**Working Paper Series** 751
(ISSN 2788-0443)

# Extrapolative Income Expectations and Retirement Savings

**Marta Cota**

CERGE-EI
Prague, April 2023

**ISBN 978-80-7343-558-5 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)**
**ISBN 978-80-7344-675-8 (Národohospodářský ústav AV ČR, v. v. i.)**

# Extrapolative Income Expectations and Retirement Savings

Marta Cota*

CERGE-EI†

**Abstract**

Why do employees' retirement contributions gradually increase throughout their careers? This paper uses a structural life-cycle model based on household expectations data to explain workers' retirement contribution decisions. The Michigan Survey of Consumers data shows that young households extrapolate from their recent income realizations and overstate the persistence and volatility of their future income. The structural life-cycle model with extrapolative expectations quantifies the difference in retirement contribution rates compared to rational expectations. Contrary to rational workers, extrapolative workers' contributions match the data on retirement contributions over the life cycle. Consequently, mandating automatic enrollment yields negligible effects on retirement savings.

***Keywords:*** extrapolative expectations, forecast errors, illiquid savings, retirement contribution
J.E.L. classification: E21, J26, J32

*I thank the participants of seminars at the Copenhagen Business School departments of Finance and Economics for their useful comments, especially Kathrin Schlafmann for advice and guidance during my research stay. I thank John Beshears and Taha Choukhmane for their constructive comments on the paper. I thank the participants at the 4th Behavioral Macro Workshop and the 1st PhD Workshop in Expectations for their comments. I thank Jeremy Lise, Fatih Guvenen, and other participants of their workshops at the University of Minnesota for their comments, especially Mariacristina De Nardi's continuous support. I am grateful to my advisor, Marek Kapicka, Ctirad Slavik and CERGE-EI faculty for their comments at CERGE-EI workshops and seminars.*

*†CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences Sciences, Politických vězňů 7, Prague, Czech Republic, e-mail: marta.cota@cerge-ei.cz.*
*This output was supported by the NPO 'Systemic Risk Institute' number LX22NPO5101, funded by European Union - Next Generation EU (Ministry of Education, Youth, and Sports, NPO: EXCELES.)*

# 1 Introduction

Income expectation biases arising from pessimism or optimism may affect the extent of saving, resulting in lower savings rates than rates predicted by the standard rational model of wealth accumulation. This paper investigates the effects of income expectation bias on retirement savings rates over workers' careers. I build a structural life-cycle model with income forecast biases that generates patterns of retirement savings plan contributions during a worker's career.

The structural life-cycle model relies on three findings in the expectation data of the Michigan Survey of Consumers. First, income forecast biases change sign across the income distribution, moving from pessimistic low-income to optimistic high-income households. I extend the Michigan Survey of Consumers' data analysis in Schlafmann and Rozsypal (2023) and show that households extrapolate, basing their income predictions on previous income realizations. Next, I show that as households age, their income forecast errors decrease. Lastly, I show that households overstate the probability of losing a job, regardless of age and education level. While all households overstate persistent income volatility, their income growth expectations differ, separating pessimists and optimists.

Using the retirement contribution data in the Survey of Consumer Finances, I show that the liquid-to-retirement savings ratio persists over the life cycle for low- income workers. In contrast, younger high-income workers keep more than half of their savings in liquid accounts. With age, high-income workers reallocate their savings to illiquid savings accounts, including retirement savings. These findings are consistent over time. In line with these findings, Parker et al. (2022) find steadily increasing contribution rates over the life-cycle, across all cohorts. Using model simulations, I show that rational expectations do not match these contribution patterns. In contrast, the extrapolative expectations model, as households learn to correct their biases, they gradually increase their contributions. Moreover, in line with persistent contribution rates across cohorts (Parker et al., 2022), I show that retirement plan reforms do not affect contribution rates with extrapolative households.

I motivate the extrapolation mechanism analytically in a stylized three-period model. Pessimistic workers contribute less to their illiquid retirement accounts, and allocate their savings to liquid accounts that they can tap into. The mechanism works through the fear of being borrowing-constrained by persistently low income in the near future. On the other hand, high-income workers postpone their retirement contributions due to optimistic income expectations.

I reinstate the mechanism and develop a structural life cycle model with extrapolative expectations and two types of assets, built on Schlafmann and Rozsypal (2023). Workers face persistent and transitory income shocks and choose when to start and how much to contribute to their private retirement plans, and how much to keep in liquid savings accounts. Retirement plans are illiquid and imitate private retirement accounts (such as the 401(k)) in the U.S., and therefore include employer's match and tax deferrals. As this paper focuses on the intensive margin, all workers in the model are eligible to participate in the retirement plan.

I calibrate the income-forecast misperception to match survey responses from the Michigan Survey of Consumers using the Method of Simulated Moments. The trade-off between saving for the near future and saving for retirement differs from the rational benchmark case. Pessimistic workers are not willing to forego their liquid buffers and hence save less in retirement savings, regardless of forms of their incentives.[1] On the other hand, optimistic workers postpone their contributions relative to rational counterparts. However, over the work life, both types of workers catch up with retirement savings by increasing contribution rates towards retirement, reflecting firm-level data findings (Parker et al., 2022; Choukhmane, 2021).

