A Computation

A.1 Computing the value function

A.1.1 Setting up the state and control grids

Rather than include consumption as a control variable and the budget constraint as a test of feasibility, we save on computation cost by excluding consumption from the control space and computing consumption as the value that causes the budget constraint to bind. The two variables of interest are then h and b. The remaining variables are discrete, while these two variables have to be discretized. We take $\overline{b}$ to be the maximum net wealth that the household can hold. Given $\overrightarrow{p}$, $\overline{h}$ is set at that value such that $\overline{h}p_J = \overline{b}$. We set the number of points on the housing grid to 121 with the housing grid points evenly spaced.

The wealth grid is set conditional on the housing grid and the discrete choices such that the net wealth of the household does not exceed $\overline{b}$. The lower bound on the wealth grid is 0 for renters and given by the borrowing constraint for owners. For the mover's problem, the grid for cash-in-hand, b^m, has its lower bound set to $-\bar{l}$, since a household has to have positive cash-in-hand in order to move. The upper bound is set to $\overline{b}$.

The number of island productivities is finite, and so is the set of idiosyncratic household states since households are finitely lived. The fixed effect is drawn from the distribution over the fixed effects as given in the calibration section. Finally, the initial wealth of newborn households is drawn from the distribution of b_0. Section (A.1.2) explains why we do not have to pick the initial wealth from a discrete grid. Rather, we treat it as a continuous variable.

All the grid sizes and densities are tested to ensure that results do not change significantly when grids are made finer.

A.1.2 Using the Golden Section Search routine

The standard solution involves discretizing all continuous variables. We improve the efficiency and the accuracy of the value function computation by considering savings to be a continuous variable. The innovation involves using the Golden Section search to optimize over the savings space given all other states and choices. As the grid size of savings increases, using golden section search is more efficient than a brute force search over all grid points. More importantly, treating savings as a continuous choice variable means that we can get more accurate results for the optimization procedure.

For the rest of the section on computation, when we refer to *discrete states* $(\tilde{s}^d)$ or *discrete choices* $(\tilde{y}^d)$, we mean all state variables or choice variables excluding savings.

A.1.3 The value of moving and staying

As outlined in the standard method, we use backward iteration to compute the value function. At every state s, the value of staying is calculated using the golden section search routine. Similarly we also compute the value of moving, V^m, at every state, s^m and choice y^m.

Next we compute the value at every $(s, \tilde{y}^d)$ pair. The value of staying comes directly from the calculations above. When calculating the value of moving, the first issue we encounter is that the cash-in-hand of the household need not be on the b^m-grid. We use linear interpolation between states here, which is standard in the literature. The second issue concerns the optimal discrete choice, $\tilde{y}^d$. The standard method does the following: For the two closest points on the b^m grid, calculate the optimal choice and the value of moving. Then interpolate to get the value at s. We found that for a computationally feasible grid size this technique was too inaccurate. That is, increasing the grid size gave significantly different results for the discrete choice.

Instead, we use the following technique: The value of every discrete choice at b^m is calculated as the linearly interpolated value of discrete choices at the two closest points on the b^m-grid. The optimal discrete choice is then calculated as the argmax of this matrix. We then compute the optimal saving using the actual cash-in-hand, b^m, and the optimal discrete choice. Testing with finer grids shows that this technique gives more accurate and robust results than the standard method.

We note that a useful feature of this technique is that households can be followed over the life cycle. Using the standard technique, when the household has wealth outside the grid, using interpolation to calculate the expected value means that there might be no corresponding optimal choice. This is particularly a concern when the two closest values of b^m imply different optimal discrete choices. Our technique provides an more accurate alternative where we can follow households over the life cycle.

If a state point is infeasible, that is, there are no feasible choices at that state, a value of $-\infty$ is assigned, and the optimal control defaults to the same wealth level as the state. The choice of optimal control does not affect the computation since the expected value at any prior date that includes this state will be $-\infty$ as well.

