Inferring income properties from portfolio choices

Gualtiero Azzalini*

IIES

Date: May 9, 2023

Click here for most recent version

Abstract

Two main views exist on the nature of the labor income process: according to one, income shocks are very persistent and agents face similar life-cycle profiles - *Restricted Income Profiles* (RIP); according to the other, income shocks are not very persistent and life-cycle profiles are individual-specific - *Heterogeneous Income Profiles* (HIP). This paper studies the implications of these two views in a portfolio choice model in order to discover identification restrictions allowing to discern between them. I find that HIP and RIP imply different life-cycle patterns of the participation and conditional risky share choices but similar patterns of consumption and saving. Crucial for this result is the inclusion of cyclical skewness in the stochastic process for income, which enables us to correctly estimate the part of income risk deriving from the persistence of the shocks.

*Institute for International Economic Studies, Stockholm University, e-mail address: gualtiero.azzalini@iies.su.se. I am grateful to Per Krusell, Paolo Sodini, Kurt Mitman and Kieran Larkin for advice, discussions and support. I also wish to thank Timo Boppart, Tobias Broer, Richard Foltyn, José-Elías Gallegos, John Hassler, Stefan Hinkelmann, Philipp Hochmuth, Markus Kondziella, John Kramer, Claire Thurwächter, Zoltán Rácz, Roine Vestman, Joshua Weiss and participants at the IIES Macro Group for helpful feedback and comments.

1 Introduction

Income risk is a crucial determinant of agents' choices and it is therefore of paramount importance to understand its nature in order to correctly evaluate the effect of economic policies.

While the literature on this topic is vast, two main hypotheses have emerged: according to one, income shocks are very persistent and agents face similar life-cycle profiles - *Restricted Income Profiles* (RIP); according to the other, income shocks are not very persistent and life-cycle profiles are individual-specific - *Heterogeneous Income Profiles* (HIP). Income data alone do not easily allow us to empirically discern which of the two is right. As shown by Guvenen (2009), identification is achieved by looking at income covariances at far apart ages in the life cycle. These covariances are however hard to compute due to natural attrition. A further problem is the known issue that the researcher's information set is a subset of the agent's. Several papers have, therefore, taken another approach in order to tackle this question: because diverse types of income risk imply different economic choices, by looking at the latter the researcher can infer properties of the income process. For instance, Blundell et al. (2008), Guvenen (2007) and Guvenen and Smith (2014) consider consumption/saving choices.

This paper contributes to this literature by looking at another decision of the agent that is determined by income risk, namely portfolio choice. While the canonical portfolio choice model does not, in general, admit a closed form solution for the share of wealth detained in risky assets when labor income is not deterministic, several papers (e.g. Catherine, 2021; Chang et al., 2018a; Cocco et al., 2005) have analyzed the portfolio choice implications of income risk with numerically solved models. The main finding is that the properties of the income process - and especially higher order moments - are relevant for the portfolio allocation. Starting from this result, this paper investigates to what extent portfolio choice models can be used to understand the nature of the income process. Specifically, i) I estimate HIP and RIP versions of a rich stochastic process including cyclical skewness of income shocks, ii) I use them as inputs to solve a state of the art portfolio choice model, and iii) I find identifying restrictions from differences in model outcomes.

My findings are the following. First, cyclical skewness needs to be included in the income process in order to correctly estimate the amount of risk deriving from the persistence of the shocks. Indeed, in a model without cyclical skewness, HIP overestimates the share of risk attributed to heterogeneity in life-cycle profiles and underestimates the share deriving from persistence.

Second, because including cyclical skewness results in similar estimates for the persistence of the shocks for both HIP and RIP, I find that the profiles of the mean and variance of consumption over the life cycle are very much alike. In light of this, these two schedules, which have previously been considered by the literature (e.g. Guvenen, 2007), do not have a strong identification power to infer income properties.

Third, I find that HIP and RIP imply different average life-cycle profiles for participation and conditional risky share. This is due to the fact that agents' choices for these two quantities are different in the two models depending on their expected income growth rate during working life. Intuitively, the latter is more similar across agents in RIP and more heterogeneous in HIP. As a consequence, agents in HIP expecting higher lifetime income choose a lower risky share if human capital is risky, especially when they do not have enough wealth to self-insure. Specifically, due to the effect of cyclical skewness on the riskiness of human capital, the HIP process implies much less heterogeneity in participation rates across people with different income growth rates, and a "butterfly pattern" for the conditional risky share. The latter means that people with high income growth rates choose lower conditional risky shares in young ages compared to individuals with low growth rates, and they catch up at around 40 years old when the order of this pattern is reversed.

Comparing the model-generated profiles and their empirical equivalents using Swedish administrative data, I find that the latter provides slightly stronger support for the RIP than the HIP hypothesis, especially because no evidence of a "butterfly pattern" was found.

Related literature. First, this paper is related to the area of research that studies the nature and properties of the income process. Given the central role of the latter in economics and finance, this literature is unsurprisingly vast (e.g. Arellano et al., 2022; Busch et al., 2022; Guvenen, 2009; Guvenen et al., 2021, 2014; Lillard and Willis, 1978; MaCurdy, 1982; Meghir and Pistaferri, 2004; De Nardi et al., 2019; Storesletten et al., 2004). My contribution here is showing that once cyclical skewness of labor income shocks is taken into account in the estimation of the income process, the estimated persistence coefficient of the shocks is not much different between HIP and RIP.

Within this literature, a series of papers (e.g. Arellano et al., 2017; Blundell and Preston, 1998; Blundell et al., 2008; Deaton and Paxson, 1994; Hall and Mishkin, 1982; Kaplan and Violante, 2010) have used revealed choices to make inference on the properties of the income process. The main idea is that, because the latter have a relevant impact on agents' decisions, it should be possible to infer something about the income risk individuals face by looking at their choices. As mentioned, this approach also helps solve the well-known issue that the econometrician's information set is a subset of the agent's. Nevertheless, these papers have mainly looked at consumption choices. My contribution is, thus, to use this method on another decision that is influenced by income risk, namely portfolio choice. In using a structural model to make inference, my paper is most closely related to the approach followed by Guvenen (2007) and Guvenen and Smith (2014).[1]

Furthermore, this work relates to the literature on portfolio choice models and income risk. That income risk has an impact on portfolio choices has been already investigated empitically and theoretically (e.g. Benzoni et al., 2007; Catherine, 2021; Chang et al., 2018a,b; Cocco et al., 2005; Fagereng et al., 2017; Gomes and Michaelides, 2005, 2008; Huggett and Kaplan, 2016; Merton, 1969; Storesletten et al., 2007). A recent contribution by Catherine (2021) in this area has shown that modelling the procyclicality of the skewness of labor income shocks improves the fit with portfolio choices in the data. While keeping this channel in my model, this paper is the first to study the effects of cyclical skewness of labor income shocks and of different assumptions on the heterogeneity of life-cycle income profiles jointly.

Finally, as the framework in this paper is used to make inference on the properties of the income process, this paper is also connected to the emerging household finance literature in (Calvet et al., 2021) using portfolio choice models for structural estimation.

Structure of the paper. The paper is structured as follows. Section 2 presents the model, Section 3 deals with model estimation and calibration, Section 4 presents the results and Section 5 concludes.

2 Model

The model is a refined version of the canonical portfolio choice model by Cocco et al. (2005). Specifically, I build on recent work from Catherine (2021), who shows that including cyclical skewness of persistent income shocks is crucial to match well portfolio choices in the data. Compared to Catherine (2021), my framework additionally includes heterogeneity in life-cycle income profiles: in my model the constant and the coefficient of the age trend in the individual-specific part of income are state variables.[2]

Preferences. Let t denote age. Each agent enters into the model at age T_{start}, lives at maximum until age T and works until retirement age K. Each agent has Epstein-Zin preferences, similar to the forms specified in Inkmann et al. (2010) and Gomes and Michaelides (2005):

$$U_{i,t} = \left\{(1-\delta)c_{i,t}^{1-\psi} + \delta\mathbb{E}_t\left[p_t U_{i,t+1}^{1-\gamma} + b(1-p_t)a_{i,t+1}^{1-\gamma}\right]^{\frac{1-\psi}{1-\gamma}}\right\}^{\frac{1}{1-\psi}} \tag{1}$$

[1]Among these two papers, it is linked especially to the former as my objective is also to disentangle HIP and RIP.

[2]Catherine (2021) estimates the parameters governing the income process assuming that individuals have different life-cycle income profiles but does not include this heterogeneity when solving the model.

where $c_{i,t}$ is individual i's consumption at t, $a_{i,t}$ is individual i's wealth at t, δ is the discount factor, p_t is the probability of being alive at $t+1$ conditional on being alive at t, γ is the coefficient of RRA, $1/\psi$ the EIS and b determines the strength of the bequest motive. Because there is a bequest motive, the terminal condition for the recursion is:

$$U_{i,T+1} = ba_{i,T+1}^{1-\gamma} \tag{2}$$

Financial assets. Agents can invest in two financial assets, one risky with time-varying gross return R_{t+1} and one safe with constant gross return R_f. Letting small letters indicate log returns, r_{t+1} is given by the following expression:

$$r_{t+1} = r_{1,t+1} + r_{2,t+1} - \kappa_m \tag{3}$$

The effective return an individual gets by investing in the risky asset is the sum of two systematic components, one co-varying with labor market conditions (r_1) and the other that does not (r_2), and is net of a management cost κ_m, that is thus paid conditional on holding the risky asset. The systematic components are modelled as in Catherine (2021). Specifically, to take into account stock market crashes, r_1 is modelled as a mixture of Normals:

$$r_{1,t+1} = \begin{cases} \underline{r}_{1,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(\underline{\mu}_r, \sigma_{r_1}^2\right) & \text{w.p.} \quad p_r \\ \overline{r}_{1,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(\overline{\mu}_r, \sigma_{r_1}^2\right) & \text{w.p.} \quad 1 - p_r \end{cases} \tag{4}$$

Without loss of generality, it is possible to interpret p_r as the probability of stock market crashes and $\underline{\mu}_r$ the expected log return during crashing periods. Similarly, $1 - p_r$ is the probability of normal periods and $\overline{\mu}_r$ the average log return during normal periods. r_2 is, instead, a simple Normal shock:

$$r_{2,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_{r_2}^2\right)$$

Finally, investing in the risky asset is subject to a fixed participation cost κ_f that is paid every time the agent chooses to hold the risky asset. In addition, agents are allowed to borrow up to a borrowing limit on their total savings proportional to the exogenously set parameter $\overline{s}$. The repayment rate per unit of borrowing is set to be equal to the risk free rate.

Income process. I closely follow Catherine (2021) for specifying the labor income process. Let $L_{i,t}$ be individual i's real income. The logarithm of $L_{i,t}$ is the sum of an aggregate income component w_t and of an idiosyncratic component $y_{i,t}$:

$$\log\left(L_{i,t}\right) = w_t + y_{i,t} \tag{5}$$

The aggregate component follows a random walk with drift, driven by shocks to the market return through a parameter λ_{rw}:

$$w_t = g + w_{t-1} + \lambda_{rw} r_{1,t} + \phi_t \tag{6}$$

where $\phi_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_\phi^2\right)$.

The idiosyncratic component is the sum of a deterministic life-cycle component $\bar{f}(t)$, of a persistent component $z_{i,t}$ and of a transitory component $\nu_{i,t}$:

$$y_{i,t} = f_{i,t} + z_{i,t} + \nu_{i,t} \tag{7}$$

The persistent component is an AR(1) process:

$$z_{i,t} = \rho z_{i,t-1} + \varepsilon_{i,t} \tag{8}$$

with innovations drawn from a mixture of Normals:

$$\varepsilon_{i,t} = \begin{cases} \underline{\varepsilon}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(\underline{\mu}_{\varepsilon,t}, \underline{\sigma}_\varepsilon^2\right) & \text{w.p.} \quad p_\varepsilon \\ \overline{\varepsilon}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(\overline{\mu}_{\varepsilon,t}, \overline{\sigma}_\varepsilon^2\right) & \text{w.p.} \quad 1 - p_\varepsilon \end{cases} \tag{9}$$

Without loss of generality, it is possible to interpret p_ε as the probability of tail events and $\underline{\mu}_{\varepsilon,t}$, $\underline{\sigma}_{\varepsilon,t}$ the expected value and standard deviation of persistent income shocks during tail events, respectively. A similar interpretation holds for the parameters governing the distribution of normal events. To match the cyclicality of skewness, $\underline{\mu}_{\varepsilon,t}$ is defined as:

$$\underline{\mu}_{\varepsilon,t} = \mu_\varepsilon + \lambda_{\varepsilon w}(w_t - w_{t-1}) \tag{10}$$

Thus, tail events imply on average higher persistent shocks during expansions and vice versa during recessions. In addition, because persistent idiosyncratic shocks have zero mean, it must hold:

$$p_\varepsilon \underline{\mu}_{\varepsilon,t} + (1 - p_\varepsilon)\overline{\mu}_{\varepsilon,t} = 0 \tag{11}$$

The transitory shock is a pure innovation following a Normal distribution, whose variance depends on whether the persistent shock was drawn from the tail distribution or not:

$$\nu_{i,t} = \begin{cases} \underline{\nu}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \underline{\sigma}_\nu^2\right) & \text{if} \quad \varepsilon_{i,t} = \underline{\varepsilon}_{i,t} \\ \overline{\nu}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \overline{\sigma}_\nu^2\right) & \text{if} \quad \varepsilon_{i,t} = \overline{\varepsilon}_{i,t} \end{cases} \tag{12}$$

Finally, the deterministic component is specified differently for the HIP vs. RIP processes. Mathematically, the two specifications are summarized as follows:

$$f_{i,t}^{\text{model}} = \begin{cases} \bar{f}(t) + \alpha_i + \beta_i t & \text{if model = HIP} \quad (13) \\ \bar{f}(t) & \text{if model = RIP} \quad (14) \end{cases}$$

where $\bar{f}$ is a function of experience that is common to all individuals and to both models. In the HIP case, α_i and β_i are drawn from an i.i.d. bivariate Normal distribution with zero mean, variances σ_α^2, σ_β^2 and covariance $\sigma_{\alpha\beta}$. The RIP model, instead, only includes the function of experience and has no individual-specific components. Therefore, the RIP process is a restricted versions of the HIP specification with $\beta_i = \alpha_i = 0$ for all i.