In contrast to the rational expectations benchmark that predicts decreasing liquid saving rates across wealth percentiles, my model solution shows that liquid savings from income remain flat after the 20th wealth percentile, consistent with empirical findings (Fagereng et al., 2019; Nardi and Fella, 2017). In addition, differences in retirement savings rates between wealth quantiles are larger and thus highlight the importance of unrealized capital gains in illiquid retirement accounts. Wealthier workers contribute more to retirement plans throughout the work lifespan and when capital gains are finally realized, once workers retire, they consume at significantly higher levels.

After aligning contribution patterns with the extrapolative expectations solution, I test the implications of the 401(k) automatic enrollment policy, recommended by the U.S. government in the employer guide[2]. All employees are enrolled in a plan with automatically, and employers match their employees' contributions up to a certain level (typically, up to 3%). Model simu-

[1] I included tax deferrals and benefit functions of different kinds.

[2] Published Guidance from the IRS can be found at `https://www.irs.gov/retirement-plans/published-guidance`.

lations show that workers contribute less when approaching retirement and offset their higher contribution rates at the start of their working lives. As a result, adjustments to retirement account enrollment do not result in substantial welfare gains.

Voluntary participation savings paths with extrapolative expectations serve as a baseline for worker behavior and show that automatic participation policy correlates with findings in event study estimates (Choukhmane, 2021; Goda et al., 2020). Initially, contribution rates increase, only to remain low to offset the distortion on impact. Consequently, my structural model offers insight regarding less distorting policies, such as auto-escalating contribution rates in retirement plans.

## 2 Related literature

Studies by Grevenbrock et al. (2021) and De Nardi et al. (2009) use life expectancy biases to motivate retirement saving decisions. In both of these studies, the consumption paths of assets during retirement are closer to the data than in standard rational expectation models. However, saving for retirement through retirement contribution accounts adheres to the life-cycle income path and other savings decisions from the start of the work life. Duarte et al. (2021) build a rational expectations model with retirement saving portfolio allocation and find that almost all young workers add to their retirement accounts and invest their savings primarily in equity[3]. Contrary to default options, rational workers opt for equity funds and face significant losses over the life cycle.

In contrast to the rational expectations portfolio choice solution, empirical studies that use firm-level data find dominantly low contribution rates and default fund choices (Blanchett et al., 2021; Parker et al., 2022). Contributions are low a few years after the tenure begins, (Choi et al., 2003; Choukhmane, 2021; Devlin-Foltz et al., 2015). As a result, modeling retirement plan contributions includes money (opt-out) costs as a tool for saving for retirement to match the retirement savings patterns in the data (Choukhmane, 2021; Dahlquist et al., 2018; Love, 2006). Moment targeting produces opt-out that can be immensely high (DellaVigna, 2018).

This paper adds to retirement contribution studies using a behavioral assumption to rec-

[3]Once the worker chooses how much to contribute, she can opt for a type of fund to invest in: equity, bonds or a mixed target-date (TDF) fund. TDF is usually set as a default option in retirement plans.

oncile the dynamics of contribution patterns with firm-level data. The data analysis builds on that in Schlafmann and Rozsypal (2023) and sets the ground for the structural life-cycle model. Schlafmann and Rozsypal (2023) focus on future income expectations and find that people overestimate their future income growth persistence. I find evidence suggestive of extrapolation, adding to evidence from panel data in the U.K. (Cocco et al., 2022) and the Netherlands (Massenot and Pettinicchi, 2019). Moreover, data estimates on subjective unemployment probabilities overstate the actual probabilities across the whole sample population, adding to increased precaution documented in the data Blundell et al. (2008). Correspondingly, an increase in idiosyncratic risk shifts beliefs towards pessimism, capturing the mechanism outlined in the theoretical model of expectation bias in Bhandari et al. (2016).

The model in this paper connects expectation biases to workers' retirement plan contributions in the U.S. Recent data findings in Ghilarducci et al. (2018) show that workers extrapolate from their recent past and adjust their savings to ensure living standards. Moreover, Goda et al. (2020) show that contribution behavior varies significantly with financial literacy, while behavioral biases such as present bias and exponential growth bias remain insignificant. Similarly, I highlight the effect of understanding one's income on yearly contribution rates throughout their career.

Aligning early retirement contribution behavior often necessitates behavioral assumptions implying passive behavior, i.e., adding at default rates [4] (Bernheim et al., 2015; Ameriks et al., 2007; Benartzi and Thaler, 2007). In a dynamic setting, studies incorporating time-preference biases cannot reconcile the retirement contributions found in the data and often turn to contribution adjustment costs (Choukhmane, 2021; Dahlquist et al., 2018). However, cost estimates are large and amount to thousands of dollars each year, discussed Choukhmane (2021) and DellaVigna (2018). The expectation bias in this paper implies staggered contributions early in the work life across the income distribution, which is explained by extrapolative expectations driving the delay.

Separating extrapolation from other household characteristics adds to the growing behavioral finance literature. Experimental studies Krijnen et al. (2022) and Goda et al. (2020) argue that correcting workers' expectations regarding savings growth may increase contribution rates.