A.2 Simulating the economy

Given the (expected) value function, we simulate the economy by forward iteration. We follow $200,000$ households over their life cycle. The initial state (the wealth endowment, b_0 and the fixed effect) as well as the shocks through the households' lifetime are drawn from the distribution as specified in the calibration section. Since we are interested in the economy at a steady state, we observe the following: the distribution of households over states at a particular age a is approximated by the distribution of the $200,000$ simulated households at age a. This allows us to get a snapshot of the economy which is composed of different cohorts, at any point in time by just using this set of simulated households. Since the economy is in a steady state, this snapshot does not change over time. We compute the aggregate macroeconomic variables using this set of simulated households.

A.3 The equilibrium price vector

The final step is determining the prices in the general equilibrium. Once we have determined the steady state distribution of agents, we can calculate the demand and supply in the housing, labor and capital markets. We iterate to find the equilibrium prices using a simplex search algorithm. This has the advantage of being robust even when the size of the parameter vector is large. We use the relative Euclidean distance between the demand and supply as the criterion function with the identity matrix as the weighting matrix.

B Calibration

B.1 Sample Selection Criteria

For the PSID, we remove the SEO sample (id>3000) and drop students and those households with top-coded labor income of head or wife or total household income less than $500 in any year. For the CPS, we drop households with top-coded data on income from alimony and other sources or if no one in the family is working or if total income is less than $500. For the SCF, we drop households with financial assets outliers (less than $0 or greater than $10 million).

B.2 Household life-cycle and preferences

Survival probabilities, $\lambda()$, are taken from the National Center for Health Statistics, United States Decennial Life Tables for 1989-1991. This table gives the mortality rate of the population as measured in the 1990 Census. We use the life table for the whole population from 1989. We set the retirement age to 66.

B.3 Family size equivalence

Figure B.1 shows the profiles of family size from the CPS for 1970-1993. Family size increases sharply when the household is young, peaking at age 38. In the model, we compute the family size transition matricies $\pi_f(f'|f,a)$ using data from the PSID. We compute separate matricies for the following age-bins (allowing for more bins where the slope of the family size profile is higher): 21-25, 26-30, 31-35, 36-40, 41-50, 51-60, 61-70, 71+. For tractability and since some transitions have few observations[1], we place the following restrictions on the transitions: any age household may transition to the single ($f = 1$) state from any other state, the maximum size household is $f = 5$, households can at most gain or lose one family member per year unless they transition to $f = 1$, no household under age 41 may transition from $f = n$ to $f = m > 1$ if $n > m$, and no household over age 60 can increase in family size.

In order to adjust the household's housing and consumption stream we use a household equivalence scale. The objective of an equivalence scale is to measure the change in consumption needed to keep the welfare of the family constant as the family size varies. Note that using per capita consumption assumes that the family converts consumption expenditure into utility flow following constant returns to scale. Lazear and Michael [1980] point to the existence of family goods, economies of scale and complementarities, which are all factors that they show to be significant. We therefore use a household equivalence scale that is not constant returns to scale. Table 1 lists some equivalence scales. L-M stands for Lazear and Michael [1980], US Dept of Commerce refers to US Department of Commerce [1991] and F-V&K stands for Fernandez-Villaverde and Krueger [2007]. Lazear and Michael's scale takes greater account of common or public goods, so that the impact of family size is less then other equivalence scales (compare, for instance, Orshansky [1965]). We use the housing equivalence scale used by Fernandez-Villaverde and Krueger [2007].

[1]For instance 2 out of 866 families that had $f = 2$ and were between ages 26-30 reported having $f' > 4$.

Family Size	L-M	Orshansky [1965]	US Dept of Commerce	F-V&K
1	100	100	100	100
2	106	126	128	134
3	128	151	157	165
4	147	189	201	197
5	169	223	237	227

Table 1: Family size equivalence scale

Figure B.1 shows the average family size from the CPS and the equivalent, normalized family size over the life cycle.