Payroll taxes. As in Catherine (2021), working agents are subject to a 12.4% payroll tax on income up to a maximum taxable amount set by the Social Security Administration (SSA), which is roughly 2.5 times the average wage index.[3] Specifically, the total tax paid is:

$$T_{i,t} = .124 \cdot \min\{L_{i,t}, 2.5 \cdot e^{w_t}\} \tag{15}$$

Retirement income. It is common practice in standard portfolio choice models (e.g. Cocco et al., 2005) to assume that retirement income is a constant fraction of the income received in the last working period. However, in this model, large shocks in the last working period would have enormous effects on total retirement income, which is counterfactual. To solve this issue, I follow Catherine (2021) and assume that retirement income depends on the whole income history of an agent.[4] Let $\bar{L}_{i,t}$ keep track of average income as follows:

$$\bar{L}_{i,t} = \frac{1}{t - T_{\text{start}} + 1} \sum_{j=T_{\text{start}}}^{t} \min\{e^{y_{i,j}}, 2.5\} \tag{16}$$

Then, during retirement ($t > K$) agents have the following income stream:

$$\log(L_{i,t}) = \log(\zeta_{i,K}) + w_K \tag{17}$$

that is, retired agents enjoy a fraction ζ of the average wage index at the time of retirement. Following Catherine (2021), $\zeta_{i,K}$ depends on average income at retirement $\bar{L}_{i,K}$ as follows:

$$\zeta_{i,K} = \begin{cases} 0.9 \cdot \bar{L}_{i,K} & \text{if} \quad \bar{L}_{i,K} < 0.2 \\ 0.116 + 0.32 \cdot \bar{L}_{i,K} & \text{if} \quad 0.2 \leq \bar{L}_{i,K} < 1 \\ 0.286 + 0.15 \cdot \bar{L}_{i,K} & \text{if} \quad \bar{L}_{i,K} \geq 1 \end{cases} \tag{18}$$

Safety net. In order not to overstate the real level of income risk that agents face, it is important to model some traits of the welfare system. Therefore, following Catherine (2021), I model the Supplemental Nutrition Assistance Program and, for retired individuals, the Supplemental Security Income Program.

Only individuals with low wealth can participate in these programs. Specifically, only agents with less than about 2000 USD, which is roughly 5% of the average wage index. After retirement, eligible individuals receive supplemental income such that their total income reaches at least 20% of the average wage index. Before retirement, eligible individuals with earnings below 20% of the wage index receive benefits equal to 6% of the wage index minus 30% of their earnings. Mathematically:

$$N_{i,t} = \begin{cases} \max\{0.06 \cdot e^{w_t} - 0.3 \cdot L_{i,t}, 0\} & \text{if} \quad a_{i,t} < 0.05 \cdot e^{w_t}, L_{i,t} < 0.2 \cdot e^{w_t}, t \leq K \\ \max\{0.2 \cdot e^{w_t} - L_{i,t}, 0\} & \text{if} \quad a_{i,t} < 0.05 \cdot e^{w_t}, t > K \end{cases} \tag{19}$$

where $a_{i,t}$ is wealth as defined below.

[3] In 2010 the average wage index was 41673.83 USD and the maximum taxable amount 106800 USD.

[4] The downside is that this implies taking care of another state variable in the maximization problem.

The optimization problem. Let $\Xi_{i,t} = (\alpha_i, \beta_i, a_{i,t}, z_{i,t}, w_t, \bar{L}_{i,t})$ denote the state variable during working life and $\Xi^{\mathrm{R}}_{i,t} = (a_{i,t}, \bar{L}_{i,K}, w_K)$ the state variable during retirement.Also, let $R^e_{t+1} := \exp(r_{t+1}) - R_f$ denote the excess return and $L^d_{i,t} = L_{i,t} - T_{i,t} + N_{i,t}$ denote disposable income. Agent i chooses consumption $c_{i,t}$, savings $s_{i,t}$, a dummy variable $F_{i,t}$ equal to one in case she decides to hold risky assets and, conditional on participation, the share of savings invested in risky assets $\xi_{i,t}$ to maximize:

$$V_{i,t}(\Xi_{i,t}) = \max_{\xi_{i,t}, c_{i,t}, s_{i,t}, F_{i,t}} \left\{ (1-\delta) c_{i,t}^{1-\psi} + \delta \mathbb{E}_t \left[p_t V_{i,t+1}^{1-\gamma}(\Xi_{i,t+1}) + b(1-p_t) a_{i,t+1}^{1-\gamma} \right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} \tag{20}$$

subject to:

$$a_{i,t} = c_{i,t} + s_{i,t} + F_{i,t}\kappa_f \cdot \exp\{w_t\} \tag{21}$$

$$a_{i,t+1} = \left[R_f + \xi_{i,t} R^e_{t+1} \right] s_{i,t} + L^d_{i,t+1} \tag{22}$$

$$s_{i,t} \geq \bar{s} \cdot \exp\{w_t + \bar{f}(t)\} \tag{23}$$

and to the equations governing the exogenous processes outlined before for the different specifications governing the deterministic part of idiosyncratic income. During retirement, the agent's problem becomes:

$$V_{i,t}(\Xi^{\mathrm{R}}_{i,t}) = \max_{\xi_{i,t}, c_{i,t}, s_{i,t}, F_{i,t}} \left\{ (1-\delta) c_{i,t}^{1-\psi} + \delta \mathbb{E}_t \left[p_t V_{i,t+1}^{1-\gamma}(\Xi^{\mathrm{R}}_{i,t+1}) + b(1-p_t) a_{i,t+1}^{1-\gamma} \right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} \tag{24}$$

subject to the constraints (21), (22), (23), to the terminal condition (2) and to the equations governing the exogenous processes during retirement.[5] Appendix C describes in detail how the model is solved numerically.

3 Estimation and calibration

In this section, I describe the estimation of the parameters governing the exogenous stochastic processes for aggregate shocks, for individual income in the HIP and RIP cases and how I calibrate the remaining parameters in the model. The estimation procedure for aggregate shocks and individual income processes is described in Appendix D. More details on the data used can be found in Appendix B.

Aggregate processes. Estimation of the exogenous processes governing the market return and the aggregate part of income requires to estimate eight parameters: $\underline{\mu}_r, \overline{\mu}_r, \sigma_{r_1}, \sigma_{r_2}, p_r, \sigma_\phi, g, \lambda_{rw}$. I target mean, standard deviation, third and fourth standardized moments (skewness and kurtosis) of log yearly SP500 returns and aggregate wage log growth and the correlation between these

[5]In the last period of working life, that is when $t = K$, the problem is the same as under retirement, except that $V_{i,K}$ is defined on the state for the last period of working life $\Xi_{i,K}$.

two series.[6] The time sample for the returns is 1900-2019 and the one for aggregate wage growth is 1979-2011.[7] Table 1 reports the parameter estimates and the moments in the data and in the model. Overall the match is quite satisfactory.

Panel A: estimated parameters

	Stock market returns					Aggregate income		
	$\underline{\mu}_r$	$\overline{\mu}_r$	σ_{r_1}	σ_{r_2}	p_r	g	λ_{rw}	σ_ϕ
	−0.242	0.114	0.074	0.114	0.138	0.008	0.170	0.016

Panel B: moments

	Log returns				Aggregate income shocks				
	Mean	SD	Skew	Kurt	Mean	SD	Skew	Kurt	Corr
Data	0.064	0.183	−0.635	3.352	0.019	0.029	−0.767	3.773	0.658
Model	0.065	0.183	−0.636	3.485	0.019	0.029	−0.771	3.631	0.654

Table 1: Estimated parameters for the stochastic processes governing macroeconomic aggregates and moments in the data vs. model.

Individual income process. Estimation of the stochastic process governing individual income requires finding the values of eight parameters common to both specifications, namely $p_\varepsilon, \mu_\varepsilon, \lambda_{\varepsilon w}, \underline{\sigma}_\varepsilon, \overline{\sigma}_\varepsilon, \underline{\sigma}_\nu, \overline{\sigma}_\nu, \rho$, and two additional parameters for the HIP process, namely $\sigma_\beta, \sigma_{\alpha\beta}$. I target the time series between 1978 and 2010 of the standard deviation of log earnings growth at the one and five year horizons, Kelly's skewness of log earnings growth at the one, three and five year horizons and the within-cohort variance of log earnings for ages between 25 and 60.[8] The first set of moments is used to estimate the parameters of the persistent and transitory shocks and the second for the income profiles. In total I have 155 time-series moments and 36 within-cohort variances for a total of 191 moments. I perform SMM estimation, assuming that the economy is hit by the same aggregate wage shocks found in the data between 1944 and 2011.[9][10] with diagonal weighting matrix with all entries equal to one.

[6]Specifically, the correlation is between the log return in year $t-1$ and aggregate wage log growth in year t.

[7]Data for returns are obtained from Robert Shiller's website and for aggregate wage growth from Guvenen et al. (2014). See Appendix B for more details.

[8]I take the values for standard deviation and for Kelly's skewness from Guvenen et al. (2014) and the within-cohort variances from Guvenen et al. (2021). See Appendix B for more details.

[9]I use data from Emmanuel Saez before 1979 and from Guvenen et al. (2014) afterwards. See Appendix B for more details.

[10]I simulate the income histories of 68 cohorts, the first starting in 1944 and the last in 2011 assuming that the persistent component is zero at the beginning. Note that this implies having a constant age structure between 1979

Table 2 reports the parameter estimates and Figure 1 plots the targeted moments in the different model specifications and in the data. Overall the match is satisfactory.

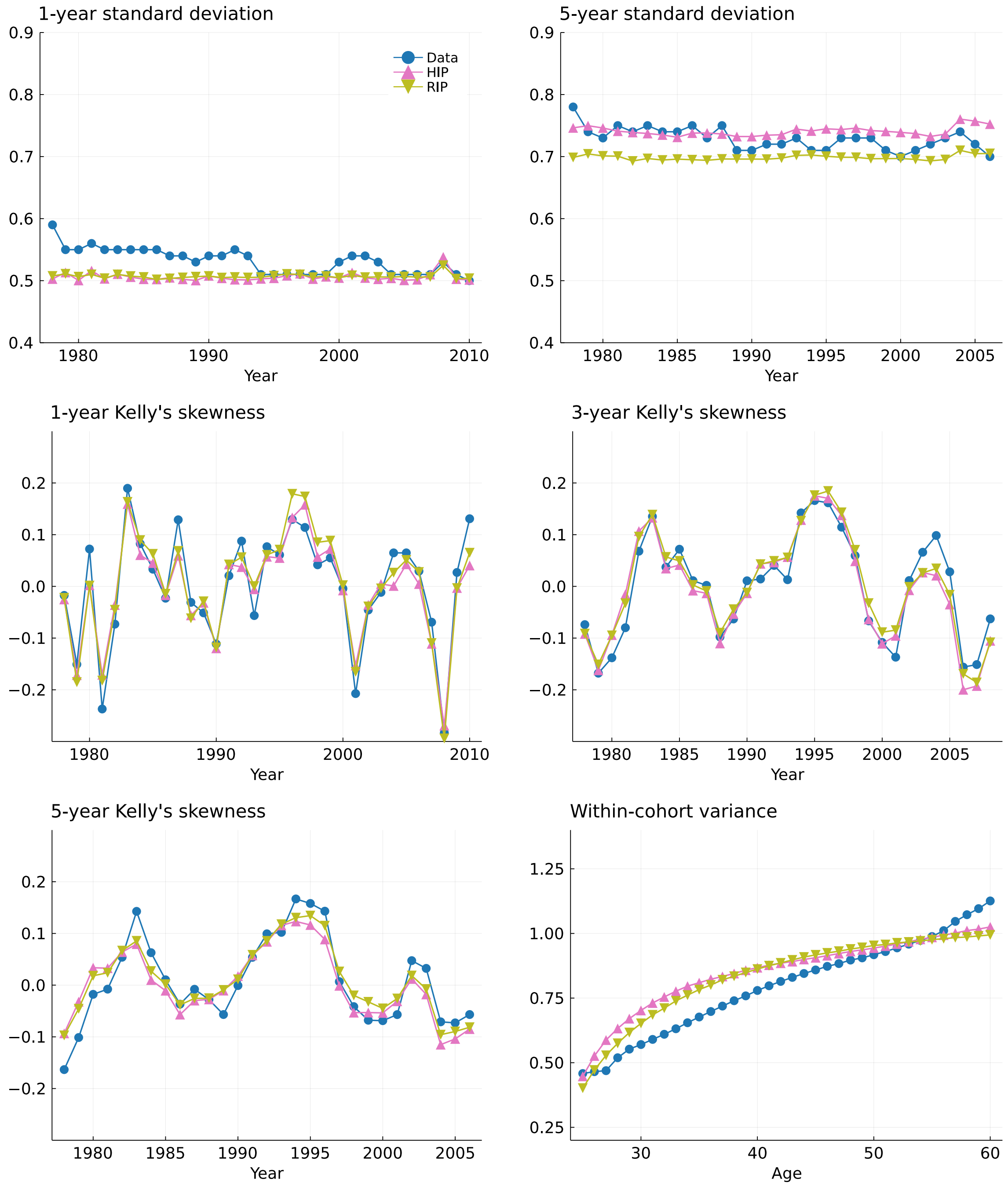


Figure 1: Individual income moments, comparison between different model specifications and data.

