Online Appendix for
"The Cross-Section of Household Preferences"

Laurent E. Calvet, John Y. Campbell,

Francisco Gomes, and Paolo Sodini*

October 2022

Section 1 describes data sources and the empirical inputs of the life-cycle model. Section 2 investigates the identification of preference parameters in simulations of the life-cycle model. Section 3 discusses indirect inference estimation and testing procedures. Section 4 develops the estimation method used to measure cross-sectional heterogeneity in the preference parameters. Section 5 reports additional empirical results. A supplementary appendix, available on our websites, provides additional information on the Swedish pension system.

*Calvet: EDHEC Business School and CEPR; e-mail: laurent.calvet@edhec.edu. Campbell: Harvard University and NBER; e-mail: john_campbell@harvard.edu. Gomes: London Business School and CEPR; e-mail: fgomes@london.edu. Sodini: Stockholm School of Economics and CEPR; e-mail: Paolo.Sodini@hhs.se. We thank Azar Aliyev,Nikolay Antonov, Huseyin Aytug, and Yapei Zhang for dedicated research assistance.

Contents

1 Data, Income, and Pensions

1.1 Data Sources and Household Variables

Statistics Sweden (SCB) has a parliamentary mandate to collect detailed data on demographics, income and wealth of all Swedish residents using tax returns and information from third parties, such as financial intermediaries, employers and welfare agencies. The demographic characteristics include age, gender, education, and marital status. SCB provides a household identifier since 1991 that allows us to define a household as a family living together with the same adults over time. We define the household head as the adult with the highest average non-financial disposable income; or, if the average income is the same, the oldest; or, if the other criteria fail, the man in the household.

The income registry is available since 1983 and contains information on employment sector and the income components necessary to construct the non-financial disposable income of each Swedish resident. We use income tax and its base to estimate the average income tax rate by education level. Information on retirement income and student allowances allows us to identify retirees and students in the population. The data do not distinguish between DC and DB pension payouts, but they provide pension-qualifying income and thus enable us to impute DC pension contributions and DB pension rights as explained in the main text, Section 1.2 of this appendix, and the supplementary appendix.

The wealth registry reports debt and disaggregated worldwide financial and real estate holdings at year-end from 1999 to 2007. Bank account balances, stock and mutual fund investments, and real estate holdings are observed at the level of each account, security, or property.[1] Even though the wealth registry does not provide the value of security holdings, it reports ISIN identification codes, which allows us to value and classify each financial asset by using FINBAS, a financial database maintained

[1]Bank account balances are reported if the account yields more than 100 Swedish kronor during the year (1999 to 2005 period), or if the year-end bank account balance exceeds 10,000 Swedish kronor (2006 and 2007). We impute unreported cash balances by following closely the method reported in Bach, Calvet and Sodini (2020).

by the Swedish House of Finance. For securities not covered by FINBAS, we use data from Citygate, Morningstar, Datastream, and the stock exchanges NGM and OMX. Positions in fixed income securities, capital guaranteed products and capital insurance are instead reported at market values. The data do not contain information on defined contribution retirement wealth, which we impute as explained in the main text, Section 1.2 of this appendix, and the supplementary appendix.

Real estate wealth includes residential real estate properties (i.e., primary and secondary residences), and commercial properties (i.e., rental, industrial, and agricultural properties) serving as business or investment vehicles. Real estate prices are compiled by Statistics Sweden from two main sources. Every 3 to 7 years, tax authorities assess the tax value of housing properties using detailed property characteristics and hedonic pricing. In addition, Statistics Sweden continuously collects data on every real estate transaction in the country, which permits the construction of sales-to-tax-value multipliers for different geographic locations and property types. The transaction data are also used to value apartments at the level of each residential building.

Debt is the sum of mortgages and all other liabilities to financial institutions.[2] Because we do not observe durable goods (such as appliances, cars and boats), the value of household debt can exceed the value of the assets we observe for some households. To avoid this problem, the debt variable is defined as the minimum of the total debt and real estate wealth reported in the registry. This approach is consistent with the fact that we proxy the borrowing rate by the average mortgage rate offered by Swedish institutions.

We explain in the main text the definition of the composite asset, which is a weighted portfolio of liquid financial wealth, real estate wealth, DC retirement wealth, and debt. In aggregate Swedish data in 1999, the shares of these four components in total net wealth are 36%, 76%, 13%, and -25%, respectively.

[2] We exclude student debt because it is exclusively provided by the state and heavily subsidized during our sample period.

1.2 Pensions

We explain the organization of the Swedish pension system and the imputation methodology we use in the main text. We refer the reader to the supplementary appendix for a full discussion of these topics.

1.2.1 The Swedish Pension System

The Swedish pension system consists of three pillars: state pensions, occupational pensions, and private pensions.

The *state pension* system requires each worker in Sweden to contribute 18.5% of their pension qualifying income: 16% to the pay-as-you-go defined benefit (DB) system and the remaining 2.5% to a defined contribution (DC) system called premiepension system. DC contributions are invested in a default fund, that mirrors the world index during our sample period, unless the worker opts out and chooses a portfolio of funds among those offered on the state DC platform. State DB payouts are a function of the pension qualifying income earned during the entire working life.

Occupational pensions were introduced to Sweden in 1991. They are regulated for the vast majority of Swedish residents by four collective agreements applying to blue-collar private-sector workers, white-collar private-sector workers, central government employees, and local government employees. Since these agreements specify workers' monthly pension contributions, the fraction directed to DB and DC pension plans, and the DC choices available to workers, we are able to impute both DC contributions and DB entitlements at the household level.

The collective agreements specify DC contributions as a percentage of pension qualifying income. These contributions are invested through insurance companies in either variable annuity products (called TradLiv in Sweden), or in portfolios of mutual funds, chosen by workers from a selection provided by the insurance company.

Defined contribution *private pensions* have existed in Sweden for a long time but our dataset provides us with individual private pension contributions from 1991. We assume that they are invested like occupational and state DC contributions.

1.2.2 Imputation Methodology

We now explain how pension qualifying income can be used to impute DB pension payouts and DC contribution. We also describe how we capitalize private pension savings and DC retirement wealth.

State DB payouts are a function of the pension qualifying income earned during the entire working life. Since our individual income data begin in 1983, we cannot observe the full income history for older individuals in our dataset. To handle this, we back-cast their income back to age 25 by using real per-capita GDP growth and inflation before 1983. We then use the state DB payout rules to impute state DB pension payments for each individual retiring during our sample period.

Occupational DB pension payouts can be accurately imputed because in all the collective agreements they are a function of at most the last 7 years of pension qualifying income during working life, and data on pension qualifying income are available from 1991 (the year occupational pensions were introduced).

Defined contribution *private pensions* have existed in Sweden for a long time but our dataset provides us with individual private pension contributions from 1991. More specifically, exact information on private pension contributions is available from 1994, whereas from 1991 the data reports only a capped version. We impute full contributions from 1991 to 1993 taking into account both age effects and individual savings propensities in subsequent years. We assume that these contributions are invested in the same way as occupational and state DC contributions. We follow Bach, Calvet and Sodini (2020) and allocate 58% of the aggregate stock of private pension wealth in 1991 to workers.[3] Across workers, we allocate pension wealth proportionately to their private pension contributions in 1991.

To calculate DC retirement wealth at each point in time, we accumulate contri-

[3]This allocation is chosen to satisfy the condition that imputed pension wealth should be roughly the same just before and just after retirement.

butions from all three pillars of the Swedish pension system. To do this for the state DC system, we follow the investment policy and cost of the system's default fund and assume that equity contributions are invested in the MSCI equity world index, without currency hedging, and are subject to a fee of 15 basis points. The equity share in the state DC system mirrors the allocation rules of the system's default fund: a 130% levered position in the world index up to the age of 55, which is then gradually rebalanced with age to an increasingly conservative portfolio. For occupational and private DC pensions, we assume that equity contributions are invested in the unhedged MSCI equity world index, subject to the 70 basis point fee that prevailed during our sample period. This assumption reflects the high degree of international diversification observed in Swedish equity investments (Calvet, Campbell, and Sodini 2007). The equity share in each household's occupational and private DC retirement portfolio is rebalanced with age following the representative age pattern of life-cycle funds available in Sweden during our sample period. We assume that all DC wealth not invested in equities is invested in cash.

1.3 Labor Income

Methodology. We estimate the income process from consecutive observations of household yearly income data over the period 1992 to 2007,[4] excluding the first and last year of labor income to avoid measuring annual income earned over less than 12 months.[5] We consider the total income received by all members of the household, but classify households by the head's education level and age. Since the vast majority of Swedish residents retire at 65, we consider two age groups: (i) non-retired households older than 19 and less than 65, and (ii) retired households that are at least 65.

For active households younger than 65, we estimate the coefficients a_c and b of the

[4]Since our individual income data begin in 1983, we cannot observe the full income history for older individuals in our dataset. To handle this, we back-cast their income back to the age of 25 by using real per-capita GDP growth and inflation before 1983. We then use the state DB payout rules to impute state DB pension payments for each individual retiring during our sample period.