[4] The default rate is either set to $0\%$ or $3\%$, depending on the enrollment regime.

As time inconsistency and other psychological biases require experimental data, this paper contributes to the literature by explaining passive behavior in retirement saving, based on public survey data estimates.

Ultimately, connecting income forecast errors to saving choices steps out of the finance literature that connects portfolio choices and future returns extrapolation (Bordalo et al., 2018, 2019). Adhering to investors' behavior, a new line of research uses extrapolation to account for heterogeneity in the marginal propensity to consume, making income expectation biases relevant for fiscal stimulus (Auclert et al., 2020; Choi and Mertens, 2019). In line with these studies, the model in this paper highlights the difference in the liquid-illiquid saving trade-off between extrapolative workers and their rational counterparts.

Sticking to liquid savings tools due to pessimism based on past low-income realizations translates to a negligible reaction to retirement plan adjustments for the bottom part of the income distribution. This explains persistent contribution rates across cohorts Parker et al. (2022) and ambiguous effects of retirement plan reforms (Choukhmane, 2021; Beshears et al., 2022; Bernheim et al., 2015). Similarly, the policy exercise with extrapolative workers finds insignificant effects of automatic enrollment on average retirement savings.

# 3 Data

## 3.1 Bias in the income growth forecast

The structural life-cycle model hinges on income expectation patterns in the Michigan Survey of Consumers data. My data analysis adds to previous findings in Schlafmann and Rozsypal (2023) and Das and Van Soest (1999). I incorporate multiple survey questions to establish useful facts relating to future income misperception and retirement confidence across the survey sample. Income mean and volatility forecast errors across household characteristics exhibit patterns that align with the subjective income process assumption.

Turning to income mean forecast errors, linear regression estimates highlight the importance of a worker's position in the income distribution, which consequently implies whether the consumer is pessimistic or optimistic. Moreover, logistic regression estimates serve as suggestive evidence of extrapolation, in line with other data analyses (Massenot and Pettinicchi, 2019;

dH́aultfoeuille et al., 2021). Using subjective unemployment probabilities as a proxy for perceived income volatility I argue that workers overstate job separation rates across all education levels and age groups.

### 3.2 Constructing expectation errors at the household level

Since households are re-interviewed (once) in the MSC survey, I estimate the bias in income growth expectations, using the questions

1. *"During the next 12 months, do you expect your (family) income to be higher or lower than during the past year? "*
2. *"By about what percent do you expect your (family) income to (increase/decrease) during the next 12 months?"*

and then compare the answers to realized income responses in the next survey wave. The advantage of the MSC survey is that it asks respondents to specify their income growth forecast percentage from 1986-2012. In the second interview, households are asked to report their last-period income, which may be subject to individual measurement error[5]. Since I cannot attain the objective last-period income value, I use household characteristics to compare reported income with official income statistics to check for robustness.

After denominating income values to 2010 dollars, I evaluate expectation errors using short panels, by tracking the household during their re-interview

$$\phi_{i,t} = \frac{\Delta \hat{Y}_{i,t+1}}{Y_{i,t}} - \frac{\Delta Y_{i,t+1}}{Y_{i,t}} = \hat{g}_{i,t+1|t} - g_{i,t+1},$$

where $g_{i,t+1}$ is the self-reported real income growth for the previous period (previous year) and $\hat{g}_{i,t+1|t}$ is the expected income growth. The sign of the error implies pessimism or optimism. If the error is positive, the worker expected a higher income than was realized. The negative error thus implies a pessimistic future income outlook.

The analysis does not include household incomes lower than the estimated unemployment benefits aggregated yearly. Moreover, the sample includes households with no change in socioe-

[5]Schlafmann and Rozsypal (2023) give a detailed explanation regarding the survey data and a comparison to other surveys.

conomic characteristics, such as family structure and education. Ideally, households would be responding to survey questions for two consecutive years. However, restricting the sample to respondents re-interviewed during the subsequent year does not change the regression estimates. Overall, there are 47,000 re-interviewed households, from which 30,000 respondents gave their first response in June. Specific questions regarding job uncertainty and retirement confidence came later in the MSC survey; hence, sample sizes vary from 20,000 to 37,000, depending on the question.

### 3.3 Forecast error distribution estimates

Figure 1 shows the difference between forecast error distributions for each income quantile. The mean of the error distribution shifts from left to right, implying a shift from negative to positive forecast errors or from pessimism to optimism. That is, on the low end of the income distribution workers expect their income to be lower than it actually is. In contrast, high-income workers are more optimistic. In addition, a distribution tail comparison shows a larger mass of pessimists in the first and second income quantiles. Therefore, error distributions shift gradually.