The table also reports estimated parameters when cyclical skewness is turned off.[11] In this

and 2011. The same procedure is used in Guvenen et al. (2014) and Catherine (2021).

[11] In the model without cyclical skewness the shocks ε and ν are i.i.d. Normals with zero mean and standard deviations σ_ε and σ_ν, respectively. For readability these two values are reported in the same column of the tail-event standard

case, I match only the standard deviation of log earnings growth at the one and five year horizons and the within-cohort variance of log earnings for ages between 25 and 60. Figure A.1 compares the moments generated by the estimated specifications without cyclical skewness against the data.

	Persistent						Transitory		Income profiles		
	μ_ε	$\lambda_{\varepsilon w}$	$\underline{\sigma}_\varepsilon$	$\overline{\sigma}_\varepsilon$	ρ	p_ε	$\underline{\sigma}_\nu$	$\overline{\sigma}_\nu$	σ_α	σ_β	$\sigma_{\alpha\beta}$
HIP	−0.095	4.486	0.657	0.046	0.929	0.192	0.603	0.124	0.393	0.015	−0.005
RIP	−0.068	3.500	0.611	0.048	0.967	0.163	0.744	0.085			
Without cyclical skewness											
HIP			0.223		0.830		0.357		0.665	0.022	−0.012
RIP			0.216		0.977		0.352				

Table 2: Estimated parameters for the stochastic processes governing individual income.

In addition to reporting the parameter estimates, Table 2 shows an important result: in the model with cyclical skewness, the autocorrelation coefficient ρ is quite similar between HIP and RIP, while it is very different between them when cyclical skewness is turned off. The explanation is the following. Without cyclical skewness, the model attributes the total income risk faced by the agents on the persistence coefficient (ρ), on the variances of the transitory and persistent shocks ($\sigma_\nu, \sigma_\varepsilon$) and, in the HIP case, to the heterogeneity of the parameters governing the life-cycle income profiles ($\sigma_\alpha, \sigma_\beta, \sigma_{\alpha\beta}$). As it is evident by looking at the estimates, in the RIP case the risk due to heterogeneity in life-cycle profiles is almost entirely loaded on the autocorrelation coefficient, since the estimates for the variances of the shocks are very similar. However, by comparing these results with the higher part of the table, it is possible to see that, when taking into account skewness, the HIP model attributes a lower part of risk to the life-cycle profile component and a higher part to persistence. The estimate for ρ in the RIP case, instead, is lower but does not change much. The conclusion is, therefore, that including skewness is crucial not to overestimate the role played by the heterogeneity in life-cycle profiles in explaining income risk against the role played by persistence of the shocks.

Average life-cycle income profile. I model the average life-cycle income profile $\bar{f}(t)$ as a 3^{rd} degree polynomial of age. I fit the polynomial on average log income by age reported by Guvenen et al. (2021) for the age range 25-60 from which, following Catherine (2021), I subtract

deviation in the full specification.

a 15% average income tax.[12][13] Table 3 reports the estimated coefficients while Figure A.2 plots the estimated profile against the data.

Other parameters. The remaining parameters are chosen as reported in Table 3. Agents' starting age is set to 23, they retire at 64 and they die with certainty at age 100.[14] Survival probabilities are taken from US life tables provided by SSA. I set the discount factor δ to 0.927 and the elasticity of intertemporal substitution $1/\psi$ to 0.336 as in Catherine (2021). I choose a standard value in the literature, namely 5 for risk aversion γ. The parameter governing the strenght of the bequest motive b is set to 2.5 as in Gomes and Michaelides (2005). As in Catherine (2021) the risk-free rate is set to 2% and the management fee to 1%. I set the fixed participation cost to 1.5% of the average wage. Finally, the borrowing limit is set to zero.

[12] I do not normalize age, that is, the coefficients are estimated using the age range 25-60.

[13] The values for ages outside this range are then extrapolated from the estimated model.

[14] The reference for these parameters is Catherine (2021), but these values are standard in the literature.

Parameter	**Value**	**Description**	**Source/Target**
Life-cycle			
T_{start}	23	Initial age	Catherine (2021)
K	64	Retirement age	Catherine (2021)
T	100	Maximum life span	Catherine (2021)
p_t	US life tables	Survival probabilities	SSA
Constant	−5.087	Average life-cycle income polynomial coefficients	Average log income by age from Guvenen et al. (2021)
t	0.249		
$t^2/10$	−0.042		
$t^3/100$	0.002		
Preferences			
δ	0.927	Discount factor	Catherine (2021)
γ	5	Risk aversion	Preset
ψ	1/0.336	Inverse EIS	Catherine (2021)
b	2.5	Bequest motive	Gomes and Michaelides (2005)
Financial markets and borrowing limit			
r_f	0.02	Risk-free rate	Catherine (2021)
κ_m	0.01	Management fee	Catherine (2021)
κ_f	0.015	Fixed participation cost	Preset
$\bar{s}$	0	Borrowing limit	Preset

Table 3: Other parameters.

4 Results

4.1 Life-cycle profiles

Figure 2 plots the average simulated life-cycle profiles for participation, conditional and unconditional risky share and wealth during working life for three-year age groups.[15][16] I compare the model-generated patterns against data from the eleven waves of the Survey of Consumer Finances (1989-2019). More details on the data and the computation of the empirical profiles can be found in Appendix B.

[15] I use savings $s_{i,t}$ as measure of wealth in the model.

[16] Figure A.3 plots the profiles for a longer life-cycle horizon and also for consumption and earnings.

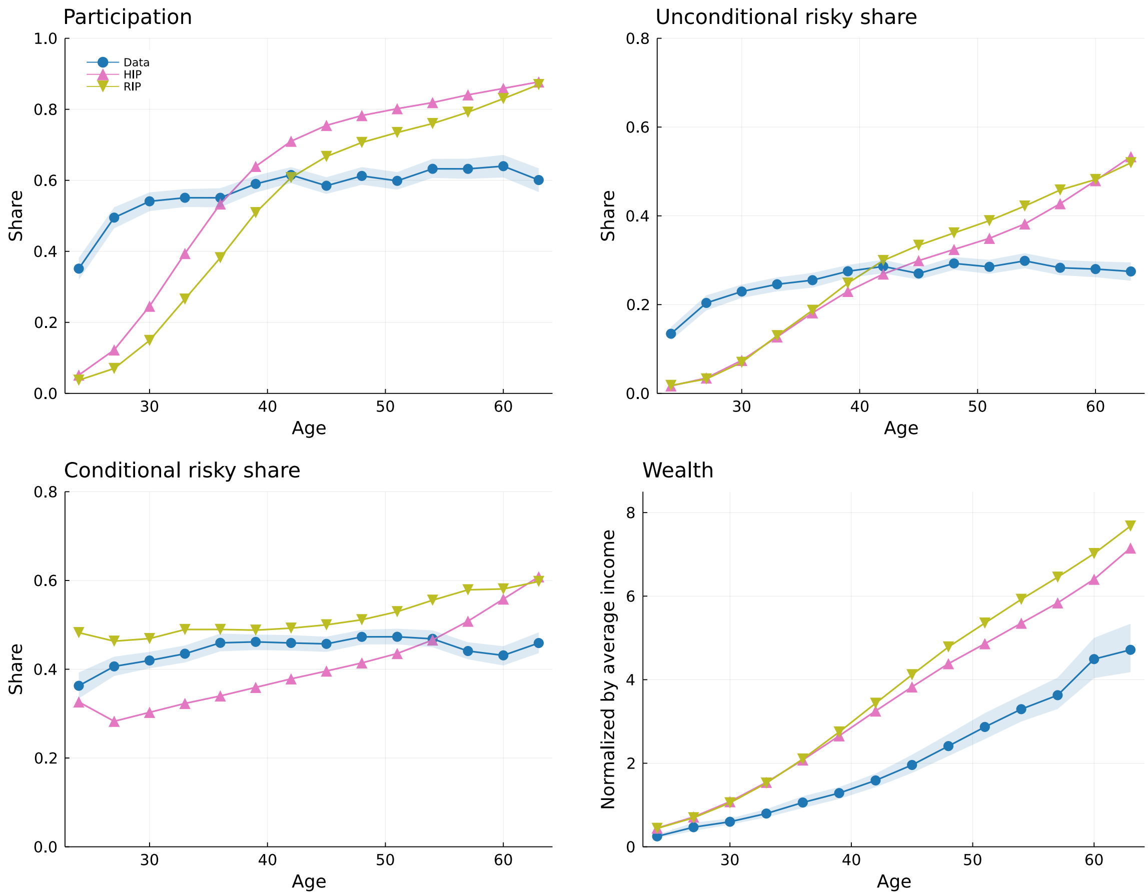


Figure 2: Average simulated life-cycle profiles.

As expected, the shapes of the profiles for both RIP and HIP are overall similar to those found by Catherine (2021). Participation is lower at young ages and slightly higher after forty years old compared to the data, but it has the same shape. The same holds for the unconditional risky share. The conditional risky share, instead, matches quite well both the level and the shape of its empirical counterpart. Wealth is higher than in the data across the whole life cycle.[17]

Analysing the differences between HIP and RIP, it is clear from the picture that there are no significant discrepancies between the two specifications when looking at the average life-cycle profile of wealth and unconditional risky share. Instead, participation is higher and the conditional risky share lower in the HIP model compared to the RIP case across the whole working life.

The reason why the wealth (and consequently consumption, as explained better in the next subsection) profiles are similar between HIP and RIP is related to the result that the estimated autocorrelation coefficients ρ of the AR(1) process governing the persistent component of income

[17]Given the known difficulties of these models to replicate the empirical patterns and the fact that in this version I do not estimate preference parameters, the match with the data is relatively good.

are very similar in both models, as previously described in Section 3. As shown by Carroll (1992) and Gourinchas and Parker (2002), consumption growth parallels income growth in a life-cycle model where income shocks are very persistent, which translates into very similar wealth profiles.

Turning to the conditional risky share, it is lower in the HIP model because of the heterogeneity in income growth rates implied by the different β_i, which determine the slope of the age trend in the individual-specific part of the income process. Specifically, agents in the HIP model with higher than average β_i - recall that the βs are known from the beginning of life and there is no uncertainty on them - while having higher expected income in levels, because of cyclical skewness they also have a larger part of their total wealth (defined as financial wealth plus human capital) that is risky. Thus, as shown in Catherine (2021), to hedge against this, they optimally choose a lower risky share. In turn, because these agents participate more in the market for the risky asset, this results in a lower average conditional risky share.

Regarding participation, the reason why it is higher in the HIP model has to do again with the heterogeneity in income growth rates. More in detail, imagine there is a wealth-to-income level above which it is optimal for agents to enter the market of the risky asset. If - as in the HIP model - there is a fraction of agents who have lower income growth than the average, following the same argument in the previous paragraph, because they have both lower expected income but also a smaller risky part of total wealth, they have a lower wealth-to-income threshold, and they thus participate more in the market for the risky asset. In turn, because the opposite mechanism for people with higher than average income growth is not as strong (as they accumulate wealth faster), these individuals push upwards the average participation rate.

Although the evidence presented in this section has shown that HIP and RIP imply different profiles of average participation and conditional risky share over the life cycle, just looking at these two schedules is not sufficient to tell them apart. Indeed, while the RIP model matches slightly better both the level and the slope of the conditional risky share, it also overshoots it more at older ages and it also undershoots more participation at the beginning of the life cycle.

The next two sections will thus investigate more in detail the reasons behind the similarities and differences in the profiles generated by the HIP and RIP models just described.

4.2 Consumption mean and variance over the life cycle

Even though life-cycle moments of consumption are not usually the interest of portfolio choice models, since previous literature has used the life-cycle patterns of the cross-sectional mean and variance of consumption to discern between HIP and RIP (Guvenen, 2007), in this Section I look more in detail at the model response of these two quantities.

Figure 3 plots these profiles. Average log consumption features the usual hump-shaped pattern due to consumption smoothing and the variance is increasing over the life cycle. [18]

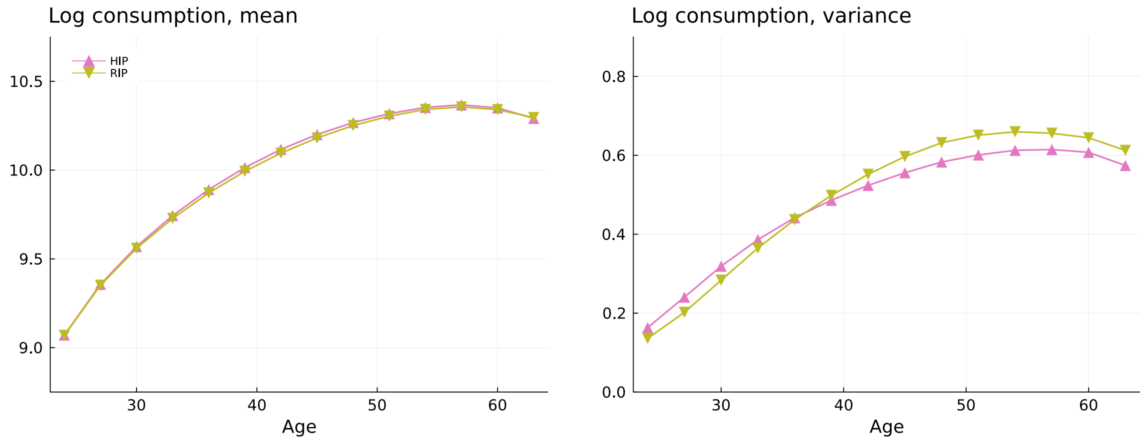


Figure 3: Life-cycle profiles of cross-sectional mean and variance of log consumption.