[5]In each year, we winsorize non-financial real disposable income to a minimum level of 1000 kronor or about $150. We also winsorize the pooled data from above at the 0.01% level to take care of extreme outliers at the top of the income distribution.

labor income process, $\log(Y_{h,t}) = a_c + b'x_{h,t} + \nu_{h,t} + \varepsilon_{h,t}$, by running pooled regressions for each of the three education levels. The vector of explanatory variables $x_{h,t}$ includes age dummies, which we then regress on a third-degree polynomial in age and use the fitted third-degree polynomial in our life-cycle model. By construction, the residual, $y_{h,t} = \log(Y_{h,t}) - a_c - b'x_{h,t}$, satisfies

$$y_{h,t} = \nu_{h,t} + \varepsilon_{h,t} = \xi_t + z_{h,t} + \varepsilon_{h,t}.$$

We use the sample mean, $\bar{y}_t = \sum_h y_{h,t}/N$, as an estimate of the permanent aggregate component ξ_t. We estimate the variance of permanent aggregate shocks, σ_u^2, by the sample variance of $\bar{y}_t - \bar{y}_{t-1}$ $(t = 2, \ldots, T)$. Let $y^*_{h,t} = y_{h,t} - \bar{y}_t$ denote the idiosyncratic component of income. We estimate the variance of the permanent and transitory idiosyncratic labor income shocks, σ_w^2 and σ_e^2, as in Carroll and Samwick (1997).

For retired households, we impute the state and occupational after-tax pension benefit of each individual from 1999 to 2007, as explained in section 1.2 of this online appendix. We fill forward the imputed pension benefit in real terms until 2007 at individual level, and aggregate income at the household level in each year. The replacement ratio is estimated for each education group as the fraction of the average income of non-retired 64-year-old households to the average income of retired 65-year-old households across the 1999 to 2007 period.

Results. Table A.1 presents the size of education and income risk categories and Table A.2 reports the employment sectors of education and income risk categories. The patterns are intuitive, with relatively little income risk in the public sector and in mining and quarrying, electricity, gas, and water supply, and relatively high income risk in hotels and restaurants, real estate activities, construction for less educated workers, and the financial sector for more educated workers.

Figure A.1 plots the distribution of wealth-income ratios and risky shares across Swedish households. Figure A.2 illustrates the estimated age-income profiles for our three education groups. The profiles are steeper than profiles estimated in the US.[6]

[6] Dahlquist, Setty, and Vestman (2018) estimate income profiles for Sweden with a pronounced

Table A.3 shows that educated households, particularly those with higher education, face higher transitory income risk and lower idiosyncratic permanent income risk than less educated households. This result is a likely consequence of the following features of the Swedish labor market. First, it is straightforward for companies to downsize divisions, but extremely difficult for them to lay off single individuals unless they have a high managerial position. Second, companies that need to downsize typically restructure their organizations by bargaining with unions. Third, unions are nationwide organizations that span large areas of employment and pay generous unemployment benefits. Fourth, the pay cut due to unemployment is larger for better paid jobs. After an initial grace period, an unemployed person will be required to enter a retraining program or will be assigned a low-paying job by a state agency. All these features imply that unemployment is slightly more likely and entails a more severe proportional income loss for workers with higher levels of education.

The correlation between permanent income shocks and wealth shocks is a key ingredient of our model. In the main text, we report that the average value of this correlation across the nine education-sector categories. Table A.4 reports these correlations for each of the nine categories and for three types of wealth. Risky liquid financial wealth is only weakly correlated to permanent income shocks, except for educated households employed in high-risk sectors. Real estate wealth and non-cash net wealth exhibit substantially stronger correlations, which are even more pronounced for educated households in high-risk employment sectors. As we noted in the main text, the correlation between the numeraire risky asset and individual income growth is much smaller because most individual income risk is idiosyncratic.

hump shape and lower income towards the end of working life. They use a model that excludes cohort effects, thereby estimating the age-income profile in part by comparing the incomes of households of different ages at a point in time. This procedure is biased if different cohorts receive different lifetime income on average. We obtain similar estimates when we exclude cohort effects from our model of income.

2 Identification Strategy

2.1 Intuition

To explain the intuition underlying our identification strategy, we consider an Epstein-Zin investor who can trade a riskless asset and a risky asset every period. The Euler equation for the return on the optimal portfolio is given by

$$1 = \mathrm{E}_t \left[\widetilde{\delta}_{t+1} \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \left(\frac{V_{t+1}}{\mu(V_{t+1})} \right)^{\frac{1}{\psi}-\gamma} R^P_{t+1} \right] \tag{1}$$

where $\widetilde{\delta}_{t+1} = \delta p_{t,t+1}$, $R^P_{t+1} = R_f + \alpha R^e_{t+1}$, and $\mu(V_{t+1})$ denotes the certainty equivalent of V_{t+1}.[7] Under the usual assumption of conditional joint lognormality, we obtain

$$\begin{aligned} \mathrm{E}_t g_{t+1} &= \psi[\mathrm{E}_t r^P_{t+1} - \log(\widetilde{\delta}_{t+1})] + (1-\gamma\psi)\,\mathrm{E}_t \widetilde{v}_{t+1} + \frac{1}{2\psi}\sigma^2_{g,t} + \sigma_{gr,t} \\ &\quad + \frac{\psi}{2}\left[\left(\frac{1}{\psi} - \gamma \right)^2 \sigma^2_{\widetilde{v},t} + \sigma^2_{r,t} + \left(\frac{1}{\psi} - \gamma \right) \sigma_{\widetilde{v}r,t} \right] + \left(\frac{1}{\psi} - \gamma \right) \sigma_{g\widetilde{v},t}, \end{aligned} \tag{2}$$

where lower case letters denote logs of upper case letters, $g_{t+1} = \log(C_{t+1}/C_t)$, and $\widetilde{V}_{t+1} = V_{t+1}/\mu(V_{t+1})$.

Equation (2) highlights the identification problem. If the expected portfolio return, the time discount factor, and the conditional variances are constant over time, then the expected consumption growth rate $\mathrm{E}_t g_{t+1}$ is constant and for any value of ψ there is a corresponding time discount factor δ that delivers the same level of $\mathrm{E}_t g_{t+1}$. Without additional restrictions on δ or ψ these two parameters cannot be separately identified, as shown by Kocherlakota (1990) and Svensson (1989).

Equation (2) also suggests three possible solutions. First, one can exploit time-variation in variance terms, which arises in life-cycle models with undiversifiable risky labor income such as ours. However these changes tend to be more substantial early

[7] This Euler equation holds with equality even though our model has borrowing constraints, because with labor income risk and a Bernoulli utility function that satisfies $u'(0) = \infty$ the agent will always choose to hold some financial assets. Our model also has short-sales constraints on risky asset holdings, but these do not bind for the middle-aged households we are considering.

in life, when households have less wealth to smooth shocks (Gomes and Michaelides 2005). A second channel is time variation in the expected portfolio return. Even though our model has no exogenous variation in expected asset returns, we have endogenous variation driven by changes in the agent's portfolio as a function of age. The third channel is time variation in the effective time discount factor $\widetilde{\delta}_{t+1} = \delta p_{t,t+1}$, driven by the survival probabilities $p_{t,t+1}$ which are also a function of age.

All three sources of variation imply that the profile of the wealth-income ratio is affected in different ways by the TPR and the EIS, at different ages. Our identification strategy builds on this intuition, as we now explain.

2.2 Regressions on Simulated Data

We illustrate the promise of our identification strategy by running a series of regressions based on simulated data from the model. More specifically we regress the underlying preference parameters that were used to generate those simulations against a series of moments from the simulated data. The values for the preference parameters are the same grid points that we consider in our estimation: 1,848 combinations of 12 values of RRA ranging from 3 to 12, 11 values of the TPR ranging from -0.05 to 0.22, and 14 values of the EIS ranging from 0.1 to 2.5. The exact grid points are provided in Section 3.1 of this appendix. For each of these 1,848 preference parameter combinations we consider all 4,276 combinations of the initial wealth-income ratio and other group characteristics that we observe in the data.[8]

To build intuition we consider four moments in our regressions. The first moment is the initial wealth-income ratio which determines the initial conditions in our simulations ($(W/Y)_{i0}$). The second moment is the average risky share for household i over the 8 years in our sample:

$$\overline{\alpha}_i = \left(\frac{1}{8} \sum_{t=1}^{8} \alpha_{it} \right), \tag{3}$$

[8]Simulated moments are obtained by averaging 10,000 simulations. In this exercise, unlike our empirical analysis, we use the observed wealth-income ratio only in the first year, and take wealth-income ratios in subsequent years from the simulated data rather than from the observed data.

which should provide strong identification of the risk aversion parameter. The third moment that we consider is the 8-year cumulative growth of the wealth-income ratio,

$$grWY_i = \left[\left(\frac{W}{Y} \right)_{i8} / \left(\frac{W}{Y} \right)_{i0} \right]. \tag{4}$$

Finally, to capture age variation in the rate of wealth accumulation, the fourth moment we consider is

$$convexWY_i = \frac{\frac{1}{2}\left[\frac{W}{Y}_{i0} + \frac{W}{Y}_{i8} \right]}{\frac{W}{Y}_{i4}} - 1. \tag{5}$$

This measures the convexity of the wealth-income ratio as a function of age.

As an alternative to these four moments, we also consider the average risky share and the average wealth-income ratio in every year, thus giving us a total of sixteen moments ($\{\alpha_{it}\}_{t=1,8}$ and $\{(W/Y)_{it}\}_{t=1,8}$). These are the moments we actually use in our estimation.

We first consider the risk aversion parameter. The average risky portfolio share is an intuitive moment to explore here, so we first run the following regression:

$$\gamma_i = k_\gamma^0 + k_\gamma^1 \overline{\alpha}_i + e_i. \tag{6}$$

Panel A of Table A.5 reports the estimation results. Confirming that the average risky share is a very good moment for identifying the risk aversion parameter, the adjusted R^2 from this regression is 81.5%. This is an extremely high number since we are estimating a linear regression and imposing the same coefficients across groups. We know that the true relationship is non-linear and also depends on the initial wealth-income ratio. In our second specification we add the cumulative growth rate of the wealth-income ratio as an explanatory variable; the adjusted R^2 statistic then increases to 82.2%. Finally, in the last column we report results using all 16 moments from our estimation ($\{\alpha_{it}\}_{t=1,8}$ and $\{(W/Y)_{it}\}_{t=1,8}$). The adjusted R^2 is now 90.9%, an impressive value for simple linear regressions.