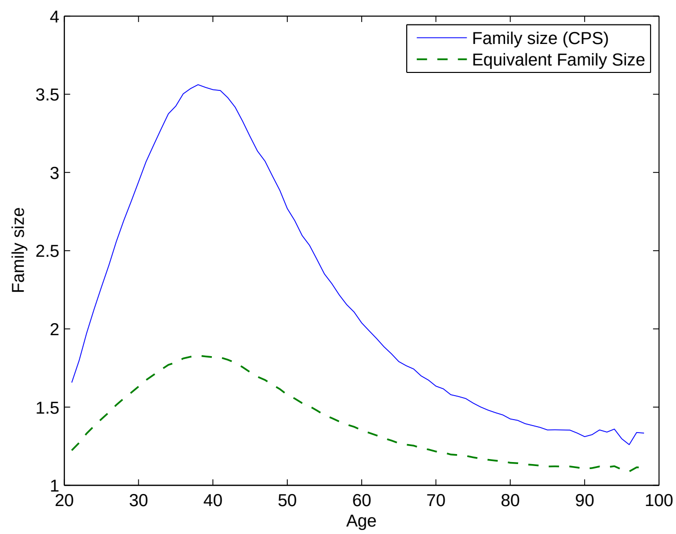

Figure B.1: F-V&K adjusted family size profile (yearly bins)

B.4 Initial wealth distribution

We calibrate the wealth distribution of newborns using the distribution of wealth among 21-25 year olds in the Survey of Consumer Finances (SCF) waves from 1989-2001. We drop households with negative wealth, those who report financial assets in excess of $10 million and students from the sample and use the sample weights and imputations provided by the SCF. We parametrize the initial wealth distribution as an exponential distribution.

That gives us one parameter that we have to match.

$$f(b_0) = \lambda_w e^{-\lambda_w b_0}$$

where b_0 is the initial wealth, and λ_w is the parameter to estimate in the exponential distribution. We estimate λ_w by matching the mean of the initial wealth distribution.

$$\lambda_w = \frac{1}{\bar{b}_0}$$

This gives us $\lambda_w = 0.00589$. We convert the initial wealth distribution in the data to model terms by scaling by the ratio of average labor earnings at age 21 in the model to average labor earnings at age 21 in the data.

B.5 Technology

Following Cooley and Prescott [1995], we calibrate δ using the law of motion for the capital stock: $K' = K(1-\delta) + I$, where I is investment. In the steady state (adjusting for growth), the capital stock remains constant and investment is used only to replace depreciated capital: $\delta = \frac{I}{K}$, We calculate K from the *Historical-cost Stock of Private Non-Residential Fixed Assets* in the NIPA. I is calculated from the *Historical-cost Investment in Private Non-Residential Fixed Assets*. We use data from the period $1970-1993$[2]. δ is computed as the growth adjusted average of $\frac{I}{K}$ over this period. This gives us $\delta = 0.13$. We set the capital share of output, α, at 0.34.[3]

We calibrate housing depreciation using the law of motion of housing capital. In the model we assume that housing supply is fixed and that home owners pay a maintenance cost to replace depreciated housing capital. So, the (growth-adjusted) relationship between housing depreciation in the model and housing investment in the national accounts is

$$\delta_h = \frac{I_h - \Delta(pH)}{pH}$$

For the value of housing, pH, we use non-farm owner-occupied housing from NIPA's

[2]This period is chosen in part to remain consistent with our decision to look at the economy before the housing boom-bust period circa 2001-2011 and also due to changes in the PSID after 1993 that would complicate extending the sample period by just a few years.

[3]Heathcote et al. [2009] set the value of α at 0.33 after surveying the literature. Cooley and Prescott [1995] set $\alpha = 0.4$. Greenwood et al. [1995] set $\alpha = 0.29$, which is followed by Gervais [2002].

Historical-Cost Net Stock of Residual Fixed Assets table. Investment in housing is computed using non-farm owner-occupied housing from NIPA's *Historical-cost Investment in Residential Fixed Assets.* This gives $\delta_h = 0.02$. These values from NIPA are the value of the structures and do not include the value of land.

B.6 Productivity process

For purposes of calibration, we assume that the household does not voluntarily move to a new U.S. state in the data.