Importantly, the graph confirms the similarity of the two profiles between HIP and RIP previously found for wealth. As previously explained, when the persistence of income shocks is high, in these kinds of models consumption growth closely tracks income growth. Because the effect on average consumption coming from agents with higher than average income growth is compensated by the opposite one coming from agents with lower than average income growth, what drives the patterns in both specifications are average income growth individuals which, in turn, results in the very similar profiles depicted in the picture.

To sum up, when cyclical skewness is properly taken care of, the estimates of the persistence of the income shocks are very similar between HIP and RIP which, in turn, translates into very similar profiles for the mean and variance over the life-cycle of consumption. Consequently, the identification power to discern HIP and RIP coming from these two series is limited.

4.3 Identifying restrictions from portfolio choices

As already described in Section 4.1, the average life-cycle profiles of participation and conditional risky share differ between HIP and RIP because of the diverse income growth rates distributions implied by the two models. In this section, I explain more in detail the sources of the differences and present identifying restrictions that can be used to test the two in the data.

[18] Figure A.4 also compares them against consumption data obtained from Krueger and Perri (2006). Despite the fact that the levels of the model-generated profiles do not match the data - the estimation matched the life-cycle variance of earnings from a different dataset featuring higher levels than those in their paper - the shapes are right. More details on the data and on the computation of the empirical profiles can be found in Appendix B.

To disentangle the average effects previously found, Figure 4 plots the life-cycle patterns for participation and conditional risky share for HIP and RIP averaging across agents classified according to their average income growth rate during working life.[19][20] The number of the group is increasing in the average growth rate: agents with the lowest growth rates are in group 1. Note that the values depicted for the conditional risky share are expressed in relative terms to group 3.[21]

Considering first participation, the pictures clearly show that the levels for individuals in the high and medium income growth groups are very similar while individuals in the low part of the distribution of income growth rates participate more in the HIP model. The graphs, therefore, confirm the explanation provided in section 4.1: because of cyclical skewness, in the HIP model individuals with lower than average income growth have an expected smaller risky component of total wealth, which is balanced by higher participation rates. At young ages, when the human capital component of total wealth is large, this force is so strong that there is almost no heterogeneity across the groups. Conversely, because in both models individuals with high income growth rates accumulate wealth fast - and thus reach faster the relevant wealth-to-income threshold for participation - their rates are not very much different between the two specifications. Another interesting difference is the dispersion of the profiles: agents tend to have more similar profiles across groups in the HIP model.

Inspecting the conditional risky share, the graph reveal instead a "butterfly pattern" in the HIP model which is absent in the RIP case. Indeed, individuals in higher income growth groups have lower conditional risky shares than agents in lower groups until around age 40, when this pattern is reversed. The picture, therefore, supports the explanation provided when looking at the average life-cycle profiles: in the HIP model, because of cyclical skewness, individuals with higher than average income growth have an expected higher risky part of total wealth, which is hedged with lower risky share. The opposite holds for agents in low income growth groups. Again, this mechanism is stronger at young ages, when the human capital component of total wealth is large. As individuals age and accumulate wealth, however, they self-insure and, together with the fact that the share of human capital in total wealth gets smaller, they take more risk, which explains the reversion of the pattern at around age 40.

---

[19] Figure A.6 plots the profiles for all ages and also for the other variables.

[20] I split the distribution of income growth rates in the two models into five groups: agents below the 20^{th} percentile, between the 20^{th} and the 40^{th} percentiles, between the 40^{th} and the 60^{th} percentiles, between the 60^{th} and the 80^{th} percentiles and above the 80^{th} percentile.

[21] Figure A.6 reports the levels for all variables and a longer life-cycle horizon.

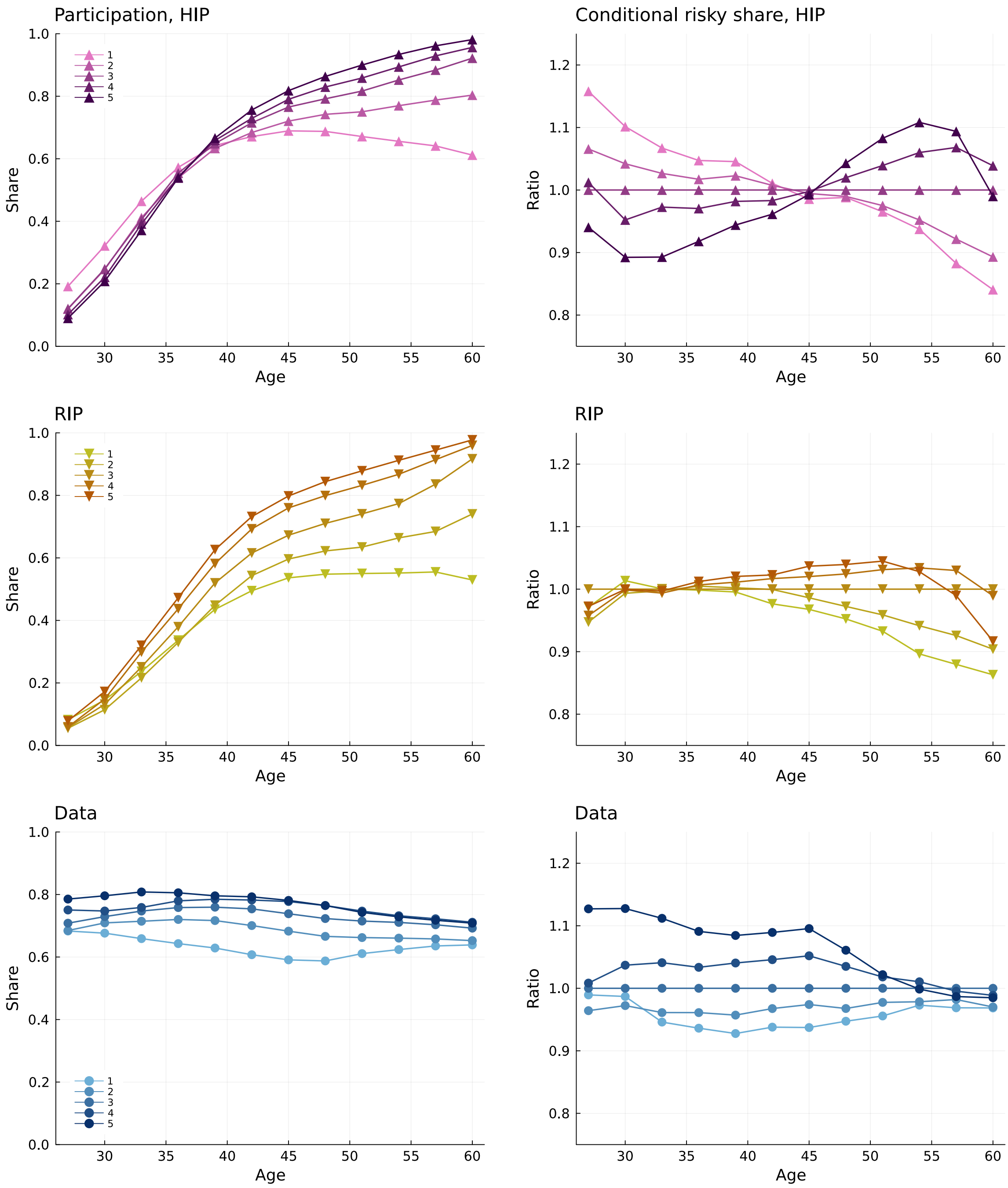


Figure 4: Average simulated life-cycle profiles conditioning on average income growth rate over individuals' working lives in HIP and RIP vs. data. Group numbers are increasing in the growth rate, with 1 containing the individuals with growth rates in the lowest part of the income growth rate distribution and 5 in the highest part. Values for conditional risky share are expressed relatively to group 3.

Summing up, this section has shown that, while - as explained in Section 4.2 - there does not seem to be much identification power to test HIP and RIP using consumption moments over the life-cycle, testable identifying restrictions can be found by looking at portfolio choices of participation and conditional risky share over the distribution of life-cycle average income growth rates. Testing these restrictions in the data is exactly what I do in the next section.

4.4 Testing the restrictions in the data

The model-based implications of the two specifications depicted in Figure 4 and described above can be tested empirically with panel data on individual income and wealth. Because the Survey of Consumer Finances is constructed as a repeated cross-section, it cannot be directly used for this purpose. Thus, the content of this section relies on data from the Swedish Wealth and Income Registry spanning the period 1994-2015.[22][23] These data include yearly variables on demographic characteristics, income, and wealth holdings at the individual level for the whole universe of Swedish residents.[24] For the purpose of my analysis, the variables needed are age, a measure of income, and a measure of the risky share, all at the individual level. Regarding income, the variable used is a series of non-financial disposable income based on the definition by Statistics Sweden, which spans the period 1994-2015. Following Catherine et al. (2021), instead, the risky share is defined as the ratio between the sum of wealth invested in stocks and funds over the sum of these two variables with cash.[25][26]

Given the large number of individuals in the simulation and the fact that agents know from the beginning of the life cycle the parameters governing the deterministic part of their income, classifying them into different average income growth groups can be done straightforwardly by computing the average yearly income log growth rate over their working life. The same is not true in the data because of the presence of other confounding factors (e.g. agents' information, cohort and year effects, etc.). Therefore, the following procedure is adopted. First, the average log growth of income in all the working age years (from 23 to 65 years old) is computed for each individual, using the whole sample of available income data, i.e. 1994-2015. Second, this measure is regressed on dummy variables controlling for the agent's age in the first

[22]As the model is estimated on US data, the underlying assumption behind this comparison is that the patterns in the Swedish context are not very different from those in the US.

[23]For more details on this comprehensive dataset, see Catherine et al. (2021) and the references cited therein.

[24]The wealth data used cover the years 2000-2007. They include bank accounts, mutual funds, and holdings of stocks, bonds, and derivatives and they were collected due to a wealth tax. A detailed description of this dataset is available in Bach et al. (2020).

[25]Financial wealth and its components are defined as in Bach et al. (2020).

[26]The final sample used includes individuals between 23 and 91 years old, without business income, with non-financial disposable income and financial wealth above 10000 SEK (in 2015 terms) and below the 0.999 quantile of their respective distributions. Furthermore, only individuals without large real estate transactions and present in all the years in which income and wealth data, respectively, are available are considered. Finally, individuals with income-to-financial wealth ratios below and above the 0.15 and 0.85 quantiles of the distribution (respectively, about 0.5 and 10) of this variable are also excluded.

year of available wealth data. The resulting residual is then used as the relevant measure for classifying individuals in different average income growth groups. Finally, the life-cycle profiles for participation and conditional risky share are computed using the procedure described in Appendix B for all the five groups.[27]

The bottom part of Figure 4 reports the results.[28] Participation is relatively stable across the life cycle, and, in general, slightly increasing at young ages. The pattern emerging across the different income growth groups is clear: participation for group 1 is always the lowest and the other groups follow in increasing order. A similar trend (except for group 1 at the very beginning and for above 50 years old, where the lines start to overlap) is also visible when considering the conditional risky share.

What can be inferred from these empirical moments? Despite the fact that the level in both model specifications at the beginning of the life cycle is lower than in the data[29], the order and variation of the participation rate between groups at young ages in the data resemble more the RIP case. Analysing the conditional risky share, instead, reveals that the data clearly do not support the "butterfly pattern" found for the HIP specification. Even though the RIP case shows a counterfactual overlap between the groups at young ages, in this specification an order and variation of the schedules among groups more in line with the data is achieved earlier in the life cycle.

Summing up, although testing the restrictions in the data in the way described in this section has not delivered a conclusive answer, the data seem to support slightly more the RIP hypothesis, especially because no evidence of a "butterfly pattern" for the conditional risky share was found.

5 Conclusion

In this paper I have investigated what inference on the properties of the income process can be drawn from a state of the art portfolio choice model.

First, I have documented that cyclical skewness needs to be included in the stochastic process for income in order to correctly estimate the amount of risk deriving from the persistence of the shocks. Indeed, the HIP model without cyclical skewness overestimates the share of risk attributed to the heterogeneity in life-cycle profiles and underestimates the share deriving from

[27] I kindly thank Paolo Sodini for sharing these moments.

[28] The levels for the conditional risky share are reported in Figure A.7. For both participation and the conditional risky share the values obtained are in line with what reported in Catherine et al. (2021).

[29] This is even more clear with Swedish data where, compared to the pattern obtained with data from the Survey of Consumer Finances previously described, participation is higher at young ages. The level of the conditional risky share, instead, is quite similar.

persistence.

Second, when the income process includes cyclical skewness, I find that the estimated autocorrelation coefficient in the AR(1) process for the persistent component of income is similar for HIP and RIP. Because consumption growth parallels income growth in life-cycle models when the persistence of the shocks is high this, in turn, implies very similar consumption and wealth profiles in both specifications. Therefore, the cross-sectional mean and variance of the wealth and consumption profiles over the life-cycle do not have enough identification power to disentangle between HIP and RIP.

Third, I have documented that the patterns of participation and conditional risky share across the distribution of average working life income growth rates have identification power to discern between HIP and RIP. Specifically, compared to the RIP case, the distribution of income growth rates in the HIP process determines less heterogeneity across income groups for participation and a "butterfly pattern" for the conditional risky share, which can both be tested empirically. Although the data did not deliver a conclusive answer, more support was found for the RIP case, especially because no evidence of the "butterfly pattern" was discovered.