In panel B of Table A.5, we turn our attention to the identification of the TPR. In the first specification we regress the TPR on the cumulative growth rate of the wealth-income ratio and obtain an adjusted R^2 of 23.1%. When we add the risky share and the convexity of the wealth-income ratio, the adjusted R^2 increases to 51.4%. The adjusted R^2 reaches 61.9% when we consider all 16 moments.

Panel C of Table A.5 focuses on the EIS. We first consider a specification that includes only the cumulative growth rate and the convexity of the wealth-income ratio. The adjusted R^2 is then only 2.9%. When we add the average risky share, the explanatory power is still only 2.9%. When we consider all sixteen moments ($\{\alpha_{it}\}_{t=1,8}$ and $\{(W/Y)_{it}\}_{t=1,8}$), so that we fully capture the evolution of the wealth-income ratio over time, the adjusted R^2 drops to 2.4%.

In panel D of Table A.5, we relax the linearity assumption and estimate regressions within seven different ranges of values of the initial wealth-income ratio. The values for the adjusted R^2 statistics vary between 17.2% and 64.0%. Overall, our simulation results confirm that the EIS is also well identified in our framework.

3 Indirect Inference

3.1 Estimation Procedure

The indirect inference procedure explained in the main text is implemented as follows. For each group g, we simulate the lifecycle model on a grid of preference parameters, as Section 3.2 of this appendix explains.

The grid is defined by 12 values of the RRA ranging from 3 to 12, 11 values of the TPR ranging from -0.05 to 0.22, and 14 values of the EIS ranging from 0.1 to 2.5. The grid values of the RRA are 3, 4, 5, 5.5, 6, 6.5, 7, 7.5, 8, 9, 10 and 12. To construct the grid values of the TPR, we assume that the patience parameter $\delta \in \{0.80, 0.85, 0.90, 0.92, 0.94, 0.96, 0.97, 0.98, 0.99, 1.00, 1.05\}$, so that TPR$= -\ln(\delta)$ is contained in $\{-0.05, 0, 0.01, 0.02, 0.03, 0.04, 0.06, 0.08, 0.11, 0.16, 0.22\}$. The grid values of the EIS are 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1.2, 1.4, 1.6, 1.8, 2, and 2.5. Overall, the grid contains 1,848 ($= 12 \times 11 \times 14$) parameter vectors θ.

For each vector θ on the grid, we calculate the value of the objective function:

$$[\tilde{\mu}_S^g(\theta) - \hat{\mu}^g]' \ \Omega \ [\tilde{\mu}_S^g(\theta) - \hat{\mu}^g] \, ,$$

where Ω is the weighting matrix defined in the main text and S is the number of Monte Carlo simulations used to compute $\tilde{\mu}_S^g(\theta)$.[9]

We next evaluate the objective function on a finer grid defined as follows. The RRA grid has a grid step of 0.01 and contains 81 equally-spaced grid points ranging from 2 to 10. The EIS grid has grid step of 0.01 and contains 241 values of the EIS ranging from 0.1 to 2.5. We consider an evenly spaced grid of the patience parameter δ ranging from 0.8 to 1.05 with a grid step of 0.001, which generates a TPR grid containing 251 values of the TPR ranging from -0.05 to 0.22. The finer grid therefore contain 4,899,771 (= $81 \times 241 \times 251$) preference parameters. We evaluate the objective function on the finer grid by interpolating the values computed on the original grid. We use modified Akima cubic Hermite interpolation, which is known to reduce interpolation overshoots and oscillations compared to standard spline methods.

For each group, we obtain the indirect inference estimator by determining the parameter vector on the finer grid that minimizes the objective function. This value may occasionally be slightly negative due to interpolation error.

For inference purposes, such as calculations of the root mean squared error or the Jacobian matrix, we compute each of the 16 auxiliary statistics by a separate interpolation. These interpolations are based on a cubic spline using not-a-knot end conditions.

3.2 Simulation Details

For a given group g and a given preference parameter θ, we compute the estimator of the binding function, $\tilde{\mu}_S^g(\theta)$, by Monte Carlo simulations. The period-t simulations take as given the group's wealth-income ratio at the end of year $t-1$, the realized

[9] See Section 3.2 of this appendix for further details.

group-level income shock, and the numeraire asset return during the year, which are all set equal to their respective empirical values. The simulations then proceed in four steps.

i. We simulate $S = 10,000$ households/paths in the group over year t. For each simulated unit $i \in \{1..., S\}$, we simulate idiosyncratic (permanent and transitory) labor income shocks and thereby obtain labor income, $\tilde{Y}_{i,t}$, and permanent income, $\tilde{Y}_{i,t}^P$, in period t. We set wealth at the beginning of period t, $\tilde{W}_{i,t-1}$, equal to $\tilde{Y}_{i,t}^P$ times the group's average wealth-income ratio at the end of period $t-1$. Using the lifecycle model's policy functions $\alpha_t^*(\,\cdot\,)$ and $C_t^*(\,\cdot\,)$, we compute the risky share, $\tilde{\alpha}_{i,t-1} = \alpha_t^*(\tilde{Y}_{i,t}, \tilde{W}_{i,t-1}, \tilde{Y}_{i,t}^P; \theta)$, and consumption, $\tilde{C}_{i,t} = C_t^*(\tilde{Y}_{i,t}, \tilde{W}_{i,t-1}, \tilde{Y}_{i,t}^P; \theta)$, of each simulated unit during year t. Consistent with the model, the simulated unit sets these quantities at the end of year $t-1$ and keeps them constant during year t.[10]

ii. We compute the predicted wealth of each simulated unit at the end of year t :

$$\hat{W}_{i,t} = (R_f + \tilde{\alpha}_{i,t-1} R_{N,t}^e)(\tilde{W}_{i,t-1} + \tilde{Y}_{i,t} - \tilde{C}_{i,t}). \tag{7}$$

The prediction incorporates the empirical return on the numeraire in year t.

iii. We obtain the group's predicted wealth-income ratio at the end of year t:

$$\widetilde{\mu}_{1,t}^g(\theta) = \frac{\sum_{i=1}^S \hat{W}_{i,t}}{\sum_{i=1}^S \tilde{Y}_{i,t}}. \tag{8}$$

iv. We observe the information set available at the end of year t, we sample S households, and we compute the group's predicted risky share at the end of year t:

$$\widetilde{\mu}_{2,t}^g(\theta) = \frac{\sum_{i=1}^S \tilde{\alpha}_{i,t} \tilde{W}_{i,t}}{\sum_{i=1}^S \tilde{W}_{i,t}}. \tag{9}$$

We stack the resulting values into the column vector $\tilde{\mu}_S^g(\theta)$.

[10]This methodology exploits the homogeneity of $\alpha_t^*(\,\cdot\,)$ and $C_t^*(\,\cdot\,)$, with respect to $(\tilde{Y}_{i,t}, \tilde{W}_{i,t-1}, \tilde{Y}_{i,t}^P)$.

3.3 Asymptotic Properties of Our Estimator

If our model is correctly specified, the indirect inference estimator $\hat{\theta}^g$ converges to the true preference parameter vector as the number of households in each group increases, as we now explain.

The empirical auxiliary estimator $\hat{\mu}^g$ is asymptotically normal:

$$\sqrt{N^g}\,[\hat{\mu}^g - \mu^g(\theta)] \xrightarrow{d} \mathcal{N}(0, W_g) \tag{10}$$

as the group size N^g goes to infinity. This result follows from the delta method and the fact that the auxiliary statistics (defined in equations (9) and (10) of the main text) can be interpreted as ratios of sample moments. We estimate the asymptotic variance covariance matrix of $\hat{\mu}^g$ by the jackknife estimator

$$\frac{\hat{W}_g}{N^g} = \frac{N^g - 1}{N^g} \sum_{i=1}^{N^g} (\hat{\mu}^J_{[i]} - \overline{\mu^J})(\hat{\mu}^J_{[i]} - \overline{\mu^J})', \tag{11}$$

where $\hat{\mu}^J_{[i]}$ is the auxiliary estimator obtained by excluding the i^{th} household, and $\overline{\mu^J} = (N^g)^{-1} \sum \hat{\mu}^J_{[i]}$.

The indirect inference estimator is asymptotically normal:

$$\sqrt{N^g}\,\left(\hat{\theta}^g - \theta^g\right) \xrightarrow{d} \mathcal{N}(0, V^g), \tag{12}$$

as Gouriéroux, Monfort, and Renault (1993) show. Furthermore, the asymptotic variance-covariance matrix is given by

$$V^g = (1 + s_g^{-1})\,(D_g\,\Omega\,D_g')^{-1}\,D_g\,\Omega\,W_g\,\Omega\,D_g'(D_g\,\Omega\,D_g')^{-1}, \tag{13}$$

where the ratio $s_g = S/N_g$ accounts for simulation noise and $(D_g)' = \partial\mu^g(\theta^g)/\partial\theta'$ is the Jacobian matrix of the binding function $\mu^g(\,\cdot\,)$ evaluated at the true parameter θ^g. In practice, we estimate the asymptotic variance-covariance matrix of V_g by its sample equivalent $\hat{V}^g = (1+s_g^{-1})\,(\hat{D}_g\,\Omega\,\hat{D}_g')^{-1}\,\hat{D}_g\,\Omega\,\hat{W}_g\,\Omega\,\hat{D}_g'(\hat{D}_g\,\Omega\,\hat{D}_g')^{-1}$, where $\hat{D}_g$ is a finite-difference approximation of D_g.