B.6.1 Base wage over the life-cycle

The estimation of the productivity process is split into two steps. The first step involves estimating household earnings conditional on age, around which the household gets shocks. Figure (B.2) shows the base wage of households from the Social and Economic Supplet-

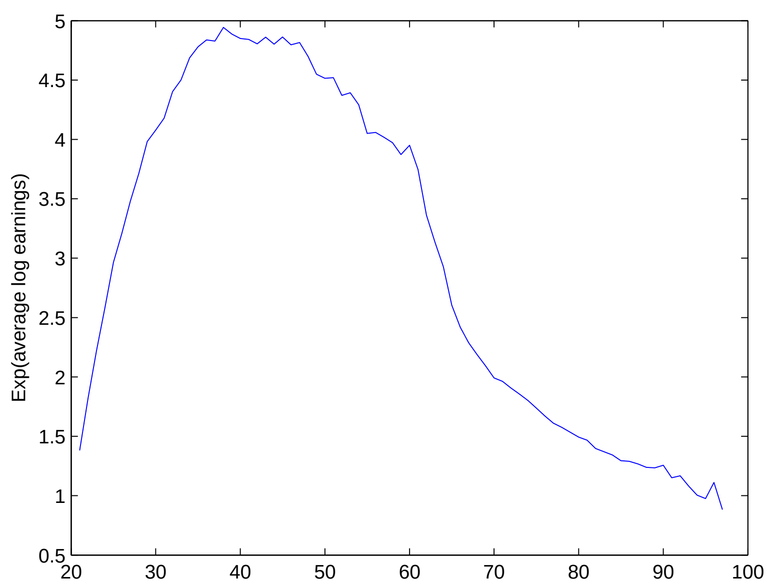

Figure B.2: Income over the life cycle (CPS - yearly bins)

ment to the Current Population Survey (CPS) for data from 1970-1993. *Base wage* is defined as the *exponential of the average log labor earnings* controlling for year effects. It is computed by taking log earnings from the CPS and regressing them on year and age dummies, with the household weights supplied by the CPS. The coefficients of the age dummies are the base wage for that age. The following regression on log income gets us the base wage, $\exp(d_k)$:

$$y_{ijat} = \sum_{k=1}^{A} 1_k d_k + \sum_{t'=1970}^{1993} 1_{t'} d_{t'} + w_{ijat}$$

where the subscripts are as follows: i indexes the individual, j indexes the state in which the individual lives, and a is the age of the head of the household. 1_k is the dummy variable that takes value 1 when $k = a$ and 0 otherwise, and d_k is the coefficient on the age-dummy variables. 1_t is the year dummy, and w_{ijat} is a mean-zero residual. The coefficients d_k are the (log) base wage conditional on age. In the model, we use a quartic approximation of the age dummies for ages 21 to 80 and then assume that the average log wage declines linearly at the rate of the difference between age 80 and 81 thereafter.

B.6.2 The earnings residual

The second step involves estimating the idiosyncratic and regional productivity shocks. The shocks are calibrated to match the variance profile of earnings over the life cycle. Since a component of the productivity shocks is persistent we use the PSID, which is a panel data set.

Our object of analysis is the residual of log earnings after controlling for age, year and family size. Since family size is also a function of age and year we first generate the family size residual by regressing family size on age and year dummies.

$$F_{ijat} = \sum_{k=1}^{A} 1_k d_k + \sum_{t'=1970}^{1993} 1_{t'} d_{t'} + \phi_{ijat}$$

where ϕ_{ijat} is the family size residual.

The log earnings residual, w_{ijat}, is defined by

$$y_{ijat} = \sum_{k=1}^{A} 1_k d_k + \sum_{t'=1970}^{1993} 1_{t'} d_{t'} + \beta^F \phi_{ijat} + w_{ijat}$$

and is the residual from the regression of log income on age dummies, year dummies and the family size residual. Henceforth we will refer to w_{ijat} as log earnings residual and log earnings interchangeably.

B.6.3 Parametric model for earnings

We model log earnings as

$$\begin{aligned} w_{ijat} &= \sigma_f f_i + \sigma_\nu \nu_{iat} + \iota_{iat} + \varepsilon_{jt} \\ \iota_{it} &= \rho_\iota \iota_{i,a-1,t-1} + \sigma_\iota e_{it} \\ \varepsilon_{jt} &= \rho_\varepsilon \varepsilon_{j,t-1} + \sigma_\varepsilon e_{jt} \end{aligned}$$

where i indexes the individual household, j indexes the state where the household resides, a indexes the age of the household, and t indexes time. $\sigma_f f_i$ is the fixed effect, where $f_i \sim \mathcal{N}(0,1)$ and σ_f is the standard deviation of the fixed effect shock. $\sigma_\nu \nu_{iat}$ is the temporary shock, where $\nu_{iat} \sim \mathcal{N}(0,1)$ and σ_ν is the standard deviation of the temporary shock. ι_{it} is the persistent idiosyncratic shock and ε_{jt} is the persistent regional shock. $e_{it}, e_{jt} \sim \mathcal{N}(0,1)$.

