Demographic Characteristics						
Age	-0.002	-5.42	0.000	0.15	-0.003	-1.57
Household size	0.022	5.97	0.023	1.66	-0.002	-0.12
Married dummy	-0.019	-1.35	-0.060	-1.21	0.044	0.86
High school dummy	-0.034	-3.30	-0.016	-0.41	-0.019	-0.48
Post-high school dummy	0.006	0.57	0.028	0.65	-0.022	-0.50
House Price Characteristics						
Average growth	-0.292	-2.51	-0.369	-1.34	0.081	0.27
Volatility	-0.120	-1.83	0.018	0.11	-0.151	-0.86
Adjusted R-square	31.78%					
Number of observations	2,466					

Table IA.4
Panel Regression of Housing Shares
Neighborhood-Specific Regressors

The table reports the panel regression of the housing share $\alpha_{H,i,t}$ on the vector $X_{i,t}$, which includes (i) the baseline characteristics, and (ii) an additional set of neighborhood-specific characteristics. The regression is estimated between 1984 and 2009 using micro-level data from the PSID,

$$\alpha_{H,i,t} = a_t + s'_U X_{i,t} + (b_t + s'_C X_{i,t}) D_{i,t}^{NL} + \varepsilon_{i,t},$$

where $D_{i,t}^{NL}$ is a dummy variable that equals 1 if the household is a non-landlord homeowner and 0 otherwise. We refer to s_U and s_C as the slope vectors of the unconstrained and constrained components of the housing demand. We refer to $s_T = s_U + s_C$ as the slope vector of the total housing demand. Panel A reports, for each demand component, (i) the variance ratio of its predicted values relative to those of the total demand (e.g., $var(s'_U X_{i,t}) / var(s'_T X_{i,t})$), (ii) the standard deviation of its predicted values (e.g., $stdev(s'_U X_i)$), and (iii) the statistical significance of the hypothesis test that its slope coefficients are jointly equal to zero, (e.g., H_o: $s_U = 0$). Panel B reports the individual slope coefficients for s_T, s_U, and s_C (baseline characteristics only) and their t-statistics in separate columns. We cluster observations for each household to compute the variance of the coefficients, and standardize each coefficient by the standard deviation of its associated characteristic among non-landlords.

Panel A: Variance Decomposition and Joint Tests

	Ratio of Variances	Std. Dev.	Joint Test	
			F-stat	p-value
Total demand	100%	0.117	33.630	<.0001
Unconstrained demand	102%	0.119	27.800	<.0001
Constrained demand	19%	0.052	2.630	<.0001

Panel B: Regression Slope Coefficients of the Housing Share

	s_T		s_U		s_C	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Financial Characteristics						
Household income (log)	-0.021	-4.44	-0.008	-0.70	-0.015	-1.18
Gross worth (log)	-0.106	-20.63	-0.111	-10.16	0.004	0.32
Demographic Characteristics						
Age	-0.002	-7.48	-0.001	-1.41	-0.001	-1.01
Household size	0.018	8.55	0.004	0.56	0.014	2.00
Married dummy	-0.015	-1.90	-0.006	-0.25	-0.010	-0.40
High school dummy	-0.002	-0.31	-0.025	-1.36	0.024	1.29
Post-high school dummy	-0.003	-0.44	-0.007	-0.37	0.003	0.15
House Price Characteristics						
Average growth	0.605	7.79	0.366	2.05	0.252	1.31
Volatility	-0.293	-9.21	-0.148	-2.03	-0.157	-1.99
Adjusted R-square	25.26%					
Number of observations	18,133					

Table IA.5
Panel Regression of Housing Shares
Net Worth

The table reports the panel regression of the housing share $\alpha_{H,i,t}$ measured using net worth on the baseline vector of characteristics $X_{i,t}$. The regression is estimated between 1984 and 2013 using micro-level data from the PSID,

$$\alpha_{H,i,t} = a_t + s'_U X_{i,t} + (b_t + s'_C X_{i,t}) D_{i,t}^{NL} + \varepsilon_{i,t,}$$

where $D_{i,t}^{NL}$ is a dummy variable that equals 1 if the household is a non-landlord homeowner and 0 otherwise. We refer to s_U and s_C as the slope vectors of the unconstrained and constrained components of the housing demand. We refer to $s_T = s_U + s_C$ as the slope vector of the total housing demand. Panel A reports, for each demand component, (i) the variance ratio of its predicted values relative to those of the total demand (e.g., $var(s'_U X_{i,t})$ / $var(s'_T X_{i,t})$), (ii) the standard deviation of its predicted values (e.g., $stdev(s'_U X_i)$), and (iii) the statistical significance of the hypothesis test that its slope coefficients are jointly equal to zero, (e.g., H_o: $s_U = 0$). Panel B reports the individual slope coefficients for s_T, s_U, and s_C and their t-statistics in separate columns. We cluster observations for each household to compute the variance of the coefficients, and standardize each coefficient by the standard deviation of its associated characteristic among non-landlords.

Panel A: Variance Decomposition and Joint Tests

	Ratio of Variances	Std. Dev.	Joint Test	
			F-stat	p-value
Total demand	100%	0.448	132.670	<.0001
Unconstrained demand	48%	0.312	26.960	<.0001
Constrained demand	14%	0.169	9.320	<.0001

Panel B: Regression Slope Coefficients of the Housing Share

	s_T		s_U		s_C	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Financial Characteristics						
Household income (log)	0.060	4.67	0.046	1.75	0.015	0.54
Gross worth (log)	-0.303	-27.65	-0.247	-11.52	-0.063	-2.65
Demographic Characteristics						
Age	-0.018	-25.49	-0.009	-6.12	-0.009	-5.37
Household size	0.049	5.95	0.059	2.61	-0.010	-0.41
Married dummy	-0.059	-2.38	-0.089	-1.79	0.030	0.54
High school dummy	-0.067	-3.45	-0.102	-2.63	0.036	0.85
Post-high school dummy	0.014	0.62	-0.048	-1.25	0.068	1.57
House Price Characteristics						
Average growth	0.206	0.99	-0.611	-1.40	0.879	1.83
Volatility	-0.744	-8.18	-0.138	-1.00	-0.656	-3.99
Adjusted R-square	26.25%					
Number of observations	22,009					

Figure IA.1
Non-Parametric Relationship Between Wealth and Housing Share

The figure reports the average share of housing in the portfolio for different groups of landlords (grey line) and non-landlord homeowners (black line) sorted by their log gross worth. The results are based on a representative sample of households in the PSID between 1984 to 2013. For each year in the survey, we sort landlords and non-landlords into 10 deciles based on their gross worth and compute the (equally-weighted) average share of real estate out of gross worth. We then compute, for each decile, the time-series average of the housing share and log gross worth over the years when the survey is conducted. Landlords are defined as homeowners that received rental income on real estate properties during the year prior to the survey.