When the size of each group g is large, we could achieve efficient estimation by setting the second-stage weighting matrix equal to the inverse of the jackknife estimator: $\Omega^{(2)} = \hat{W}_g^{-1}$, and then solving the optimization problem defined in equation (11). of the main text Efficient estimation, however, is not feasible in our sample because most groups are too small to obtain a reliable estimator of W_g^{-1}. The median group size is 63, while the symmetric matrix W_g contains 136 ($= 16 \times 17/2$) distinct elements. A related problem is that in many groups, the weighting matrix $\Omega^{(2)} = \hat{W}_g^{-1}$ assigns almost all the weight to the risky share, while the wealth-income ratio plays essentially no role in estimation. Efficient estimation is therefore unsatisfactory in our sample on statistical and economic grounds.[11] For these reasons, we henceforth focus on one-step estimation based on the diagonal weighting matrix Ω defined in the main text. Since this approach does not provide global specification tests based on the value of the objective function (11), we focus on measures of fit based on root mean squared error or economic significance.

3.4 Hypothesis Testing

We now explain the methodology used to test hypotheses about preference parameters in Table 6 of the main text.

3.4.1 Group-Specific Tests

We consider a null hypotheses of the form $\mathbf{H_0} : R(\theta^g) = 0$, where $R(\theta^g)$ is a column vector function of dimension r. Since $\sqrt{N^g}(\hat{\theta}^g - \theta^g) \xrightarrow{d} \mathcal{N}(0, V^g)$, the delta method implies that

$$\sqrt{N^g} R(\hat{\theta}^g) \xrightarrow{d} \mathcal{N}\left[0, \frac{\partial R}{\partial \theta'}(\theta^g)\ V^g\ \frac{\partial R'}{\partial \theta}(\theta^g)\right]$$

under the null. In Table 6 of the main text, we report the results of χ^2 tests based on

$$N^g\ R(\hat{\theta}^g)'\ \left[\frac{\partial R}{\partial \theta'}(\hat{\theta}^g)\ \hat{V}^g\ \frac{\partial R'}{\partial \theta}(\hat{\theta}^g)\right]^{-1} R(\hat{\theta}^g) \xrightarrow{d} \chi^2(r), \tag{14}$$

[11] These difficulties are consistent with the finite-sample inaccuracy of two-step generalized method of moments studied in Hwang and Sun (2018).

where $\hat{V}^g$ is defined in Section 3.3 of this appendix.

Expected utility. Following this methodology, we test in each group g the null hypothesis that households in the group exhibit expect utility:

$$\mathbf{H_0} : \psi^g = 1/\gamma^g.$$

The restriction function is $R(\theta^g) = \gamma^g \psi^g - 1$, where $\theta^g = (\delta^g, \gamma^g, \psi^g)'$. The Jacobian matrix of R is the 1×3 vector

$$\frac{\partial R}{\partial \theta'}(\theta^g) = (0, \psi^g, \gamma^g)$$

under the null hypothesis. The corresponding χ^2 tests are reported in Table 6 of the main text.

Time Preference Rate. Since $TPR^g = -\ln(\delta^g)$, the delta method implies that the variance of TPR is the variance of beta divided by the squared of beta: $Var(\widehat{TPR^g}) \approx Var(\hat{\delta}^g)/(\hat{\delta}^g)^2$. This result allows us to test hypotheses on the time preference rate reported in Table 6 of the main text.

3.4.2 Tests Involving All Groups

In Table 6 of the main text, we also test restrictions involving the preference parameters of all groups. For instance, we assess if the vector of preference parameters, θ^g, or each of its components, are homogeneous across groups. These tests are conducted as follows.

Let $N = N^1 + ... + N^G$ denote the total number of observations, and let

$$k^g = N^g/N$$

denote the fraction of observations in group g. We stack the group-level parameters into $\theta = [(\theta^1)', \dots, (\theta^G)']'$ and $\hat{\theta} = [(\hat{\theta}^1)', \dots, (\hat{\theta}^G)']'$. Since $\sqrt{N^g}(\hat{\theta}^g - \theta^g) \xrightarrow{d} \mathcal{N}(0, V^g)$, we infer that $\sqrt{N}(\hat{\theta}^g - \theta^g) \xrightarrow{d} \mathcal{N}(0, V^g/k^g)$ and therefore

$$\sqrt{N}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, V),$$

where $V = \operatorname{diag}(V^1/k^1, \ldots, V^G/k^G)$.[12] A finite-sample estimator of the asymptotic variance-covariance matrix V is given by $\hat{V} = \operatorname{diag}(\hat{V}^1/k^1, \ldots, \hat{V}^G/k^G)$, where $\hat{V}^1, \ldots, \hat{V}^G$ are defined in Section 3.3 of this appendix.

We consider the null hypothesis

$$\mathbf{H_0} : R(\theta) = 0, \tag{15}$$

where $R(\cdot)$ is a column vector function of dimension r. Under the null, the delta method implies that

$$\sqrt{N} R(\hat{\theta}) \xrightarrow{d} \mathcal{N}\left[0, \frac{\partial R}{\partial \theta'}(\theta)\ V\ \frac{\partial R'}{\partial \theta}(\theta)\right] \tag{16}$$

and

$$N\ R(\hat{\theta})' \left[\frac{\partial R}{\partial \theta'}(\hat{\theta})\ \hat{V}\ \frac{\partial R'}{\partial \theta}(\hat{\theta})\right]^{-1} R(\hat{\theta}) \xrightarrow{d} \chi^2(r). \tag{17}$$

The result holds for fixed proportions $k^1, \ldots, k^G$, and for a total number of observations N going to infinity.

Equality of Group-Level Parameters to the Cross-Sectional Mean. Following this methodology, we test the null hypothesis:

$$\mathbf{H_0} : \theta^g = \sum_{i=1}^{G} k^i\ \theta^i,$$

where θ^i and k^i denote the parameter vector and relative size of each group i. The restriction function is

$$R^g(\theta) = \theta^g - \sum_{i=1}^{G} k^i\ \theta^i.$$

[12]We denote by $\operatorname{diag}(A^1, \ldots, A^n)$ the block diagonal matrix with diagonal blocks $A^1, \ldots, A^n$.

Let I_3 denote the identity matrix of size 3, let e_g denote the row vector of dimension G with g^{th} element equal to unity and other elements equal to 0, and let $k = (k^1, \ldots, k^G)$. We note that $\theta^g = (e_g \otimes I_3)\theta$, and $\sum_{i=1}^{G} k^i \theta^i = (k \otimes I_3)\theta$. The restriction function can therefore be rewritten in matrix form as

$$R^g(\theta) = A^g \ \theta,$$

where $A^g = (e_g - k) \otimes I_3$ for every g. The corresponding tests, based on equations (16) and (17) of this appendix, are reported in Table 6 of the main text.

4 Measuring Parameter Heterogeneity

We denote by $\mathbb{E}$ the expectation operator computed across random realization of income shocks and by E^* the cross-sectional expectation operator. The mean preference parameter vector in the population is $\mu_\theta = E^*(\theta) = \sum_{g=1}^{G} k^g \ \theta^g$, where k^g is the population share of group g and θ^g is the preference vector of households in the group. The cross-sectional variance-covariance matrix of the preference vector is $V_\theta = E^*[(\theta - \mu_\theta)(\theta - \mu_\theta)']$. Our objective is to estimate V_θ from the group-level indirect inference estimators $\hat{\theta}^g$ $(g = 1, \ldots, G)$ defined in the main text, controlling for estimation error.

The estimation of V_θ requires us to take account of estimation noise. We let

$$\begin{aligned}(\hat{\theta}^g - \mu_\theta)(\hat{\theta}^g - \mu_\theta)' \quad &= \quad (\hat{\theta}^g - \theta^g)(\hat{\theta}^g - \theta^g)' + (\theta^g - \mu_\theta)(\theta^g - \mu_\theta)' \\ &\quad + (\hat{\theta}^g - \theta^g)(\theta^g - \mu_\theta)' + (\theta^g - \mu_\theta)(\hat{\theta}^g - \theta^g)'.\end{aligned}$$

We take the expectations across realization of income shocks:

$$\begin{aligned}\mathbb{E}\left[(\hat{\theta}^g - \mu_\theta)(\hat{\theta}^g - \mu_\theta)'\right] \quad &= \quad \mathbb{E}\left[(\hat{\theta}^g - \theta^g)(\hat{\theta}^g - \theta^g)'\right] + (\theta^g - \mu_\theta)(\theta^g - \mu_\theta)' \\ &\quad + \mathbb{E}(\hat{\theta}^g - \theta^g) \ \ (\theta^g - \mu_\theta)' + (\theta^g - \mu_\theta) \ \ \mathbb{E}(\hat{\theta}^g - \theta^g)'.\end{aligned}$$

We apply the cross-sectional expectation operator and obtain:

$$\Psi^{(1)} = \Psi^{(2)} + V_\theta + \Psi^{(3)} + [\Psi^{(3)}]', \tag{18}$$

where $\Psi^{(1)} = E^*\mathbb{E}\left[(\hat{\theta}^g - \mu_\theta)(\hat{\theta}^g - \mu_\theta)'\right]$, $\Psi^{(2)} = E^*\mathbb{E}\left[(\hat{\theta}^g - \theta^g)(\hat{\theta}^g - \theta^g)'\right]$, and $\Psi^{(3)} = E^*\left[\mathbb{E}(\hat{\theta}^g - \theta^g) \ \ (\theta^g - \mu_\theta)'\right]$. When the estimators are unbiased, equation (18) is equivalent to

$$\Psi^{(1)} = E^*\left[Var(\hat{\theta}^g)\right] + V_\theta,$$

as the law of total variance implies.