This work opens the avenue to future research in several ways. While this paper has used it to reach other conclusions, the result that the persistence of income shocks is similar across HIP and RIP when cyclical skewness is taken care of is very interesting and I am investigating it more in detail in ongoing research. Furthermore, I have focused only on a specific set of moments: additional work is needed to check whether other moments contain useful information for identification. In addition, the data patterns used for testing the restrictions considered the average agent in each income growth group: it would be interesting to test the robustness of the results for agents differently exposed to cyclical skewness (Catherine et al., 2021), or with different wealth-to-income ratios. Lastly, this study has focused on a particular question, namely inferring from portfolio choices whether the income process is more in line with the HIP or RIP hypothesis. Future work should use the approach outlined in this paper to look at income properties from revealed portfolio choices in more general terms.

References

Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme, "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework," *Econometrica*, 2017, *85* (3), 693–734.

__ , **Stéphane Bonhomme, Micole De Vera, Laura Hospido, and Siqi Wei**, "Income risk inequality: Evidence from Spanish administrative records," *Quantitative Economics*, 2022, *13* (4), 1747–1801.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini, "Rich Pickings? Risk, Return, and Skill in Household Wealth," *American Economic Review*, September 2020, *110* (9), 2703–47.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, "Portfolio Choice over the Life-Cycle when the Stock and Labor Markets Are Cointegrated," *The Journal of Finance*, 2007, *62* (5), 2123–2167.

Blundell, Richard and Ian Preston, "Consumption Inequality and Income Uncertainty," *The Quarterly Journal of Economics*, 1998, *113* (2), 603–640.

—, Luigi Pistaferri, and Ian Preston, "Consumption Inequality and Partial Insurance," *American Economic Review*, December 2008, *98* (5), 1887–1921.

Busch, Christopher, David Domeij, Fatih Guvenen, and Rocio Madera, "Skewed Idiosyncratic Income Risk over the Business Cycle: Sources and Insurance," *American Economic Journal: Macroeconomics*, April 2022, *14* (2), 207–42.

Calvet, Laurent E., John Y. Campbell, Francisco J. Gomes, and Paolo Sodini, "The Cross-Section of Household Preferences," *Working Paper*, 2021.

Carroll, Christopher D., "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 1992, *23* (2), 61–156.

Catherine, Sylvain, "Countercyclical Labor Income Risk and Portfolio Choices over the Life Cycle," *The Review of Financial Studies*, 12 2021. hhab136.

—, Paolo Sodini, and Yapei Zhang, "Countercyclical Income Risk and Portfolio Choices: Evidence from Sweden," *Working Paper*, 2021.

Chang, Yongsung, Jay H. Hong, and Marios Karabarbounis, "Labor Market Uncertainty and Portfolio Choice Puzzles," *American Economic Journal: Macroeconomics*, April 2018, *10* (2), 222–62.

—, Jay Hong, Marios Karabarbounis, and Yicheng Wang, "Income Volatility and Portfolio Choices," 2018 Meeting Papers 412, Society for Economic Dynamics 2018.

Cocco, João F., Francisco J. Gomes, and Pascal J. Maenhout, "Consumption and Portfolio Choice over the Life Cycle," *The Review of Financial Studies*, 02 2005, *18* (2), 491–533.

Deaton, Angus and Christina Paxson, "Intertemporal Choice and Inequality," *Journal of Political Economy*, 1994, *102* (3), 437–467.

Fagereng, Andreas, Luigi Guiso, and Luigi Pistaferri, "Portfolio Choices, Firm Shocks, and Uninsurable Wage Risk," *The Review of Economic Studies*, 04 2017, *85* (1), 437–474.

Gomes, Francisco and Alexander Michaelides, "Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence," *The Journal of Finance*, 2005, *60* (2), 869–904.

— and —, "Asset Pricing with Limited Risk Sharing and Heterogeneous Agents," *The Review of Financial Studies*, 11 2008, *21* (1), 415–448.

Gourinchas, Pierre-Olivier and Jonathan A. Parker, "Consumption over the Life Cycle," *Econometrica*, 2002, *70* (1), 47–89.

Guvenen, Fatih, "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *The American Economic Review*, 2007, *97* (3), 687–712.

—, "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, January 2009, *12* (1), 58–79.

— and Anthony A. Smith, "Inferring Labor Income Risk and Partial Insurance From Economic Choices," *Econometrica*, 2014, *82* (6), 2085–2129.

—, Fatih Karahan, Serdar Ozkan, and Jae Song, "What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?," *Econometrica*, 2021, *89* (5), 2303–2339.

—, Serdar Ozkan, and Jae Song, "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 2014, *122* (3), 621–660.

Hall, Robert E. and Frederic S. Mishkin, "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, 1982, *50* (2), 461–481.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "Two Views of Inequality over the Life Cycle," *Journal of the European Economic Association*, 2005, *3* (2/3), 765–775.

Huggett, Mark and Greg Kaplan, "How large is the stock component of human capital?," *Review of Economic Dynamics*, 2016, *22*, 21 – 51.

Inkmann, Joachim, Paula Lopes, and Alexander Michaelides, "How Deep Is the Annuity Market Participation Puzzle?," *The Review of Financial Studies*, 09 2010, *24* (1), 279–319.

Kaplan, Greg and Giovanni L. Violante, "How Much Consumption Insurance beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, October 2010, *2* (4), 53–87.

Krueger, Dirk and Fabrizio Perri, "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *The Review of Economic Studies*, 01 2006, *73* (1), 163–193.

Lillard, Lee A. and Robert J. Willis, "Dynamic Aspects of Earning Mobility," *Econometrica*, 1978, *46* (5), 985–1012.

MaCurdy, Thomas E., "The use of time series processes to model the error structure of earnings in a longitudinal data analysis," *Journal of Econometrics*, 1982, *18* (1), 83–114.

Meghir, Costas and Luigi Pistaferri, "Income Variance Dynamics and Heterogeneity," *Econometrica*, 2004, *72* (1), 1–32.

Merton, Robert, "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case," *The Review of Economics and Statistics*, 02 1969, *51*, 247–57.

Nardi, Mariacristina De, Giulio Fella, and Gonzalo Paz-Pardo, "Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare," *Journal of the European Economic Association*, 03 2019, *18* (2), 890–926.

Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron, "Asset pricing with idiosyncratic risk and overlapping generations," *Review of Economic Dynamics*, 2007, *10* (4), 519–548.

__ **, Chris I. Telmer, and Amir Yaron**, "Cyclical Dynamics in Idiosyncratic Labor Market Risk," *Journal of Political Economy*, 2004, *112* (3), 695–717.

A Additional figures

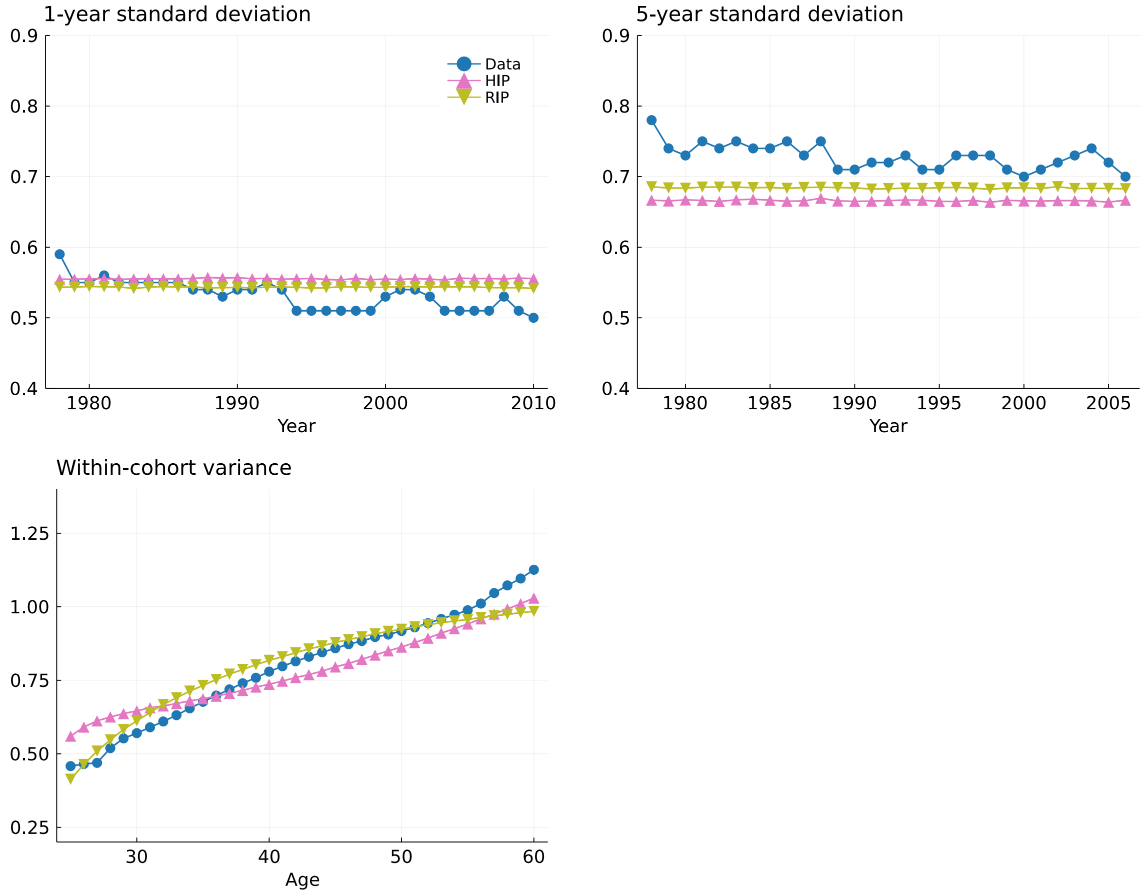


Figure A.1: Individual income moments, comparison between different model specifications without cyclical skewness and data.

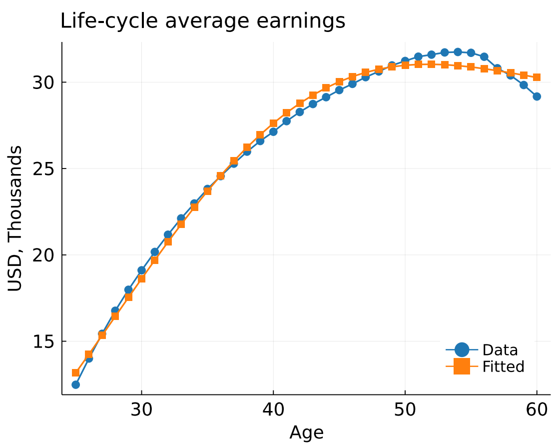


Figure A.2: Fitted age polynomial vs. data

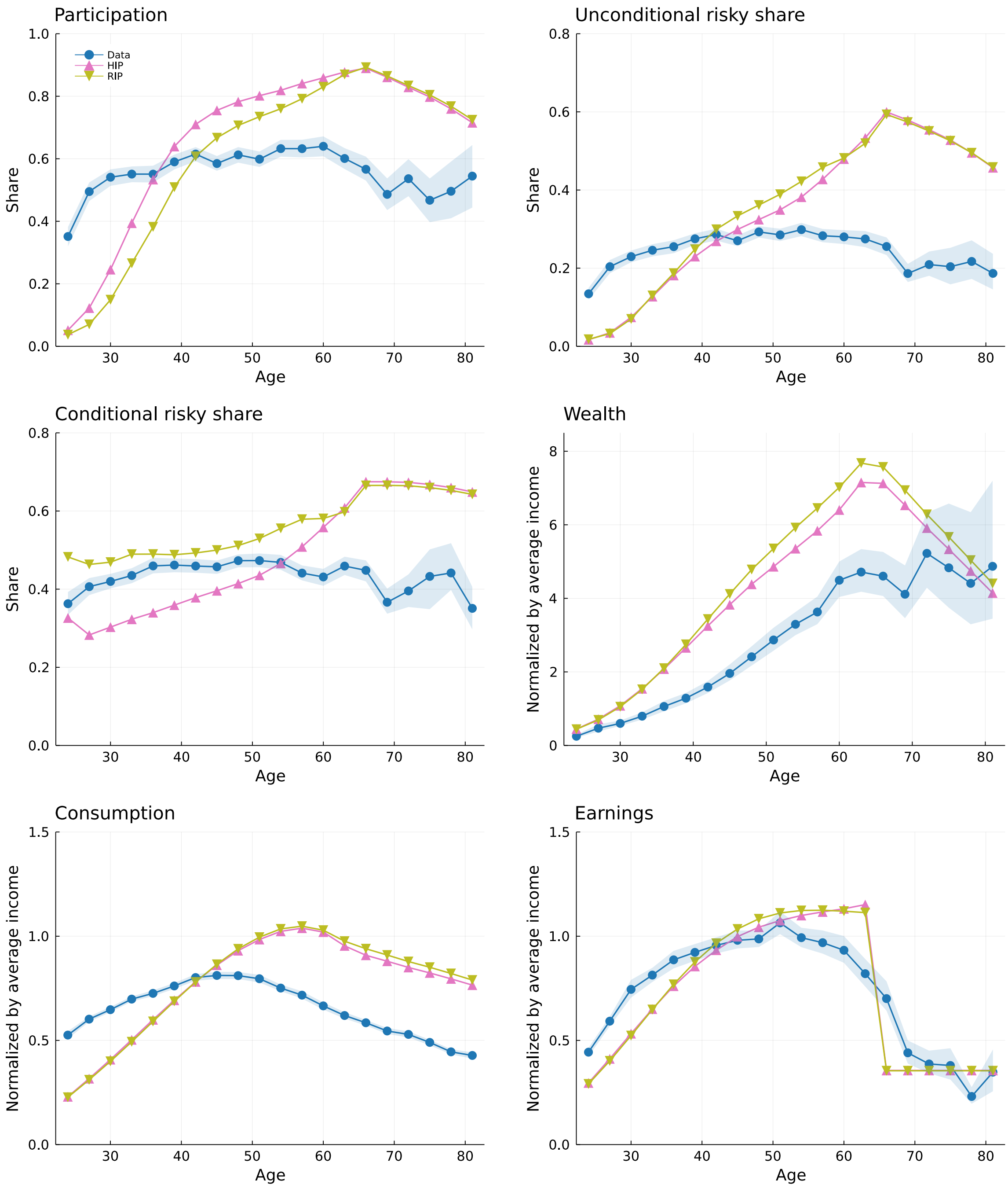


Figure A.3: Average simulated life-cycle profiles from the model, all ages and variables. Consumption data are from Krueger and Perri (2006) and data for the other variables from the Survey of Consumer Finances. More details on the data can be found in Appendix B.