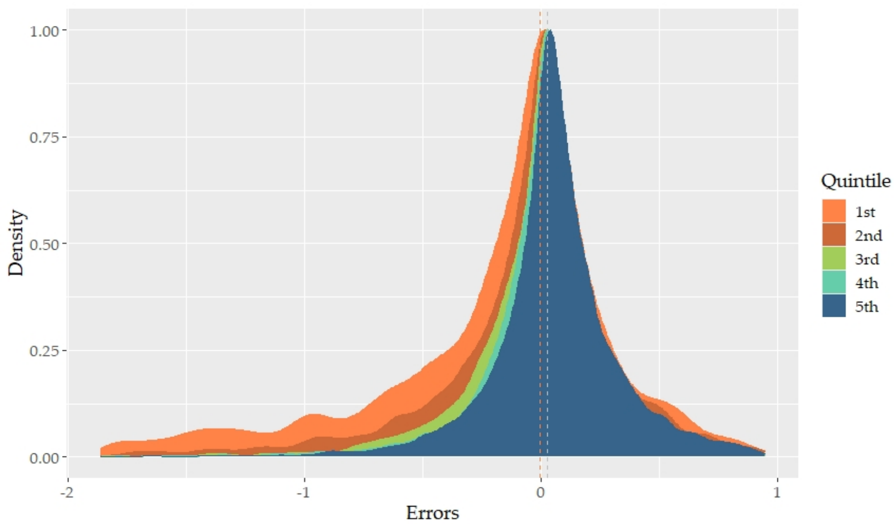


Figure 1: Income growth bias density by income quantile, MSC data

### 3.4 Linear regression results

While non-parametric estimates show differences in the error distribution across income quantiles, linear model estimates control for other household characteristics. The income quantile remains a significant predictor for forecast error while controlling for household characteristics. Estimates include month and year effects and limit the education variable to only three possibilities, clearly distinguishing between high school graduates, college graduates, and those with less education. The coefficient next to the income quantile is significant and large relative to other household characteristics (Table 1). Furthermore, predicted errors at each quantile show a decreasing influence of income as individuals move up the income distribution (Figure 2).

The majority of regressors are indicator variables, whereas $age$ and $age^2$ are rescaled, following Gelman (2008). This way the regression model does not lose interpretability and the standard errors are not downward biased [6]. As a result, age becomes a significant predictor of the income growth forecast error [7]. Therefore, the structural model incorporates lower errors in later work life, as a result of deterministic income shape with respect to work experience.

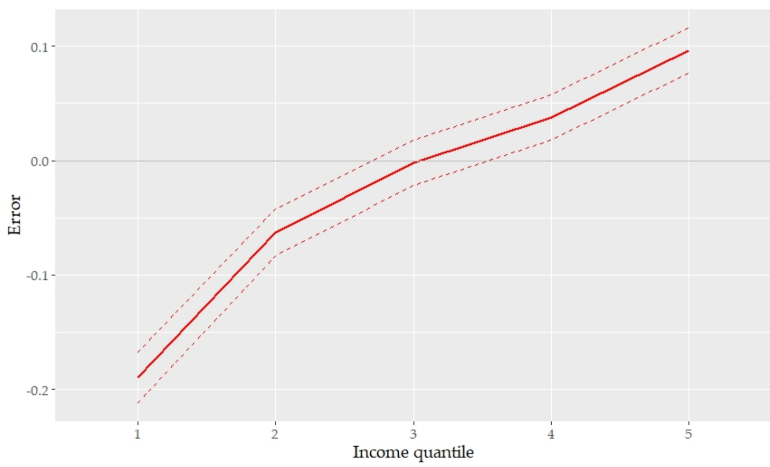


Figure 2: Income growth errors change sign when moving from lower to higher quintiles, MSC data.

[6]Reported standard errors are heteroskedasticity-robust.

[7]Furthermore, estimates of kernel density for separate age bins show that the forecast error decreases in mean and variance. Estimates are given in the appendix.

Table 1: Linear Regression Results

| | *Dependent variable:* |
|---|---|
| | Income Growth Forecast Errors |
| $q_2$ | 0.206*** |
| | (0.008) |
| $q_3$ | 0.286*** |
| | (0.008) |
| $q_4$ | 0.327*** |
| | (0.009) |
| $q_5$ | 0.393*** |
| | (0.014) |
| $male$ | −0.012* |
| | (0.006) |
| no HS | 0.039*** |
| | (0.006) |
| college | −0.046*** |
| | (0.003) |
| $age$ | −0.156*** |
| | (0.026) |
| $age^2$ | 0.152*** |
| | (0.034) |
| $1\ adult$ | 0.096*** |
| | (0.005) |
| $> 2\ adults$ | −0.035*** |
| | (0.008) |
| Constant | −0.303*** |
| | (0.012) |
| Observations | 47,341 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

## 3.5 Probability estimations - extrapolation and overstating income volatility

While forecast error signs are dominantly explained by household income, the fact that households extrapolate from the near past is not obvious. Current empirical studies find evidence of extrapolation in households expectations surveys (Massenot and Pettinicchi, 2019; Ghilarducci et al., 2018). In line with empirical findings, estimates from the MSC data show that households extrapolate based on their recent income realization (i.e, the income growth error), regardless of their income level.

During the survey, households are asked to assign probabilities to the rise in personal income during the next year and evaluate the likelihood of their five-year job retention:

- *What do you think is the percent chance that your income in the next twelve months will be higher than your income in the past twelve months?*

Using the ordered logistic model, I show that the most recent forecast error explains the worker's outlook on future income growth. In addition, high-income households are, on average, twice as optimistic about future income growth. The estimates suggest that individual income growth expectations depend on previous income realizations, owing to extrapolation from the nearest past (forecast error in Table 2), while still depending on the whole income history (income quantile in Table 2).