In finite samples, we estimate V_θ by as follows. First, we estimate μ_θ by the size-weighted mean of group estimates:

$$\bar{\theta} = \sum_g k^g \ \hat{\theta}^g,$$

where $k^g = N^g/(\sum_k N^k)$ is the share of group g in the population. The estimator $\bar{\theta}$ is a consistent estimator of μ_θ as the group sizes $N^1, \ldots, N^G$ go to infinity.

Second, we estimate $\Psi^{(1)}$ by the size-weighted variance of group estimates:

$$\hat{\Psi}^{(1)} = \sum_{g=1}^{G} k^g \ (\hat{\theta}^g - \bar{\theta}) \ (\hat{\theta}^g - \bar{\theta})'.$$

We estimate $E^*\left[Var(\hat{\theta}^g)\right]$ by the average variance-covariance matrix of $\hat{\theta}^g$:

$$\hat{\Psi}^{(2)} = \sum_{g=1}^{G} k^g \ \frac{\hat{V}^g}{N^g},$$

where $\hat{V}^g$ is the asymptotic variance-covariance matrix defined in section 3.3 of this online appendix. We therefore estimate the variance-covariance of θ by

$$\hat{V}_\theta = \hat{\Psi}^{(1)} - \hat{\Psi}^{(2)}.$$

5 Additional Empirical Results

5.1 Distribution of Preference Parameters

Figure A.3 plots the univariate distribution of relative risk aversion, the time preference rate and the elasticity of intertemporal substitution across the Swedish households.

Figure A.4 plots heats maps for estimates of RRA and its standard error (top panel), the TPR and its standard error (middle panel), and the EIS and its standard error (bottom panel) across Swedish households. The figure reveals that the asymptotic standard error of the EIS is positively correlated with the level of the estimated EIS.

5.2 Relation Between Preference Parameters and Characteristics

The lower portion of Table 4 in the main text explores correlation patterns among preference parameters and observables. Tables A.6 of this internet appendix report multiple regressions rather than univariate correlations. Most patterns are similar, but controlling for the initial wealth-income ratio, the growth of wealth-income predicts the EIS positively rather than negatively.

The results in the main text are weighted by the number of households in each group. While this is the natural weighting scheme in household finance applications, asset pricing economists may be interested in wealth-weighted average preference parameters of households. In appendix Table A.7, we weight groups by their average wealth during the sample period rather than by their size. Compared to equally-weighted averages, we find a similar mean risk aversion of 7.14, much lower mean time preference rate of 2.63%, and a somewhat higher mean EIS of 1.19. The cross-sectional standard deviations of these parameters are similar to the equally weighted case.

5.3 Fitted and Observed Life Cycle Profiles

Figure A.5 illustrates the life-cycle profiles observed in the data and in the model as a function of age. Panel A reports the average risky share and Panel B reports the wealth-income ratio. The plots are computed by averaging across all 4276 groups of Swedish households, where each group is weighted by its wealth share. The figure shows that on average, the model fits well the patterns of portfolio age and wealth accumulation over the life-cycle.

5.4 Preference Parameter Heterogeneity

Our asymptotic standard errors can be used to adjust our estimates of the heterogeneity in true preference parameters. Table 3 and Figure 1 of the main text describe the cross-sectional distribution of our parameter estimates, but this is increased by noise in the estimation procedure. Since our asymptotic standard errors estimate the noise for each group, in principle we can correct for the effect of noise on the estimated cross-sectional variance of parameters by subtracting the cross-sectional average squared standard error from the cross-sectional variance of our estimates.

A practical difficulty in doing this is that some groups have extremely high standard errors. Although these high standard errors are not pervasive enough to undermine our ability to reject homogeneous preferences for most households in the group-specific tests reported in Table 6, they do have a strong influence on the cross-sectional average of squared standard errors. In fact, if we do not limit the influence of outliers the average squared standard error is higher than the cross-sectional variance of estimates for TPR and EIS, implying a negative cross-sectional variance for true TPR and EIS.

We obtain more reasonable results if we winsorize the group-specific standard errors at the 90th percentile of the cross-sectional distribution. Table A.8 shows that this procedure implies a cross-sectional standard deviation of 1.05 for RRA, 5.82% for the TPR, and 0.55 for the EIS, as compared with the cross-sectional standard deviations of estimates reported in Table 3 of the main text which are 1.06, 6.96%, and 0.90 respectively.

Table A.9 shows how model fit deteriorates under homogeneous preferences. The mean RMSE-scaled objective function more than doubles to 16.0% if we fix RRA at its cross-sectional mean. Fixing TPR at its cross-sectional mean produces a mean RMSE-scaled objective function of 8.6%, and similarly restricting the EIS delivers a mean RMSE-scaled objective function of 7.7%. Fixing all parameters at their cross-sectional means is disastrous in the sense that it increases the mean RMSE-scaled objective function to 24.8%. A life-cycle model with homogeneous preferences, under our maintained assumption of homogeneous rational beliefs, delivers an extremely poor fit to the cross-section of household behavior.

5.5 Heterogeneous Beliefs

In Tables A.10-A.12, we consider a simple form of heterogeneity in beliefs by considering three alternative assumptions about the Sharpe ratio: the base value of 0.26, a high value of 0.40, and a low value of 0.15. Then, for each group we pick the Sharpe ratio and preference parameters that minimize the objective function. The base case Sharpe ratio is selected for groups representing 54% of households, while the low Sharpe ratio and the high Sharpe ratio are each selected for 23% of households.

In Table A.10, we report the resulting size-weighted preference parameter estimates. Allowing for heterogeneity in household beliefs has only a modest impact on the average preference parameters we estimate. Mean RRA is now 7.80, the mean TPR is 4.72%, and the mean EIS is 1.01. The cross-sectional standard deviations of the TPR and the EIS are similar to those we estimate in the homogeneous-beliefs case, but the cross-sectional standard deviation of risk aversion is over twice as large at 2.74. The explanation is that the model uses heterogeneous beliefs to better fit wealth accumulation, and offsets belief heterogeneity with RRA heterogeneity to avoid counterfactual heterogeneity in the risky share.

Table A.11 reports the cross-sectional correlations of preference parameters. and observabke characteristics. As in Table 4 of the main text, preference parameters exhibit only weak cross-sectional correlations. The EIS and the TPR have very similar correlations to observable characteristics as in Table 4, while the RRA coefficient is less correlated to wealth variables under heterogeneous beliefs.

Heterogeneous beliefs necessarily improve the fit of our model by adding free parameters, but the degree of improvement is modest. Table A.12 shows that the mean RMSE-scaled objective function declines only from 7.03 in the homogeneous-beliefs case to 6.52 in the heterogeneous-beliefs case. Importantly, the mean RMSE-scaled objective function is a disastrous 21.44 when we combine heterogeneous beliefs with homogeneous preferences.

5.6 Monte Carlo Simulations

We evaluate the finite-sample performance of our procedure by a simple Monte Carlo exercise. For each group in our sample, we simulate our model under the group's initial conditions and the preference parameters we estimated for the group. We combine simulated households into hypothetical groups each containing N_g^* households, where N_g^* is a measure of the effective empirical group size. We repeat this procedure to obtain 1,000 hypothetical groups and calculate the mean parameter estimate. A comparison of this mean with the preference parameters under which the model was simulated allows us to assess finite-sample bias in our estimation method.

This Monte Carlo analysis does not fully capture the heterogeneity in household-level data, even under the assumption that our model holds without error at the household level and that all households in each group have identical preferences. This is because we simulate each household in the group assuming that the household has the group average wealth-income ratio at the start of the period. In the data, by contrast, and in the ergodic distribution of wealth-income ratios implied by the model, different households have different income and wealth levels at each point of time reflecting the influence of past idiosyncratic income shocks. Hence, the group average wealth-income ratio is more strongly influenced by those households with higher wealth. To partially capture this effect, we adjust our simulations to set the effective group size N_g^* equal to the reciprocal of the sum of squared wealth shares of individual households in the group, rather than the number of households in the group N_g. We find that N_g^* is on average about 3/4 of N_g, with relatively little variation in this ratio across groups.

In Table A.13 we report regression coefficients of Monte Carlo mean parameter estimates on the parameter estimates that were used to generate the simulated data ("true" parameters for the purpose of this exercise). The results are very good for RRA, which has a slope coefficient of 1.005, insignificantly different from one, and an R^2 statistic of 94%. The regression for TPR has a slope coefficient of 0.917 and an R^2 statistic of 90%. Results are not quite as good for EIS, which has a slope coefficient of 0.651 and an R^2 statistic of 64%. This regression places most of its weight on the high EIS estimates, which are noisy; but results are similar for the log of the EIS. An important lesson of these results is that small-sample bias cannot explain the substantial cross-sectional heterogeneity in our preference parameter estimates. There is almost no small-sample bias for RRA, and minimal bias for the TPR; and while there is some bias in our EIS estimates, a bias correction would have little effect on the cross-sectional dispersion of the EIS.

References

Dahlquist, Magnus, Ofer Setty, and Roine Vestman, 2018, On the asset allocation of a default pension fund, *Journal of Finance* 73, 1893-1936.

Hwang, Jungbin, and Yixiao Sun, 2018, Should we go one step further? An accurate comparison of one-step and two-step procedures in a generalized method of moments framework, *Journal of Econometrics* 207, 381–405.

Table A.1: Sizes of Education and Income Risk Categories

	No High School	High School	Post-High School	All
Low	17,929	44,224	47,988	110,141
Medium	16,036	53,920	49,001	118,957
High	11,362	29,403	28,677	69,442
All	45,327	127,547	125,666	298,540

This table reports the number of households in groups with 3 levels of education and working in sectors with 3 levels of income volatility given in Table A.2, and for aggregates of these categories.