We assume that the initial values for the persistent shocks are set as follows: $\iota_{i0t} = 0 \; \forall t$ and ε_{j0} is drawn from its ergodic distribution. Storesletten et al. [2004] note that ρ_ι is very close to 1. We set $\rho_\iota = 1$ in our model, so that the persistent idiosyncratic shock follows a random walk.

B.6.4 GMM estimation

We estimate the parameters using an overidentified set of moments with the identity matrix as the weighting matrix. The set of moments is listed below:

$$\begin{aligned} m1(k) \equiv E(w_{ijat} w_{i'ja',t-k}) &= \rho_\varepsilon^k \sigma_\varepsilon^2 / (1 - \rho_\varepsilon^2) \\ m2 \equiv E(w_{ijat}^2) &= \sigma_f^2 + \sigma_\nu^2 + (a-1)\sigma_\iota^2 + \sigma_\varepsilon^2/(1-\rho_\varepsilon^2) \\ m3(n) \equiv E(w_{ijat} w_{ija-n,t-n}) &= \sigma_f^2 + (a-1-n)\sigma_\iota^2 + \rho_\varepsilon^n \sigma_\varepsilon^2/(1-\rho_\varepsilon^2) \end{aligned}$$

We use $k = \{0, 1, .., 15\}$ and $n = \{0, 1, .., 9\}$. We use households of ages $26-62$, where we have the most data, to estimate the above parameters.

Parameter	Description	Value
σ_f	Std. dev. of the fixed effect shock	0.5
σ_ι	Std. dev. of the persistent idiosyncratic shock	0.098
σ_ε	Std. dev. of the regional productivity shock	0.026
ρ_ι	Persistence of the idiosyncratic shock	1
ρ_ε	Persistence of the regional shock	0.9839

Table 2: Productivity parameters

B.6.5 Calibration of β, σ, θ_m

To calibrate β, σ, θ_m, we compute at each point of a coarse three-dimensional grid, the value of each model moment listed in section 3.7.1. As a criterion function, we use the sum of difference in logs of each data moment and the corresponding model moment, and use an identity weighting function for each moment. We then choose the point of the grid which minimizes the criterion function.

References

Thomas F. Cooley and Edward Prescott. Economic growth and business cycles. In Thomas F. Cooley, editor, *Frontiers of Business Cycle Research*, pages 1–38. Princeton University Press, 1995.

Jesus Fernandez-Villaverde and Dirk Krueger. Consumption over the life cycle: Facts from consumer expenditure survey data. *The Review of Economics and Statistics*, 89 (3):552–565, June 2007.

M. Gervais. Housing taxation and capital accumulation. *Journal of Monetary Economics*, 49:1461–1489, 2002.

J. Greenwood, R. Rogerson, and R. Wright. Household production in real business cycle theory. In Thomas F. Cooley, editor, *Frontiers of Business Cycle Research*. Princeton University Press, 1995.

Jonathan Heathcote, Kjetil Storesletten, and Giovanni L. Violante. Consumption and labor supply with partial insurance: An analytical framework. NBER Working Papers 15257, National Bureau of Economic Research, Inc, August 2009.

Edward P. Lazear and Robert T. Michael. Family size and the distribution of real per capita income. *American Economic Review*, 70:91–107, 1980.

Mollie Orshansky. Counting the poor: Another look at the poverty profile. *Social Security Bulletin*, January:3–29, 1965.

K. Storesletten, C.I. Telmer, and A. Yaron. Consumption and risk sharing over the life cycle. *Journal of Monetary Economics*, 51(3):609–633, 2004.

Bureau of the Census US Department of Commerce. Ser. p-60, no. 176-rd, table f and app. a. In *Current Population Reports*. US Government Printing Office, 1991.