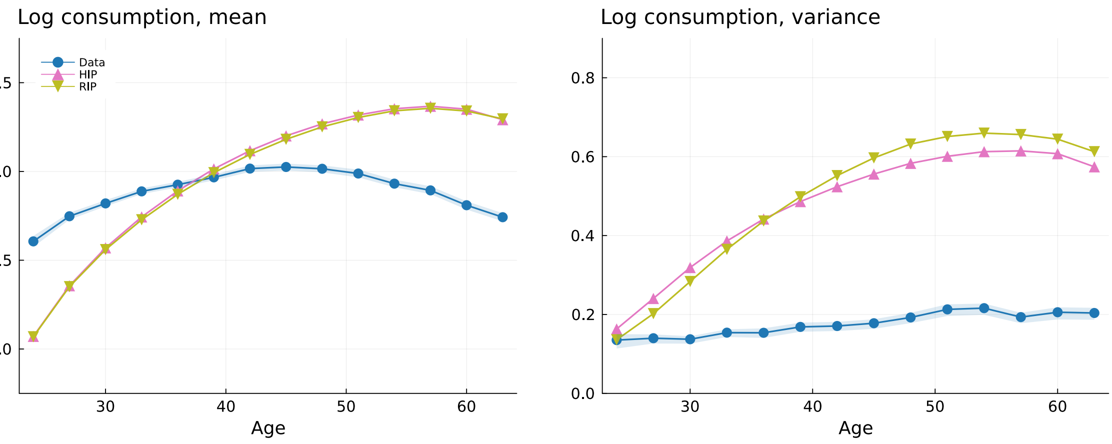
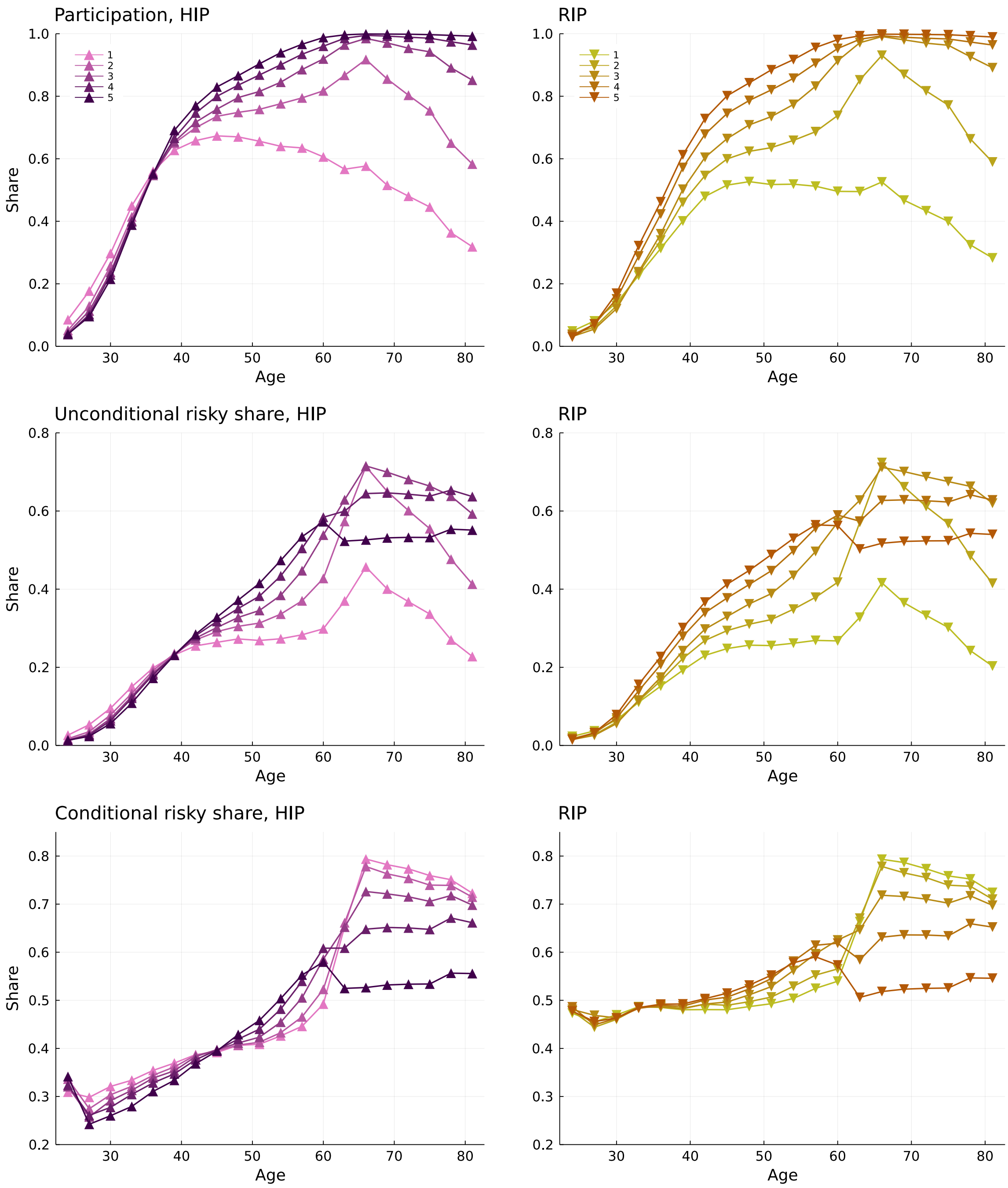


Figure A.4: Life-cycle profiles of cross-sectional mean and variance of log consumption vs. data from Krueger and Perri (2006).


Participation, HIP
RIP
Unconditional risky share, HIP
RIP
Conditional risky share, HIP
RIP
Share
Age
1
2
3
4
5

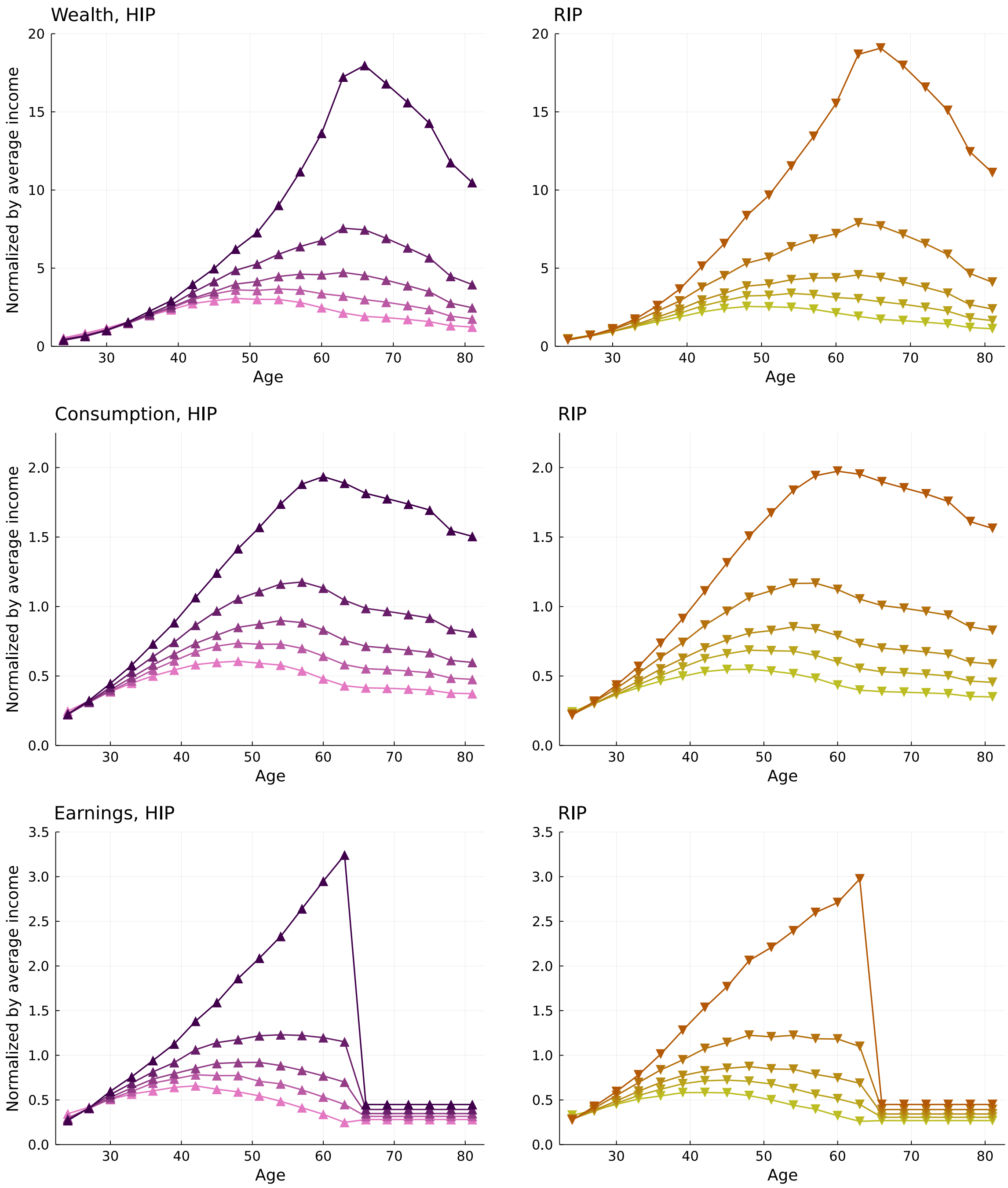


Figure A.6: Average simulated life-cycle profiles conditioning on average income growth rate over individuals' working lives, all ages and variables, HIP and RIP. Group numbers are increasing in the growth rate, with 1 containing the individuals with growth rates in the lowest part of the income growth rate distribution and 5 in the highest part.

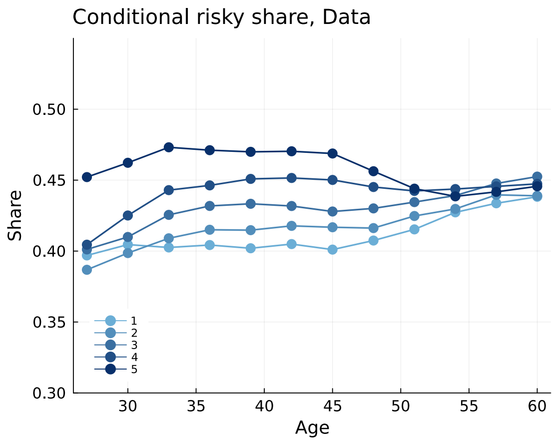


Figure A.7: Average life-cycle profile for conditional risky share conditioning on average income growth rate over individuals' working lives, data. Group numbers are increasing in the growth rate, with 1 containing the individuals with growth rates in the lowest part of the income growth rate distribution and 5 in the highest part.

B Data

This section describes in detail the data sources used in the paper.

Demographics. Survival probabilities are taken from the 2019 actuarial life table compiled by US Social Security Administration, available at this link https://www.ssa.gov/oact/STATS/table4c6.html.

Aggregate variables. Data for the return on the risky asset are taken from Robert Shiller's website (http://www.econ.yale.edu/~shiller/data/ie_data.xls). I use the variable "Real Total Return Price". I get a yearly series by taking the monthly value at the beginning of the year. The log return in year t is then the log difference between the yearly price at $t+1$ and t.

For the wage index I use two sources. From 1944 to 1978 I compute log growth rates using data from Emmanuel Saez available at this link http://eml.berkeley.edu/~saez/TabFig2012prel.xls. Specifically, I use the variable "Average wage income ($ latest year)" in Table B1. For the period 1979-2011 I use the log growth rates reported by Guvenen et al. (2014) in Table A1 (available at this link https://fatihguvenen.com/s/gos-jpe2014-data.xlsx) for the variable "Change in log earnings averaged over workers x100".

To deflate nominal variables I use the CPI index for the US (reference year 2010) from the World Bank, which is available at their online database.

Individual income estimation. For estimation of the stochastic process governing individual income I use again the values reported by Guvenen et al. (2014) (available at this link `https://fatihguvenen.com/s/gos-jpe2014-data.xlsx`). Specifically, I use the time series for standard deviation of earnings growth at the 1 and 5 year horizons reported in Table A8 and the time series for the 10th, 50th and 90th percentiles of the distribution of earnings growth at the 1, 3 and 5 year horizons reported in Table C1.

In addition, I use within-cohort variances of log earnings and average log earning by age reported, respectively, in the sheets "Figure D3" and "Figure C36" compiled by Guvenen et al. (2021), available at this link `https://fatihguvenen.com/s/gkos_2021_moments.xlsx`.

Agents' balance sheets. For variables related to agents' balance sheets I rely on the eleven waves of the Survey of Consumer Finances from 1989 to 2019. More in detail, I use the "summary extract public data", which are available at the Federal Reserve's website. In the description below, variables in italics refer to variable names in the original datasets. Additionally, to ensure comparability across different surveys, I do not use the absolute weights provided in the original data, but their rescaled version (i.e. the original weights divided by their sum in each year). The variables I focus on are: labor earnings (*wageinc*), net worth (*networth*)[30], financial wealth (*fin*), equity (*equity*)[31].