Table 2: Ordered logistic regression results; households extrapolate from their recent income realization, MSC data.

| | *Dependent variable:* |
|---|---|
| | P(income increase in t+1—t) |
| sex | $0.338^{***}$ |
| | (0.031) |
| no HS | $-0.513^{***}$ |
| | (0.089) |
| College | $0.423^{***}$ |
| | (0.034) |
| $err_{t-1}$ | $0.455^{***}$ |
| | (0.053) |
| age | $-0.804^{***}$ |
| | (0.033) |
| 1 adult | $0.138^{***}$ |
| | (0.040) |
| $>2$ adults | 0.019 |
| | (0.050) |
| $q_2$ | 0.094 |
| | (0.066) |
| $q_3$ | $0.377^{***}$ |
| | (0.065) |
| $q_4$ | $0.399^{***}$ |
| | (0.066) |
| $q_5$ | $0.590^{***}$ |
| | (0.069) |
| Observations | 14,710 |
| *Note: Controlled for year-effects.* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

Next, I use the workers' job loss predictions elicited in the MSC responses. Comparison of subjective job loss probabilities to empirical estimates in labor studies show that workers overstate the probability of losing a job, regardless of age or education levels.

The survey question

- *During the next 5 years, what do you think the chances are that you (or your husband/wife) will lose a job you wanted to keep?*

allows for estimation of subjective job loss predictions over the next five years. These estimates define a lower bound for subjective job loss probabilities one year ahead if one assumes a constant year-ahead unemployment outlook. Comparing empirical estimates from Farber et al. (2005) to the MSC data estimates implies that households overestimate their job loss probability (Tables 3)[8]. In the model, income process misperception includes overstating income volatility throughout the income distribution[9]. Before going through the retirement contribution data

| $\hat{\mathbb{P}}$ | $ed < 12$ | $12 \leq ed \leq 15$ | $ed > 15$ |
|---|---|---|---|
| age $25-34$ | 0.063 | 0.060 | 0.053 |
| age $35-44$ | 0.070 | 0.054 | 0.051 |
| age $45-54$ | 0.056 | 0.052 | 0.053 |
| age $55-66$ | 0.017 | 0.039 | 0.034 |

| $ed < 12$ | $12 \leq ed \leq 15$ | $ed > 15$ |
|---|---|---|
| 0.068 | 0.052 | 0.035 |
| 0.058 | 0.043 | 0.030 |
| 0.053 | 0.039 | 0.028 |
| 0.057 | 0.039 | 0.027 |

Table 3: Left: subjective job loss probabilities from the MSC data, right: empirical estimates (Farber et al., 2005; Love, 2006).

findings, I outline expectation assumptions that generate the error patterns in the MSC data. I relate the pattern in income growth forecast errors to the model similarly to Schlafmann and Rozsypal (2023), by assuming the misperception in the auto-regression coefficient $\lambda$.

[8] The estimates do not include NBER recession years.

[9] Of course, the income quantile does predict the probability stated in the survey. However, the probability is still higher than the true one.

### 3.6 Relating forecast errors to the model

The extrapolative expectations model hinges on expectation patterns in the data. The incorrectly perceived income process separates between low-income pessimists and high-income optimists. Throughout their career, agents can change their outlook based on their income history. That is, the subjective life-cycle income process incorporates three facts:

1. income level expectations transition from pessimistic on the left part to optimistic on the right part of the income distribution
2. persistent income volatility is overstated across all workers
3. workers extrapolate from their recent income realizations.

The structural model assumes that the true income process satisfies

$$Y_{it} = A_i G(t) \mathit{\Gamma}_{it} P_{it}, \quad \log A_i \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2), \;\; \mathit{\Gamma}_{it} \sim \log \mathcal{N}\left( -\frac{\sigma_{\mathit{\Gamma}}^2}{2}, \sigma_{\mathit{\Gamma}}^2 \right),$$

where $\log P_{it}$ follows an AR(1) process

$$\log P_{it} = \lambda \log P_{it} + \xi_{it}, \quad \xi_{it} \sim \mathcal{N}(\mu_\xi, \sigma_\xi^2).$$

Given the true income process, assume that agents mispercieve the persistence of their income, regardless of their age or individual effects. This implies

$$\hat{Y}_{it} = A_i G(t) \mathit{\Gamma}_{it} P_{it}^{\hat{\lambda}} \implies \hat{\mathbb{E}}_t Y_{i,t+1} = A_i G(t) P_{i,t}^{\hat{\lambda}}, \tag{1}$$

whereas rational agents know the true income process, so

$$\mathbb{E}_t Y_{i,t+1} = A_i G(t) P_{i,t}^{\lambda}. \tag{2}$$

$\hat{\lambda}$ implies the perceived log-income process

$$y_{it} = \alpha_i + g(t) + p_{it} + \gamma_{it} \text{ and } p_{it} = \hat{\lambda} p_{i,t-1} + \xi_{it}, \xi_{it} \sim \mathcal{N}\left(\mu_\xi, \sigma_\xi^2\right), \;\; \gamma_{it} \sim \mathcal{N}\left(\mu_\gamma, \sigma_\gamma^2\right).$$