Table A.2: Employment Sectors of Education and Income Risk Categories

	No High School	High School	Post-High School
Low	Mining and utilities	Mining and utilities	Mining and utilities
	Manufacturing	Public sector	Public sector
	Public sector	Manufacturing	Manufacturing
	Healthcare	Healthcare	Education
Medium	Education	Finance	Healthcare
	Finance	Education	Construction
	Transportation	Transportation	Other services
	Other services	Construction	Transportation
High	Construction	Wholesale	Finance
	Wholesale	Real estate	Wholesale
	Real estate	Other services	Real estate
	Hotels	Hotels	Hotels

This table reports the classification of employment sectors into 3 levels of income volatility, separately for households with 3 levels of education. Sectors are listed in ascending order of income volatility within each education level.

Table A.3: Percentage Volatilities of Income Shocks

	Total			Systematic		
	No High School	High School	Post-High School	No High School	High School	Post-High School
Low	13.86	13.69	15.65	2.68	2.80	3.27
Medium	18.06	16.54	17.06	2.95	3.16	3.33
High	21.36	20.75	22.29	2.89	3.18	3.68

	Idiosyncratic permanent			Idiosyncratic transitory		
	No High School	High School	Post-High School	No High School	High School	Post-High School
Low	7.22	6.60	5.15	11.52	11.67	14.41
Medium	8.08	7.45	4.16	15.88	14.43	16.21
High	6.03	6.45	3.91	20.29	19.46	21.63

This table reports the standard deviations of income shocks, in percentage points, for Swedish household groups with 3 levels of education and working in sectors with 3 levels of income volatility. The top left panel reports the total standard deviation of income shocks, the top right panel reports the standard deviation of systematic (group-level) permanent income shocks, the bottom left panel reports the standard deviation of idiosyncratic (household-level) permanent income shocks, and the bottom right panel reports the standard deviation of idiosyncratic transitory income shocks.

Table A.4: Percentage Correlations of Income Shocks with Wealth Shocks

	Risky Liquid Wealth			Real Estate Wealth		
	No High School	High School	Post-High School	No High School	High School	Post-High School
Low	4.29	6.05	10.46	29.87	32.90	42.61
Medium	3.07	6.21	9.09	30.18	34.10	50.94
High	2.58	5.80	23.33	32.57	35.07	47.93

	Aggregate Wealth		
	No High School	High School	Post-High School
Low	19.11	21.98	30.30
Medium	18.42	22.74	33.83
High	19.38	22.97	42.17

This table reports the correlations of income shocks with wealth shocks, in percentage points, for Swedish household groups with 3 levels of education and working in sectors with 3 levels of income volatility. The top left panel reports the correlation between the excess return on risky liquid wealth net of taxes and group average total permanent income shocks, the top right panel reports the correlation between the excess return on real estate wealth net of taxes and group average total permanent income shocks, the bottom panel reports the average yearly correlation between the excess return on non-cash aggregate net wealth and group average total permanent income shocks.

Table A.5: Regressions of Preference Parameters on Simulated Moments

Panel A. RRA Regressions.

Average RS	Yes	Yes	No
Growth of WY	No	Yes	No
Convexity of WY	No	No	No
16 moments	No	No	Yes
R^2	0.815	0.822	0.909

Panel B. TPR Regressions.

Average RS	No	Yes	No
Growth of WY	Yes	Yes	No
Convexity of WY	No	Yes	No
16 moments	No	No	Yes
R^2	0.231	0.514	0.619

Panel C. EIS Regressions.

Average RS	No	Yes	No
Growth of WY	Yes	Yes	No
Convexity of WY	Yes	Yes	No
16 moments	No	No	Yes
R^2	0.029	0.029	0.024

Panel D. EIS Regressions Part II.

WY range	≤ 1	$(1,2]$	$(2,3]$	$(3,5]$	$(5,7]$	$(7,10]$	> 10
R^2	0.604	0.640	0.634	0.534	0.463	0.343	0.172

This table reports the R^2 statistics of regressions in simulated data using all preference parameters on a grid containing 12 values of relative risk aversion (RRA) ranging from 3 to 12, 11 values of the time preference rate (TPR) ranging from -0.05 to 0.22, and 14 values of the elasticity of intertemporal substitution (EIS) ranging from 0.1 to 2.5. For each of the 1,848 combinations of preference parameters we consider all initial levels of the wealth-income ratio (WY) observed among Swedish household groups. We regress the preference parameters on simulated moments including the average risky share (RS), the initial WY, the 8-year cumulative growth of WY defined in equation (4), the convexity of WY defined in equation (5), and all 16 moments (8 values of RS and 8 values of WY) used in our empirical analysis. The dependent variable in the regressions is RRA in Panel A, the TPR in Panel B, and the EIS in Panel C. The three columns in Panels A, B, and C include different combinations of explanatory variables. Panel D runs EIS regressions separately, using all 16 moments, for simulated groups with initial WY in different bins indicated in the columns.

Table A.6: Size-Weighted Cross-Sectional Regressions of Estimated Preference Parameters on Group Financial Characteristics

	(1) RRA	(2) RRA	(3) TPR	(4) TPR	(5) EIS	(6) EIS
Average RS	-3.457***	-3.742***	0.153***	0.138***	0.525***	0.161
	(0.116)	(0.082)	(0.007)	(0.008)	(0.109)	(0.113)
Initial WY	-0.165***	-0.111***	-0.001***	-0.000	0.150***	0.153***
	(0.006)	(0.005)	(0.000)	(0.000)	(0.005)	(0.005)
Growth of WY	5.392***	9.472***	0.389***	0.509***	5.607***	6.540***
	(0.674)	(0.671)	(0.035)	(0.036)	(0.596)	(0.653)
Convexity of WY	0.172	-0.288	-0.051***	-0.052***	-0.570***	-0.393*
	(0.190)	(0.153)	(0.011)	(0.012)	(0.158)	(0.159)
Constant	4.687***	1.731*	-0.449***	-0.553***	-5.932***	-6.752***
	(0.716)	(0.709)	(0.037)	(0.039)	(0.635)	(0.699)
Control variables	No	Yes	No	Yes	No	Yes
R^2	0.436	0.832	0.388	0.438	0.208	0.234

This table reports the cross-sectional regression coefficients across estimated preference parameters and group financial characteristics. All regressions weight groups by their size, to recover the underlying cross-sectional relationships at the household level. Growth of WY is defined in equation (4) and convexity of WY is defined in equation (5). Standard errors are reported in parentheses and statistical significance levels are indicated with stars: * denotes 1-5%, ** 0.1-1%, *** less than 0.1% significance. There are 4,276 groups containing 298,540 households. Control variables are 9 income risk/education categories and cohort fixed effects.

Table A.7: Average Wealth-Weighted Cross-Sectional Distributions of Estimated Preference Parameters and Group Financial Characteristics

	Mean	Median	Std. Dev.	10%	25%	75%	90%
RRA	7.14	7.10	1.03	5.90	6.40	7.90	8.50
TPR (%)	2.63	2.12	5.12	-4.31	0.80	3.25	6.40
EIS	1.19	1.19	0.92	0.10	0.21	2.05	2.50
Average RS	0.59	0.56	0.14	0.42	0.49	0.67	0.77
Initial WY	6.76	5.17	4.71	1.73	3.09	9.09	14.70
Growth of WY	1.06	1.05	0.04	1.02	1.03	1.07	1.10
Convexity of WY	0.25	0.23	0.10	0.15	0.19	0.29	0.36

This table reports the mean, median, standard deviation, and 10th, 25th, 75th, and 90th percentiles of estimated preference parameters and group financial characteristics. All statistics weight groups by their average net wealth. Growth of WY is defined in equation (4) and convexity of WY is defined in equation (5). There are 4,276 groups containing 298,540 households.

Table A.8: Adjusting Estimated Parameter Heterogeneity for Estimation Noise

	(1) RRA	(2) TPR (%)	(3) EIS
Unadjusted	1.057	6.96	0.900
Adjusted by winsorized asymptotic standard errors	1.051	5.82	0.545
Adjusted by Monte Carlo standard errors	1.056	6.77	0.782

This table reports the cross-sectional standard deviations of parameter estimates with two alternative adjustments for estimation noise. Unadjusted standard deviations, as in Table 3 of the paper, are reported in the first row. The second row reports standard deviations subtracting the cross-sectional average squared asymptotic standard error, after winsorizing asymptotic standard errors at the 90th percentile of the cross-sectional distribution. The third row reports standard deviations subtracting the cross-sectional average squared Monte Carlo standard error. As discussed in the text, Monte Carlo standard errors are much smaller than asymptotic standard errors but this reflects the limited extent of within-group household-level heterogeneity considered in the Monte Carlo analysis.

Table A.9: Size-Weighted Cross-Sectional Distributions of RMSE-Scaled Objective Functions for Alternative Model Specifications

	Mean	Median	Std. Dev.	10%	25%	75%	90%
Unrestricted	7.03	5.95	4.48	3.40	4.40	8.48	12.23
Fixed RRA	16.03	12.60	11.61	5.39	8.26	20.06	31.25
Fixed TPR	8.55	7.06	5.11	4.07	5.20	10.35	15.39
Fixed EIS	7.71	6.52	4.50	3.81	4.80	9.37	13.02
Fixed TPR and EIS	9.80	8.30	5.47	4.81	6.10	11.81	17.38
All Parameters Fixed	24.78	18.61	21.23	7.53	11.27	30.92	47.93

This table reports the mean, median, standard deviation, and 10th, 25th, 75th, and 90th percentiles of the RMSE-scaled objective function for several alternative model specifications. All statistics weight groups by their size to recover the underlying cross-sectional distributions at the household level. The RMSE-scaled objective function is the square root of the objective function divided by 4 and multiplied by 100 to express it in RMSE-equivalent percentage units. The results in the first row are for the unrestricted model estimated in Table 7 of the main text. The results in subsequent rows are for models that fix selected parameters at their size-weighted cross-sectional means estimated in the unrestricted model. There are 4,276 groups containing 298,540 households.