I restrict the sample to households between age 23 and 82 and, in order to filter out entrepreneurs and self-employed people, I remove all the households for whom the variable *bus* is not zero. I further filter out individuals whose labor earnings are lower than 1000 USD and whose net worth is lower than 1000 USD.[32]

Additionally, I create the following variables:

- the ratio between financial wealth and average income, where the latter is defined as the yearly cross-sectional average labor income (for comparability with the model's variables) computed using survey weights, that is $fin/\mathbb{E}\left[wageinc\right]$;
- the ratio between income minus a 15% tax and average income, that is $wageinc/\mathbb{E}\left[wageinc\right]$;
- risky share defined as the ratio between stock holdings and financial wealth, that is $equity/fin$;
- participation defined as a dummy equal to one if the risky share is strictly positive.

[30] This is the sum of financial assets (cash, savings, retirement, investment accounts, . . .), businesses and residential assets, minus all debts.

[31] This is the sum of directly held stocks and stocks own indirectly through mutual funds and retirement accounts. The survey asks households whether these accounts are invested mostly into bonds or stocks and imputes a fraction of the total value of the account to the equity variable based on the response.

[32] The Survey of Consumer Finances' extracts at the time of download are in 2019 USD.

Agents' consumption. I use the data compiled by Krueger and Perri (2006).[33]. In the description below, variables in italics refer to variable names in their dataset. In addition to the sample restrictions already present in the dataset available for download (incomplete income respondents, households who report 0 USD in food consumption, households who only report only food consumption), I apply other similar restrictions as they do in their paper: I remove observations with positive labor income but no hours worked and I restrict to households completing all the interviews. Then, as they do in their paper, I classify an household as belonging to year t if the last interview was conducted between the second quarter of year t and the first quarter of year $t + 1$, I define yearly income values as those reported in the last interview and and yearly consumption as the sum of quarterly consumption reported in each of the four interviews. The income measure I use is total income before taxes (*incbetax*) and the relevand consumption measure I consider is the one constructed by the authors (*ndpbe0*). In addition, I also define yearly survey weights for each observation as the sum of survey weights for each of the four interviews. Furthermore, age is defined as the age of the reference person (*refage*) at the time of the last interview. I restrict the sample to households with at least 1000 USD of wealth (defined as the sum of financial wealth (*finwea*) and the value of owned residence (*propval*), at least 1000 USD of income, for which both the reference person and the spouse have zero business income (*refby* and *spoby* equal to zero) and for which age is between 23 and 82.

In addition to the variables already provided in the dataset, I construct an additional variable: the ratio between consumption and average income, where the latter is defined as the yearly cross-sectional average total income before taxes (for comparability with the model's variables) computed using survey weights, that is $ndpbe0/\mathbb{E}\left[incbetax\right]$.

Life-cycle profiles. To construct life-cycle profiles I use a method similar to Heathcote et al. (2005). First, I build 3-year age groups. Then, I compute $m_{a,c,t}$, that is, moment m for households in age group a, with cohort c in year t using survey weights. I then regress these moments on age group and year dummies[34] and recover the age profile for moment m by adding the unconditional average of the coefficients of the time dummies to the coefficients of the age group dummies. Standard errors on such moments are computed by boostrapping the data 1000 times at the cohort and year level.

[33]The data are available at Fabrizio Perri's website `http://www.fperri.net/research_data.htm`.

[34]This is what Heathcote et al. (2005) call the "time view".

C Numerical solution

C.1 Discretization and grids construction

Normally distributed random variables. Let X be an i.i.d. Normally distributed random variable with mean μ_x and standard deviation σ_x. I discretize X using Gaussian quadrature. Specifically, the support of X is approximated with a finite grid of values $x_1, \ldots, x_{N_q}$ computed as follows:

$$x_j = \mu_x + \sqrt{2}\sigma_x Z_j, \quad j = 1, \ldots, N_q$$

where the Z_j's are Gauss-Hermite nodes and the probability mass of each point of the discretized support is computed as:

$$p(x_j) = \omega_j / \sqrt{\pi}, \quad j = 1, \ldots, N_q$$

where the ω_j's are Gauss-Hermite weights. This procedure is used to discretize $r_{2,t}$, ϕ_t, and the distributions of $\nu_{i,t}$ and $r_{1,t}$ conditional, respectively, on tail/non-tail event and on stock market crash/normal period. N_q is the same for all shocks.

Persistent component of idiosyncratic income. I approximate the process governing the evolution of $z_{i,t}$ as follows: (i) I discretize the conditional distribution of $\varepsilon_{i,t}$, (ii) I compute the evolution of the persistent component according to equation (8) and (iii) I evaluate the model functions at the resulting value of $z_{i,t}$. The advantage of this method is that it requires to discretize just the conditional distribution of $\varepsilon_{i,t}$, which is easier than discretizing the full process of $z_{i,t}$. In particular, the crucial connections between the higher moments of $z_{i,t}$ and other variables are preserved. The disadvantage is that the resulting values of $z_{i,t}$ will very often be off grid, so I need a grid of values that captures well the behavior of the model at such points given the interpolation procedure.[35] Given the above discussion, the grids for $\varepsilon_{i,t}$ and $z_{i,t}$ are constructed as follows. The conditional distribution of $\varepsilon_{i,t}$ is discretized using the procedure described above for Normally distributed shocks with N_q points. To set up the grid for $z_{i,t}$, instead, I first construct an exponentially spaced grid of $(N_z - 1)/2 + 1$ points with minimum value equal to zero, maximum value equal to $z_{\max}$ and spacing parameter equal to spacing_z. This gives me the positive side of the grid plus the central point (which is therefore equal to zero). Then, I add the negative $(N_z - 1)/2$ values by taking the negative of the positive values just computed and obtain the full grid of N_z points.

Average income. The grid for average income $\bar{L}_{i,t}$ is an exponentially spaced grid of $N_{\bar{L}}$ points with minimum value equal to the lowest possible realization of $\bar{L}$ implied by the income process

[35] See below for more details on the interpolation method.

and the formula for average income in (16) and maximum value equal to the highest possible realization of $\bar{L}$ implied by the formula, that is 2.5. The spacing parameter is equal to $\text{spacing}_{\bar{L}}$.

Life-cycle parameters. The grid for α_i is an exponentially spaced grid of $(N_\alpha - 1)/2 + 1$ points with minimum value equal to zero, maximum value equal to $\alpha_{\max}$ and spacing parameter equal to $\text{spacing}_{\text{lc}}$. This gives me the positive side of the grid plus the central point (which is therefore equal to zero). Then, I add the negative $(N_\alpha - 1)/2$ values by taking the negative of the positive values just computed and obtain the full grid of N_α points. The grid for β_i is constructed with the same procedure, using N_β points, maximum value $\beta_{\max}$ and spacing parameter $\text{spacing}_{\text{lc}}$.

Cash on hand and savings. The grid for cash on hand is an exponentially spaced grid of $N_{\hat{a}}$ points with minimum value equal to the lowest possible realization of cash on hand implied by the model, maximum value equal to $\hat{a}_{\max}$ and spacing parameter equal to $\text{spacing}_{\hat{a}}$. The grid for savings $\hat{s}_{i,t}$ is an exponentially spaced grid of $N_{\hat{s}}$ points with minimum value equal to $\overline{s}$, maximum value equal to $\hat{s}_{\max}$ and spacing parameter equal to $\text{spacing}_{\hat{s}}$.[36]
Table C.1 summarizes the choices for the numerical parameters.

[36] Variables with a hat on top refer to normalized variables as defined in section C.2 below.

Parameter	Value	Description
Panel A: numerical parameters for model solution		
N_q	3	Number of points Gaussian quadrature
N_z	15	Number of points grid persistent/idiosyncratic income
$N_{\hat{a}}$	51	Number of points grid cash on hand
$N_{\hat{s}}$	$N_{\hat{a}}$	Number of points grid savings
$N_{\bar{L}}$	21	Number of points grid average income
N_α	5	Number of points grid life-cycle constant
N_β	5	Number of points grid life-cycle slope
$z_{\max}$	4.5	Maximum value grid persistent income
$\hat{a}_{\max}$	200.0	Maximum value grid cash on hand
$\hat{s}_{\max}$	$\hat{a}_{\max}$	Maximum value grid savings
$\alpha_{\max}$	$3\sigma_\alpha$ in HIP model	Maximum value grid life-cycle constant
$\beta_{\max}$	$3\sigma_\beta$ in HIP model	Maximum value grid life-cycle slope
spacing_z	1.5	Spacing parameter grid persistent income
$\text{spacing}_{\hat{a}}$	1.25	Spacing parameter grid cash on hand
$\text{spacing}_{\hat{s}}$	$\text{spacing}_{\hat{a}}$	Spacing parameter grid savings
$\text{spacing}_{\bar{L}}$	1.25	Spacing parameter grid average income
$\text{spacing}_{\text{lc}}$	1.25	Spacing parameter grids life-cycle parameters
Panel B: numerical parameters for model simulation		
T_{eco}	1000	Number of different time-series of aggregate shocks to simulate
N_{sim}	1500	Number of agents to simulate
Panel C: numerical parameters for estimation		
N_{glo}	1000	Number of points to evaluate in global stage[a]
N_{loc}	10	Number of points to evaluate in local stage
N_{eco}	5	Number of economies to simulate
T_{cal}	10^5	Number of time-series points to simulate for aggregate shocks
T_{dis}	1000	Number of periods to discard for moments computation
N_{cal}	1500	Number of individuals in each cohort to simulate

Table C.1: Numerical parameters.

[a] For the estimation of the aggregate processes N_{glo} is 5000 and N_{loc} is 50.

C.2 Solving the optimization problem

Whenever it does not lead to confusion I am dropping α_i, β_i and $\bar{L}_{i,t}$ from the state variable. Also, for ease of exposition, I will consider the case in which disposable income coincides with labor income and there are no participation costs: the general case is a straightforward extension. Because the aggregate component of the wage follows a random walk, it is possible to rescale the problem to get rid of w as a state variable as follows. Let $\hat{x}_{i,t} = x_{i,t}/e^{w_t+\bar{f}(t)}$ for a generic variable[37] x and $\hat{V}_{i,t}\left(a_{i,t}, z_{i,t}\right) := V_{i,t}\left(a_{i,t}, z_{i,t}, 0\right)$, so that I can write:

$$V_{i,t}\left(a_{i,t}, z_{i,t}, w_t\right) = e^{w_t+\bar{f}(t)} V_{i,t}\left(\frac{a_{i,t}}{e^{w_t+\bar{f}(t)}}, z_{i,t}, 0\right) = e^{w_t+\bar{f}(t)} \hat{V}_{i,t}\left(\hat{a}_{i,t}, z_{i,t}\right)$$

Using the above definitions, letting $\Delta w_t := w_t - w_{t-1}$ and $\Delta f_t := \bar{f}(t) - \bar{f}(t-1)$, the optimization problem during working life can be rewritten as follows:

$$\hat{V}_{i,t}(\hat{a}_{i,t}, z_{i,t}) = \max_{\xi_{i,t}, \hat{c}_{i,t}, \hat{s}_{i,t}} \left\{ (1-\delta)\hat{c}_{i,t}^{1-\psi} \right.$$
$$\left. + \delta \left[\mathbb{E}_t \left[\left(p_t \hat{V}_{i,t+1}^{1-\gamma}(\hat{a}_{i,t+1}, z_{i,t+1}) + b(1-p_t)\hat{a}_{i,t+1}^{1-\gamma}\right) e^{(\Delta w_{t+1}+\Delta f_{t+1})(1-\gamma)}\right]\right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}$$

subject to:

$$\begin{aligned} \hat{c}_{i,t} + \hat{s}_{i,t} &= \hat{a}_{i,t} \\ \hat{a}_{i,t+1} &= \left[R_f + \xi_{i,t} R_{t+1}^e\right] \hat{s}_{i,t} e^{-\Delta w_{t+1} - \Delta f_{t+1}} + e^{z_{i,t+1}+\nu_{i,t+1}} \\ \hat{s}_{i,t} &\geq \bar{s} \end{aligned}$$

For the solution, it is useful to define:

$$\tilde{V}_{i,t}(\hat{s}_{i,t}, \xi_{i,t}, z_{i,t}) = \left[\mathbb{E}_t \left[\left(p_t \hat{V}_{i,t+1}^{1-\gamma}(\hat{a}_{i,t+1}, z_{i,t+1}) + b(1-p_t)\hat{a}_{i,t+1}^{1-\gamma}\right) e^{(\Delta w_{t+1}+\Delta f_{t+1})(1-\gamma)}\right]\right]^{\frac{1-\psi}{1-\gamma}}$$

The first order condition with respect to $\xi_{i,t}$ reads:

$$\frac{\partial \tilde{V}_{i,t}(\tilde{s}_{i,t}, \xi_{i,t}, z_{i,t})}{\partial \xi_{i,t}} = 0 \iff$$

$$\mathbb{E}_t \left[e^{-\gamma(\Delta w_{t+1}+\Delta f_{t+1})} R_{t+1}^e \left(p_t \hat{V}_{i,t+1}^{-\gamma}(\hat{a}_{i,t+1}, z_{i,t+1}) \frac{\partial \hat{V}_{i,t+1}(\tilde{a}_{i,t+1}, z_{i,t+1})}{\partial \hat{a}_{i,t+1}} + b(1-p_t)\hat{a}_{i,t+1}^{-\gamma}\right)\right] = 0$$

The first order condition with respect to $\hat{s}_{i,t}$ reads:

$$(1-\delta)(1-\psi)\hat{c}_{i,t}^{-\psi} = \delta \frac{\partial \tilde{V}_{i,t}(\hat{s}_{i,t}, \xi_{i,t}, z_{i,t})}{\partial \hat{s}_{i,t}} \iff$$

$$(1-\delta)\hat{c}_{i,t}^{-\psi} = \delta \tilde{V}_{i,t}(\hat{s}_{i,t}, \xi_{i,t}, z_{i,t})^{\frac{\gamma-\psi}{1-\psi}} \times$$

$$\mathbb{E}_t \left[e^{-\gamma(\Delta w_{t+1}+\Delta f_{t+1})} \left(R_f + \xi_{i,t} R_{t+1}^e\right) \left(p_t \hat{V}_{i,t+1}^{-\gamma}(\hat{a}_{i,t+1}, z_{i,t+1}) \frac{\partial \hat{V}_{i,t+1}(\tilde{a}_{i,t+1}, z_{i,t+1})}{\partial \hat{a}_{i,t+1}} + b(1-p_t)\hat{a}_{i,t+1}^{-\gamma}\right)\right]$$

[37] That is, I also rescale the problem by $\bar{f}(t)$.