The differences in expected and realized income are

$$\begin{aligned}\mathbb{E}^*_t[y_{i,t+T}] - \mathbb{E}[y_{i,t+T}] &= \mathbb{E}^*[p_{i,t+T}] - \mathbb{E}[p_{i,t+T}] \\ &= (\hat{\lambda}^T - \lambda^T)(p_{i,t} - \mu_\xi), \forall T\end{aligned}$$

and change depending on

$$p_{i,t} \lesseqgtr \mu_\xi.$$

For a large enough realization of persistent income, the subjective future income is higher than the rational one, i.e., the agent is an optimist. In contrast, if persistent income is sufficiently low, the agent becomes pessimistic and expects lower future income. Given that the persistent component is a sum of all previous income realizations, the worker may change their outlook over their career.

Following the MSC data evidence on unemployment probability pessimism among all workers, the structural model includes the persistent component volatility. The data shows that all workers overstate their persistent income volatility, regardless of age or other characteristics. In model terms, using the AR(1) persistent income process

$$p_{it} = \hat{\lambda} p_{i,t-1} + \xi_{i,t}, \forall t = 1, \dots, T_{ret} - 1,$$

where $\hat{\lambda} > \lambda$ implies the income growth forecast errors in expression (1), conditional volatility satisfies, $\forall t = 1, \dots, T_{\text{ret}}$

$$\mathbb{V}_t\left[p_{t+T}\right] = \sigma_\xi^2 \frac{1-\lambda^{2T}}{1-\lambda} < \hat{\mathbb{V}}_t\left[p_{t+T}\right] = \sigma_\xi^2 \frac{1-\hat{\lambda}^{2T}}{1-\hat{\lambda}},$$

regardless of the previous realization of $p_t$.

For $T = 1$ and $\hat{\lambda} = \lambda + \varepsilon < 1$

$$\begin{aligned}\hat{\mathbb{E}}_t\left[p_{i,t+1}\right] - \mathbb{E}_t\left[p_{i,t+1}\right] &= \hat{\lambda} p_{i,t} - \lambda p_{i,t} \\ &= \varepsilon \left( \sum_{s=0}^{t-1} \hat{\lambda}^{t-s} \xi_{i,s} + \hat{\lambda}^t p_{i,0} \right),\end{aligned} \tag{3}$$

Thus, persistent income realizations determine the sign of the difference, separating pessimists from optimists, and corresponding to the infinite horizon outlook in Schlafmann and Rozsypal (2023). Differences in signs align with empirical separation of optimists and pessimists based on the income distribution position and the effect of recent income realization on income outlook. Over the life cycle, the deterministic component outweighs the persistent one when considering the income level. As a result, the model forecast error accounts for heterogeneity over time without imposing ad hoc constraints on the income process.

The two stage calibration of the lifecycle model entails calibrating $\hat{\lambda}$ using the MSC income growth error data. Calibration procedure is explained in detail in the next section. The estimation is robust to income growth error outlier specification and always yields $\hat{\lambda} > \lambda$.

### 3.7 Calibrating the income growth bias ($\hat{\lambda}$)

Since the MSC does not include panel data, I use the stand-in values for the income process parameters in the Panel Study of Income Dynamics (PSID) data[10]. The deterministic income growth parameters follow estimates in Cocco et al. (2005), satisfying the typical hump shape of workers' income over their life cycle. The stochastic part of the income process contains both transitory and persistent components, so I use the estimates in Storesletten et al. (2004). Specifically, the true persistence parameter is set to $0.972$. In this way, true $\lambda = 0.972$ becomes the lower bound for mispercieved $\hat{\lambda}$ [11]

The parametrized income process

$$y_{i,t} = \alpha_i + \text{const.} + g_1 t + g_2 t^2 + g_3 t^3 + p_{i,t} + \gamma_{i,t}, \text{ and } p_{i,t} = \lambda p_{it} + \xi_{i,t} \tag{4}$$

is defined with Each grid element $\hat{\lambda}$ defines the objective function for income persistence bias

Table 4: Income process parameter values.

| $\sigma^2_\alpha$ | const | $g_1$ | $g_2$ | $g_3$ | $\sigma^2_\xi$ | $\sigma^2_\gamma$ | $\lambda$ |
|---|---|---|---|---|---|---|---|
| 0.27 | -2.1700 | 0.1682 | -0.0323 | 0.0020 | 0.0737 | 0.0106 | 0.972 |

calibration. The Method of Simulated Moments minimizes the difference between simulated

[10] Schlafmann and Rozsypal (2023) show that objective income growth rates from the PSID align with patterns in the MSC.

[11] Different estimates in the literature reproduce contribution patterns fairly well. That is, model solution does not depend on the change in the parameter values.

income forecast errors implied by the current $\hat{\lambda}$ and 4, and the empirical income growth forecast errors in the MSC data.

That is, for each income quantile, the perceived persistence parameter minimizes the mean forecast error. Each grid member represents the sample's income error forecast. After taking out age effects, the residuals are used as the dependent variable in a linear regression that makes predictions for income growth forecast errors, at a given income quantile. This way, simulated residuals correspond to the income forecast error in the data and define a loss function.