Table A.10: Size-Weighted Cross-Sectional Distributions of Estimated Preference Parameters and Group Financial Characteristics, Assuming Heterogeneous Beliefs

	Mean	Median	Std. Dev.	10%	25%	75%	90%
RRA	7.80	7.50	2.74	4.40	5.90	9.00	12.00
TPR (%)	4.72	3.15	6.80	-4.40	1.01	6.19	18.39
EIS	1.01	0.58	0.90	0.10	0.20	1.83	2.50
Average RS	0.65	0.63	0.17	0.45	0.53	0.75	0.90
Initial WY	4.28	3.04	3.90	0.87	1.64	5.22	9.25
Growth of WY	1.08	1.07	0.05	1.03	1.05	1.10	1.14
Convexity of WY	0.24	0.23	0.09	0.15	0.19	0.28	0.35

This table is an equivalent of Table 4, allowing three possible household beliefs about the Sharpe ratio, that it equals 0.15, 0.26, or 0.40. For each group we pick the Sharpe ratio and preference parameters that minimize the objective function. This table reports the mean, median, standard deviation, and 10th, 25th, 75th, and 90th percentiles of estimated preference parameters and group financial characteristics. All statistics weight groups by their size to recover the underlying cross-sectional distributions at the household level. Growth of WY is defined in equation (4) and convexity of WY is defined in equation (5). There are 4,276 groups containing 298,540 households.

Table A.11: Size-Weighted Cross-Sectional Correlations of Estimated Preference Parameters and Group Financial Characteristics, Assuming Heterogeneous Beliefs

	RRA	TPR	EIS	Average RS	Initial WY	Growth of WY
RRA	1.000					
TPR	0.087***	1.000				
EIS	0.042***	-0.210***	1.000			
Average RS	-0.176***	0.550***	-0.115***	1.000		
Initial WY	-0.018***	-0.508***	0.365***	-0.501***	1.000	
Growth of WY	0.052***	0.606***	-0.161***	0.600***	-0.709***	1.000

This table is an equivalent of Table 4, allowing three possible household beliefs about the Sharpe ratio, that it equals 0.15, 0.26, or 0.40. For each group we pick the Sharpe ratio and preference parameters that minimize the objective function. This table reports the cross-sectional correlations across estimated preference parameters and group financial characteristics. Correlations weight groups by their size to recover the underlying cross-sectional correlations at the household level. Growth of WY is defined in equation (4) and convexity of WY is defined in equation (5). Statistical significance levels of correlation coefficients are indicated with stars: * denotes 1-5%, ** 0.1-1%, *** less than 0.1% significance. There are 4,276 groups containing 298,540 households.

Table A.12: Size-Weighted Cross-Sectional Distributions of RMSE-Scaled Objective Functions for Alternative Model Specifications

	Mean	Median	Std. Dev.	10%	25%	75%	90%
Homogeneous beliefs	7.03	5.95	4.48	3.40	4.40	8.48	12.23
Heterogeneous beliefs							
Unrestricted preferences	6.52	5.48	4.46	2.91	3.95	7.98	11.81
Fixed TPR and EIS	8.84	7.43	5.30	4.15	5.32	10.66	15.95
All preferences fixed	21.44	18.00	14.44	7.51	11.26	26.16	40.26
Pref vary only with beliefs	20.02	14.70	16.53	7.13	9.43	25.28	38.81

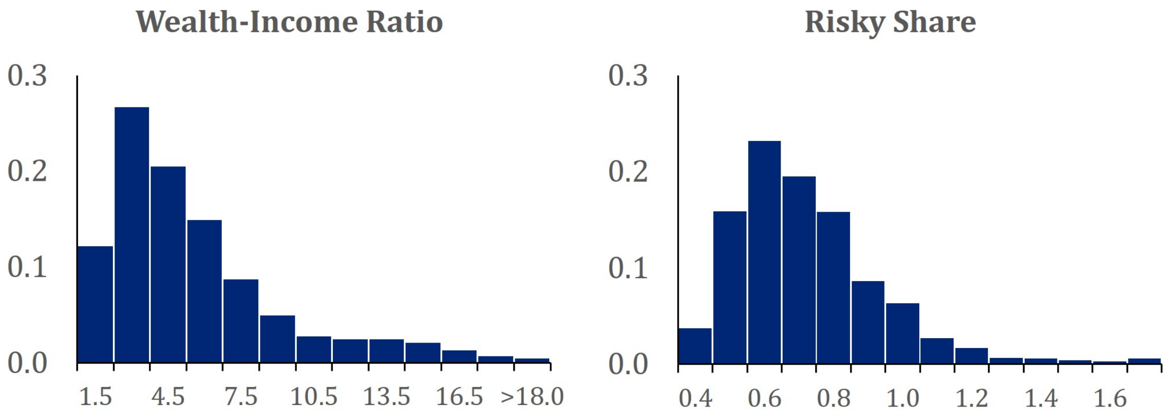
This table reports the mean, median, standard deviation, and 10th, 25th, 75th, and 90th percentiles of the RMSE-scaled objective function for several alternative model specifications. All statistics weight groups by their size to recover the underlying cross-sectional distributions at the household level. The RMSE-scaled objective function is the square root of the objective function divided by 4 and multiplied by 100 to express it in RMSE-equivalent percentage units. The results in the first row are for the unrestricted model estimated in Table 7. The results in subsequent rows are for models that fix selected parameters at their size-weighted cross-sectional means estimated in the unrestricted model. In the last row, the preference parameters are restricted to only vary with beliefs. There are 4,276 groups containing 298,540 households.

Table A.13: Size-Weighted Cross-Sectional Regressions of Mean Monte Carlo Preference Parameter Estimates on Indirect Inference Estimates

	(1) RRA	(2) TPR	(3) EIS	(4) Log EIS
Slope coefficient	1.005***	0.917***	0.643***	0.651***
	(0.003)	(0.007)	(0.010)	(0.011)
Constant	-0.030	0.006***	0.374***	0.031***
	(0.023)	(0.001)	(0.013)	(0.009)
R^2	0.936	0.897	0.630	0.636

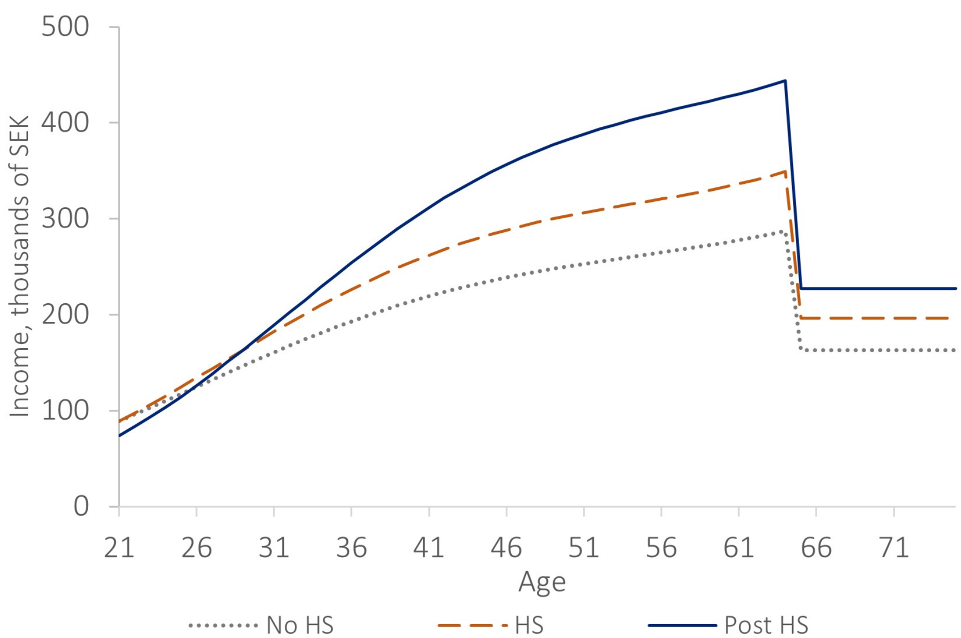
This table reports cross-sectional regressions of average Monte Carlo estimates of preference parameters on the preference parameters used to generate simulated data, which are set equal to indirect inference parameter estimates for each group. Monte Carlo simulations use the effective group size, the reciprocal of the sum of squared wealth shares of individual households in each group. Average estimates are calculated from 1,000 simulations of each group. All regressions weight groups by their size, to recover the underlying cross-sectional relationships at the household level. Standard errors are reported in parentheses and statistical significance levels are indicated with stars: * denotes 1-5%, ** 0.1-1%, *** less than 0.1% significance. There are 4,276 groups containing 298,540 households.