Finally, the envelope condition is:

$$\frac{\partial \hat{V}_{i,t}(\hat{a}_{i,t}, z_{i,t})}{\partial \hat{a}_{i,t}} = \frac{\hat{V}_{i,t}(\hat{a}_{i,t}, z_{i,t})^{\psi}}{1-\psi} \times$$
$$\left[(1-\delta)(1-\psi)\hat{c}_{i,t}^{-\psi}\frac{d\hat{c}_{i,t}}{d\hat{a}_{i,t}} + \delta\left(\frac{\partial \tilde{V}_{i,t}(\hat{s}_{i,t}, \xi_{i,t}, z_{i,t})}{\partial \hat{s}_{i,t}}\frac{d\hat{s}_{i,t}}{d\hat{a}_{i,t}} + \frac{\partial \tilde{V}_{i,t}(\hat{s}_{i,t}, \xi_{i,t}, z_{i,t})}{\partial \xi_{i,t}}\frac{d\xi_{i,t}}{d\hat{a}_{i,t}}\right)\right]$$

Using the above first order conditions and the fact that $\frac{d\hat{c}_{i,t}}{d\hat{a}_{i,t}} + \frac{d\hat{s}_{i,t}}{d\hat{a}_{i,t}} = 1$, the envelope condition reduces to:

$$\frac{\partial \hat{V}_{i,t}(\hat{a}_{i,t}, z_{i,t})}{\partial \hat{a}_{i,t}} = \hat{V}_{i,t}(\hat{a}_{i,t}, z_{i,t})^{\psi}(1-\delta)\hat{c}_{i,t}^{-\psi}$$

Replacing the envelope condition in the FOCs above I obtain:

$$\mathbb{E}_t\left[e^{-\gamma(\Delta w_{t+1}+\Delta f_{t+1})}R^e_{t+1}\left(p_t(1-\delta)\hat{V}^{\psi-\gamma}_{i,t+1}(\hat{a}_{i,t+1}, z_{i,t+1})\hat{c}^{-\psi}_{i,t+1} + b(1-p_t)\hat{a}^{-\gamma}_{i,t+1}\right)\right] = 0 \qquad \text{(C.1)}$$

$$(1-\delta)\hat{c}_{i,t}^{-\psi} = \delta\tilde{V}_{i,t}(\hat{s}_{i,t}, \xi_{i,t}, z_{i,t})^{\frac{\gamma-\psi}{1-\psi}} \times$$
$$\mathbb{E}_t\left[e^{-\gamma(\Delta w_{t+1}+\Delta f_{t+1})}\left(R_f + \xi_{i,t}R^e_{t+1}\right)\left(p_t(1-\delta)\hat{V}^{\psi-\gamma}_{i,t+1}(\hat{a}_{i,t+1}, z_{i,t+1})\hat{c}^{-\psi}_{i,t+1} + b(1-p_t)\hat{a}^{-\gamma}_{i,t+1}\right)\right] \qquad \text{(C.2)}$$

C.2.1 Special cases

Retirement. The problem's solution remains exactly the same except for the fact that $z_{i,t}, \alpha_i$ and β_i are not state variables anymore and that $\bar{L}_{i,t} = \bar{L}_{i,K}$, which is constant and with respect to, therefore, we do not need to compute expectations. Another difference is the retirement replacement ratio for labor income starting from $K+1$. Furthermore, because during retirement wages are not indexed anymore, it is not possible to scale the problem by the average wage as before. I overcome this issue by following Catherine (2021) and assuming that the average wage index remains constant after retirement.

Last period of life. Recall that in the last period of life T it holds $p_T = 0$ so that the objective function becomes:

$$\hat{V}_{i,T}(\hat{a}_{i,T}, y_{i,K}) = \max_{\xi_{i,T}, \hat{c}_{i,T}, \hat{s}_{i,T}}\left\{(1-\delta)\hat{c}^{1-\psi}_{i,T} + \delta\left[\mathbb{E}_T\left(b\hat{a}^{1-\gamma}_{i,T+1}e^{(\Delta w_{T+1}+\Delta f_{T+1})(1-\gamma)}\right)\right]^{\frac{1-\psi}{1-\gamma}}\right\}^{\frac{1}{1-\psi}}$$

subject to the same constraints as before. For the solution, it is again useful to define:

$$\tilde{V}_{i,T}(\hat{s}_{i,T}, \xi_{i,T}, y_{i,K}) = \left[\mathbb{E}_T\left(b\hat{a}^{1-\gamma}_{i,T+1}e^{(\Delta w_{T+1}+\Delta f_{T+1})(1-\gamma)}\right)\right]^{\frac{1-\psi}{1-\gamma}}$$

Proceeding as in the previous section, we get the following first order conditions:

$$\mathbb{E}_T\left[e^{-\gamma(\Delta w_{t+1}+\Delta f_{T+1})}R^e_{T+1}\hat{a}^{-\gamma}_{i,T+1}\right] = 0$$

$$(1-\delta)\hat{c}_{i,T}^{-\psi} = \delta\tilde{V}_{i,T}(\hat{s}_{i,T}, \xi_{i,T}, y_{i,K})^{\frac{\gamma-\psi}{1-\psi}}\mathbb{E}_T\left[e^{-\gamma(\Delta w_{T+1}+\Delta f_{T+1})}\left(R_f + \xi_{i,T}R^e_{T+1}\right)b\hat{a}^{-\gamma}_{i,T+1}\right]$$

In the above equations we have assumed $b \neq 0$. If $b = 0$, the problem becomes very simple:

$$\hat{V}_{i,T}(\hat{a}_{i,T}, y_{i,K}) = \max_{\xi_{i,T}, \hat{c}_{i,T}, \hat{s}_{i,T}}(1-\delta)^{\frac{1}{1-\psi}}\hat{c}_{i,T}$$

with the same constraints as before. The trivial optimal policies are thus $\hat{c}_{i,T} = \hat{a}_{i,T}$ and $\hat{s}_{i,T} = 0$.

C.2.2 Solution algorithm

I will outline the solution algorithm for the most general case, special cases can be included as straightforward extensions. The model is solved by backward induction and the endogenous grid point method with the following procedure:

1. Use the terminal condition (2) to solve for the the value function and optimal policies at $T + 1$;
2. For each $t \in [K + 1, T]$ solve for the optimal policies and value function as follows:
   - For each point in the grid for $\bar{L}$ and for each for each point in the grid for $\hat{s}$ compute:
     (a) Optimal risky share $\xi_{i,t}$ in the case of participation and of non-participation. Recall that equation (C.1) solves $\frac{\partial \tilde{V}}{\partial \xi} = 0$. If $\frac{\partial \tilde{V}}{\partial \xi} > 0$ if $\xi = 1$ then set the optimal risky share to 1 while if $\frac{\partial \tilde{V}}{\partial \xi} < 0$ if $\xi = 0$ then set the optimal risky share to 0. Otherwise, set the optimal risky share to the value that solves (C.1). In the case of non-participation the optimal risky share is trivially zero;
     (b) Optimal consumption $\hat{c}_{i,t}$ by solving equation (C.2);
     (c) Cash on hand at the beginning of the period from the normalized constraint $\hat{c}_{i,t} + \hat{s}_{i,t} + F_{i,t}\kappa_f / e^{\bar{f}(t)} = \hat{a}_{i,t}$;
     (d) The value function by inserting the optimal policies just computed in the expression of the value function $\hat{V}_{i,t}$;
     (e) Using the minimum value of cash on hand implied by the model find if the borrowing constraint binds. In the affirmative case add a point corresponding to this case;
     (f) Linearly interpolate the value function and the optimal policies on the grid for cash on hand at the beginning of the period. Note that this requires finding the switching point between participation and non-participation on the cash on hand grid point and using the solution values for non-participation below that point and for participation above that point. using the optimal quantities for the case
3. For each $t \in [T_{\text{start}}, K]$ solve for the optimal policies and value function as follows:
   - For each point in the grid for β, for each point in the grid for α, for each point in the grid for z, for each point in the grid for $\bar{L}$ and for each for each point in the grid for $\hat{s}$: repeat the same steps (a)-(e) in the list above.

C.3 Interpolation

The solution procedure outlined in section C.2 will very often require to evaluate the value function and the consumption policy at points off the grid. This also applies to model simulation when evaluating the solved policies at the points of the simulated paths. As explained in section C.1, I do not discretize the persistent component of idiosyncratic income, which implies that I need to interpolate these functions not only at points off the cash on hand grid, the life-cycle parameters α, β grids, the average idiosyncratic income $\bar{L}$ grid, but also off the grid of persistent income. In other words, I need a multidimensional interpolation procedure over the $(\alpha, \beta, a, z, \bar{L})$ grid. This is achieved by multidimensional linear interpolation.

C.4 Computing expectations

In order to solve the model, it is necessary to compute expectations of some non trivial functions. In the most general case, I need to compute expectations with respect to the shocks r_1, r_2, ϕ, ε and ν.[38] To do that, I proceed as follows: (i) for all the possible combinations of grid values of these variables, I compute the value of the function (ii) I multiply it by the probability of that particular combination of values (iii) once I have done this for all the possible combinations I sum up all the function values obtained. Note that the grid values and probabilities of the other shocks coincide with Gaussian quadrature nodes and weights[39], which enables me to compute expectations very accurately. Finally, remember that the distributions of r_1 is conditional on the realization or not of a stock market crash and, similarly, those of ε and ν on the realization of a tail event or not. This is taken into account simply by scaling the probability of the discretized conditional distributions of these variables by the probability of these events.

D Estimation

This section describes how I estimate the exogenous stochastic processes. For both aggregate variables and individual income process I follow the procedure outlined in Catherine (2021). I will now describe the part of the procedure that is common for both and then dedicate two specific paragraph for the peculiarities regarding each of the two. The numerical parameters chosen for the estimation procedure are reported in Table C.1.

Let θ be the vector of parameters that has to be estimated. θ is chosen to minimize the

[38] Cases in which I do not need to take expectations with respect to one or more of these variables can be handled by the same procedure outlined here with straightforward modifications.

[39] See section C.1.

following objective function:

$$\min_{\theta} \hat{m}(\theta)' W \hat{m}(\theta) \tag{D.1}$$

where $\hat{m}(\theta)$ is a vector of moments that depends on the parameters to be estimated and W is a weighting matrix. The procedure involves a global and a local stage. In the global stage I compute the value of the objective function for N_{glo} combination of points for the elements of the vector θ. The combinations correspond to the first N_{glo} of a Sobol sequence. At the end of the global stage, the best - in the sense of providing the lowest values of the objective function - N_{loc} points pass to the local stage. In the local stage, for each of the N_{loc} points, equation (D.1) is solved for the minimum using the Nelder-Mead algorithm with starting guess each of such points. The minimum is then the vector of parameters among the N_{loc} local points that returns the lowest value of the objective function.

Aggregate processes. To estimate the stochastic process governing the aggregate variables in the model I simulate the process for T_{cal} periods and then I compute the difference between the model generated moments and the moments in the data. I discard the first T_{dis} points from moments computation and, in order to smooth the surface of the objective function, I simulate the process for N_{eco} economies and average moments across them. Letting m indicate a generic moment, $\hat{m}(\theta)$ is defined as follows:

$$\hat{m}(\theta) = \frac{m_{\text{data}} - m_{\text{simulated}}(\theta)}{m_{\text{data}}} \tag{D.2}$$

The weighting matrix W is a unitary diagonal matrix. The actual moments I target are described in the main text.

Individual income process. I closely follow Guvenen et al. (2014) and Catherine (2021) to estimate the stochastic process governing individual income. Specifically, I simulate the income histories of 68 cohorts, the first starting in 1944 and the last in 2011 assuming that the persistent component is zero at the beginning and that the model economy is subject to the same aggregate wage shocks as in the data. Each cohort is made of N_{cal} individuals. Note that this implies having a constant age structure between 1979 and 2011.[40] Also in this case, to smooth the surface of the objective function I simulate N_{eco} economies and average moments across them. As described in the main text, the first 155 moments I match are standard deviation and Kelly's skewness of earnings growth at different time horizons. For these moments, the function $\hat{m}_t(\theta)$ is defined as follows:

$$\hat{m}_t(\theta) = \frac{m_{t,\text{data}} - m_{t,\text{simulated}}(\theta)}{\overline{m}_{\text{data}}} \tag{D.3}$$

[40]The moments provided by Guvenen et al. (2014) refer to individuals between 25 and 60.

where $\overline{m}_{\text{data}}$ is the time-series average of the absolute value of the moment under scrutiny. For the 36 within-cohort variances, instead, I use the same formula as in equation (D.2). The weighting matrix W is a unitary diagonal matrix that assigns equal weights to all moments.