The calibration includes 50 000 households with separate income processes. The loss function is

$$L(\hat{\lambda}) = \sqrt{\sum_{i=1}^{5} w_i \big(\text{err}(\hat{\lambda})_{q_i} - \text{err}(\lambda)_{q_i}\big)^2}, \tag{5}$$

where $w_i$ is the weight of a given quantile, and is inversely proportional to the life-cycle variance of the each income quantile.

The optimal $\hat{\lambda} = 0.99$ minimizes the loss function at the value $L = 0.0007$. Calibration results do not depend on the outlier criteria for empirical forecast error data. Moreover, results do not depend on the choice of the grid for $\hat{\lambda}$ and yield $\hat{\lambda} > \lambda$. Most importantly, quantile-based forecast error means change sign from low-income (pessimistic) to high-income (optimistic) quantiles (Table 5).

Table 5: Mean income growth forecast error by income quantile for true $\lambda$ and misperceived $\hat{\lambda}$.

| mean error quantile | $q_1$ | $q_2$ | $q_3$ | $q_4$ | $q_5$ |
|---|---|---|---|---|---|
| $\lambda = 0.972$ | -0.1230 | -0.0039 | 0.0506 | 0.0831 | 0.1314 |
| $\hat{\lambda} = 0.99$ | -0.0881 | -0.0106 | 0.0255 | 0.0407 | 0.0326 |

### 3.8 Retirement contributions data

The model relates income expectations to retirement contribution patterns in the data. Using public datasets, I outline two relevant facts of a contribution plan take-up in the U.S. First, the number of low-income workers eligible for a 401(k) plan is persistently increasing (Figure 3, Bureau of Labor Statistics data). Second, at the same time, low income workers tend to keep their savings in liquid accounts even when approaching retirement (Figure 4, the Survey of

Consumer Finances data). Only recently, the U.S. Government instructed employers to allow part-time workers to open up a 401(k) account. Among full-time workers, the number of eligible workers in the BLS has increased over the last couple of years. Nevertheless, only 30% of low-wage workers have access to this type of account (Figure 3).

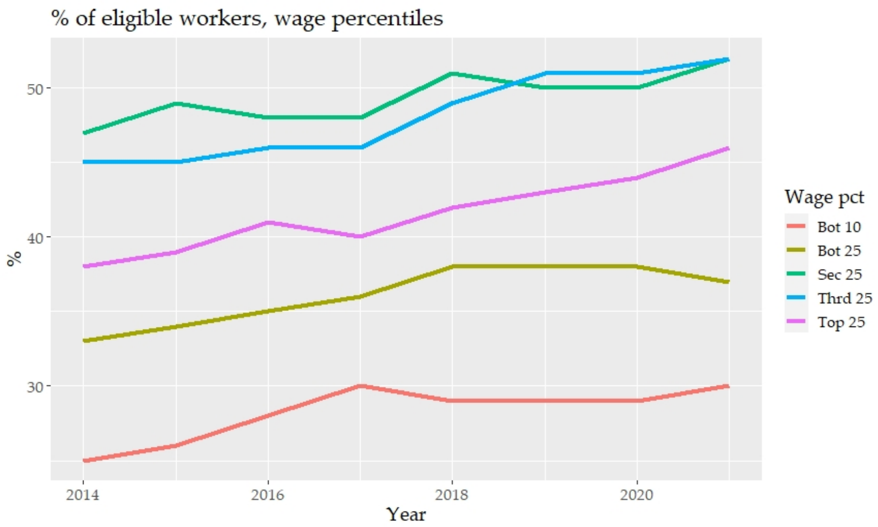


Figure 3: Workers eligible for 401(k), wage quartiles. *Bureau of Labor Statistics data*.

This paper focuses on the intensive margin and considers workers who are eligible to participate. Specifically, the life cycle model aims to capture the differences between rational and extrapolative expectations savings paths, in both liquid and illiquid retirement accounts. Expectation-driven savings paths generate the liquid savings share at the cross section. The paper assesses the performance of each expectations model by comparing liquid savings ratio to it's empirical counterpart. Therefore, the relevant data measure is the share of liquid savings in overall savings accounts, across age groups. Following Bhutta et al. (2022), liquid savings include transaction, checking and savings accounts together with directly held stocks, bonds and other financial assets. Controlling for worker demographic and financial characteristics[12], figure 4 plots the predicted share of liquid savings in all savings accounts against wage percentiles.

Figure 4 shows that the share in liquid savings remains substantial throughout the work life of low-to-middle-income workers. The flattening at the right end of the wage distribution with younger workers provides suggestive evidence for optimism-driven retirement saving delay. In

[12]Experienced bankruptcy, foreclosure, debt payments, real estate equity amount etc.


Figure 4: Share of liquid savings in overall savings by age group and wage percentile. *Survey of Consumer Finances data*, own calculations.

contrast, low- and middle- income workers draw resources from liquid savings, at the expense of attaining higher returns in illiquid accounts. Overall, the share remains high towards retirement, amounting to 40% just before retirement age.[13] That is, the data shows that workers tend to stick to their liquid savings and decrease their liquid savings share only slightly before retirement.