Figure A.1: Distribution of Wealth-Income Ratio and Risky Share Across Swedish Households



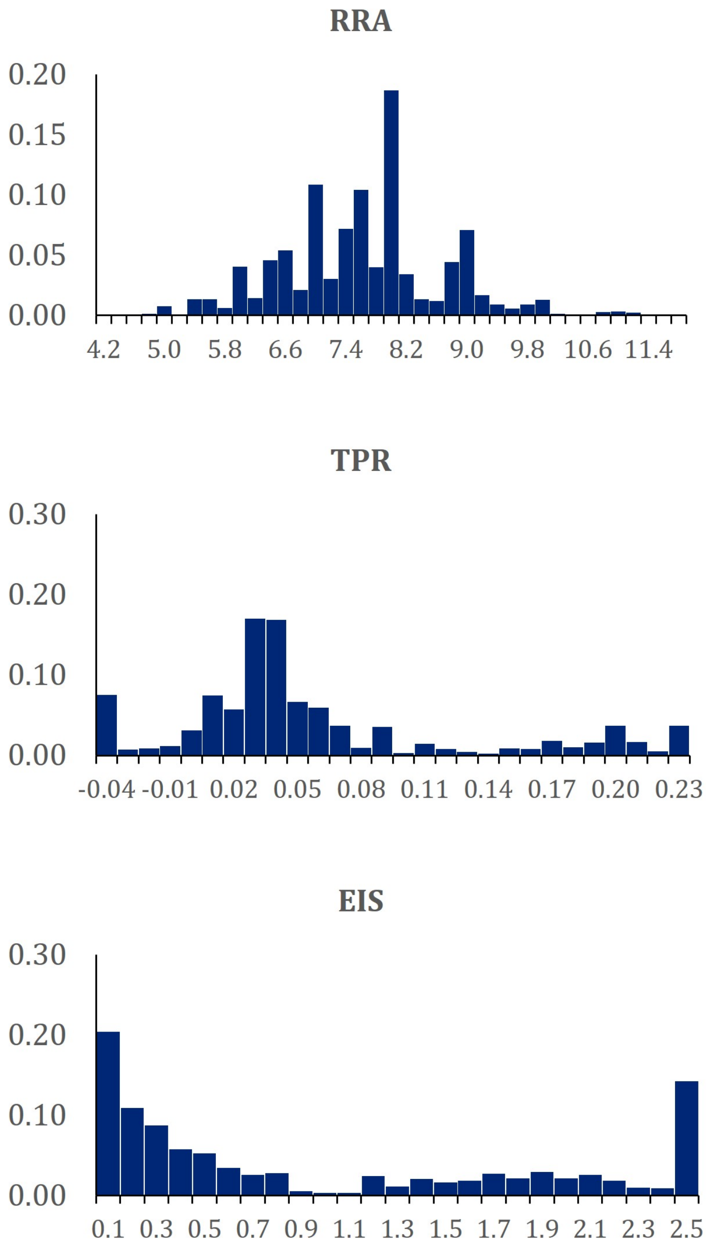
This figure presents histograms for the wealth-income ratio (WY) and risky share (RS) across 4,276 groups of Swedish households, size-weighted to recover the underlying distribution across households under the assumption that WY and RS are homogeneous within groups. Each bin is labeled on the horizontal axis with the upper cutoff value of WY or RS at the right edge of the bin, except the extreme right bin which captures all groups above the previous bin's cutoff. The vertical axis shows the size-weighted fraction of the sample in each bin.

Figure A.2: Estimated Age-Income Profiles



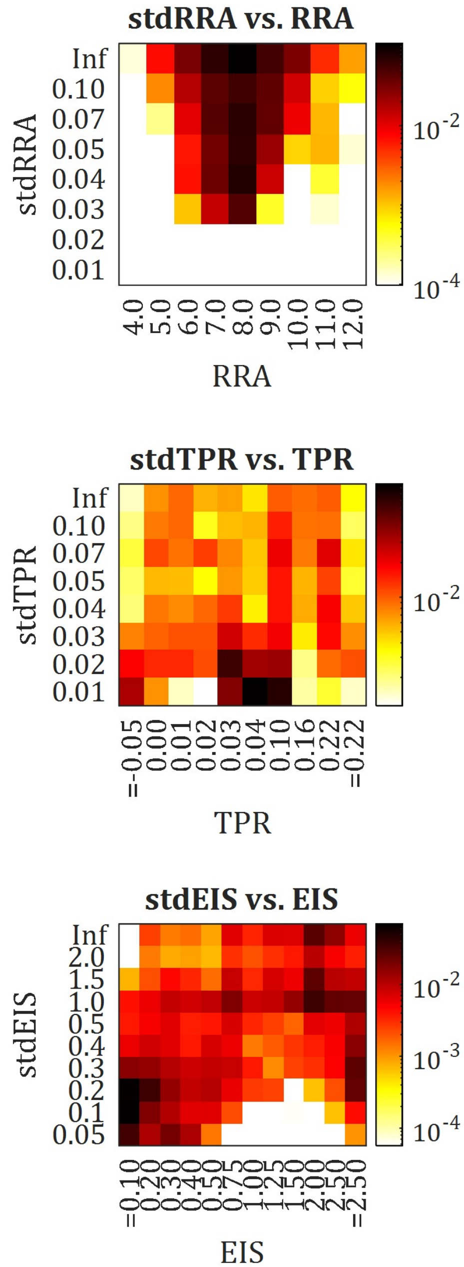
This figure presents estimated age-income profiles, including replacement ratios in retirement, for Swedish households with three levels of education: no high school (HS), high school, and post-high-school. The estimates are based on a labor income process specified in equations (5)-(6) in the main text.

Figure A.3: Distribution of Estimated Preference Parameters



This figure presents histograms for estimates of relative risk aversion (RRA), the time preference rate (TPR), and the elasticity of intertemporal substitution (EIS) across 4,276 groups of Swedish households, size-weighted to recover the underlying distribution across households under the assumption that preferences are homogeneous within groups. Each horizontal axis label shows the upper cutoff value at the right edge of the bin above the label. The vertical axis shows the size-weighted fraction of the sample in each bin.

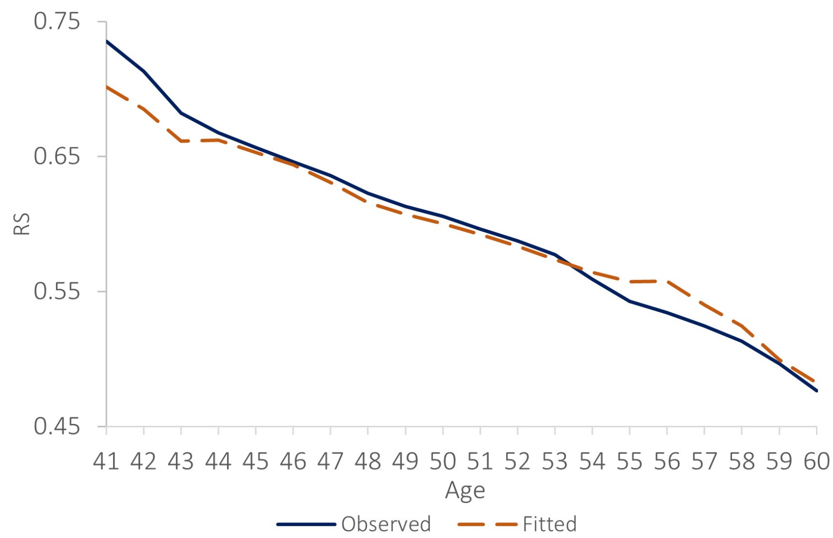
Figure A.4: Joint Distribution of Estimated Preference Parameters and Respective Standard Errors



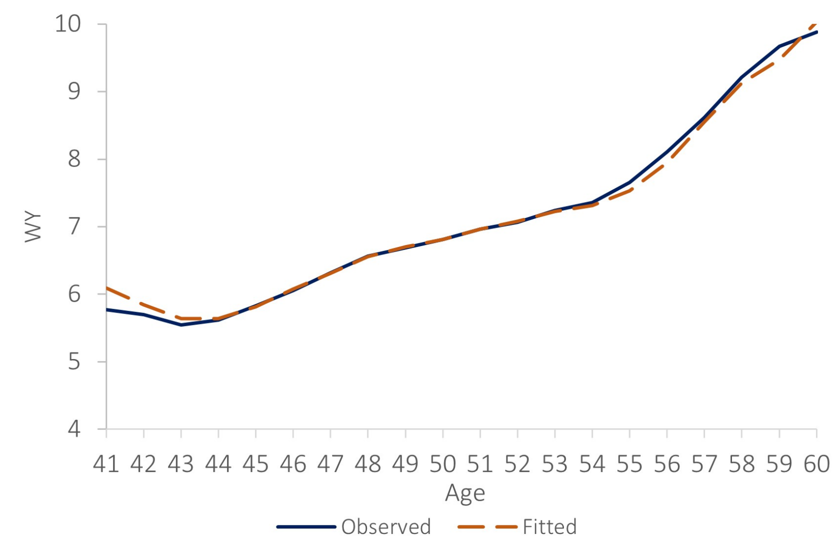
This figure presents bivariate heat maps for estimates of RRA and its standard error (top panel), the TPR and its standard error (middle panel), and the EIS and its standard error (bottom panel) across 4,276 groups of Swedish households, size-weighted to recover the underlying distribution across households under the assumption that preferences are homogeneous within groups. Each axis label shows the upper cutoff value of the corresponding bin, except for labels beginning with = which indicate that the bin contains only estimates of the exact value indicated by the label, and the label Inf which indicates that the bin has no upper cutoff but contains all values above the previous bin's cutoff. The logarithmic color scheme indicates the fraction of the sample in each bin. This fraction is equal to 11.0%, 9.7% and 7.0% for the darkest color in the top, middle and bottom panels respectively and 0.0% for the brightest color in all three panels.

Figure A.5: Empirical and Model-Implied Life-Cycle Profiles

A. Risky Share



B. Wealth-Income



This figure plots the average risky share (Panel A) and wealth-income ratio (Panel B), in the data and in the model, as a function of age. These moments are computed by averaging across all 4276 groups of Swedish households, where each group is weighted by its wealth share.