In explaining the high share in liquid savings, this paper's narrative relies on income growth expectations. In reality, low retirement contributions may not only be driven by income growth misperceptions across the income distribution. Taking care of housing is an example of a retirement saving delay. The MSC data analysis in the appendix supports abstracting from housing in the structural model. First, the retirement confidence measure does not vary significantly with home ownership. Moreover, it does not vary with home value. Renters and homeowners perceive their retirement equally.

The average worker pays off their mortgage for 30 years after buying the house. The model interpretation views mortgage and other mandatory payments as parts of the spending plan that

[13]I repeat the analysis with workers that own defined contribution accounts - the fitted lines across wealth percentiles retain similar shape.

in turn requires liquidity. The pessimistic worker understates future income and saves more to ensure liquidity. Consequently, the paper makes the case that homeowners and renters ensure liquidity in the same manner.

## 4 Three-period model

The stylized version of the model analytically proves that pessimism affects the decomposition of agents' savings (liquid-to-illiquid savings ratio). Pessimism induces workers to reallocate their savings from illiquid to liquid accounts.

Each agent is endowed with $y_1$ in period one and decides how much to consume and allocate to their savings accounts, liquid ($s_1$) and illiquid $s_1^R$. $2^{nd}$ period income follows a Bernoulli distribution

$$y_2 \sim \begin{pmatrix} y_L & y_H \\ p & 1-p \end{pmatrix}, \quad y_L < y_1 < y_H.$$

Agents can allocate part of their $2^{nd}$ period resources to liquid savings $b_2$ and consume the rest. In the third period, agents consume what they saved from both liquid and illiquid assets. When optimizing, agents form subjective expectations about the second period income. Pessimists assign greater probability to the bad outcome $y_L$, so

$$\hat{\mathbb{E}}[y_2] = \tilde{p}y_L + (1-\tilde{p})y_H, \quad \tilde{p} > p.$$

The maximization problem is

$$\max_{s_1, s_1^R \geq 0} u(c_1) + \mathbb{E}[u(c_2) + u(c_3)] \text{ such that } c_1 + s_1 + s_1^R = y_1,$$
$$c_2 + s_2 = y_2,$$
$$c_3 = Rs_1^R + s_2,$$

where $u(c) = \dfrac{c^{1-\gamma}}{1-\gamma}$.

Upon realization of the $2^{nd}$-period income, all uncertainty is resolved. If the agent does not run down her liquid assets, she is able to divide resources evenly across period 2 and 3, choosing $s_2 = \dfrac{y_H + s_1 - Rs_1^r}{2}$ so that $c_2 = c_3$. However, if she is constrained, $s_2 = 0$ and she consumes $c_2^L = y_L + s_1$. During retirement, she consumes retirement savings $Rs_1^R$, where $R > 1$. Given all the assumptions, two lemmas hold:

**Lemma 1.** *If $R < (1 + R^{\frac{\gamma-1}{\gamma}})$ and $y_2 = y_H \implies$ borrowing constraint does not bind (i.e., $s_2 > 0$, in the high income state).*

**Lemma 2.** *If the agent chooses to allocate to both liquid and illiquid savings and $R \neq 1$, the borrowing constraint binds in the low income state $y_L$.*

Agents know that they will be constrained in the low income state. Assume that agents aren't "too hungry" in the third period in the low income state $(u'(c_3^L) \leq 2u'(c_3^H) \implies c_3^L \geq \frac{c_3^H}{2^\gamma})$ for fixed $1^{st}$ period allocations $s_1, s_1^R$. Then, using the implicit function theorem it can be shown that the following result holds.

**Proposition.** *Suppose that retirement savings exhibit greater returns than liquid assets, $R > 1$, but are not too large, satisfying $R < (1 + R^{\frac{\gamma-1}{\gamma}})$. Define $s_1(p)$ an $s_1^R(p)$ as optimal liquid and illiquid savings. Given that the uncertainty in second period income is large enough, $s_1(p) > 0$, the following inequality holds:*

$$\frac{\partial s_1}{\partial p} > 0 > \frac{\partial s_1^R}{\partial p}.$$

*That is, an increase in pessimism (assigning $\tilde{p} > p$) implies an increase in liquid asset holdings, and a decrease retirement savings.*

# 5 Full life cycle model

Pessimism reinforces precautionary motives analytically in a three-period simple version of the model. The full life-cycle model exhibits more trade-offs due to multiple sources of income shocks and a longer time horizon. The model solution therefore relies on computational methods. The computational solution uses a transformation of the two-dimensional endogenous grid method, allowing for faster computation (Druedahl, 2020). This section outlines model assumptions and potential trade-offs workers face throughout their career.

## 5.1 Defined contribution account

The retirement account represents the private contribution account [14], including 401(k) and 403(b). All workers are eligible to open the DC account. A contribution rate choice $d_t$ entails transferring $d_t y_t$ to the DC account. The benefit function incentivizes small deposits [15]

$$h(d_t y_t) = \chi \log(1 + d_t y_t), \ \ \forall t \leq T_{ret}.$$

Retirement savings can be used only once the worker retires. Assets in the DC account exhibit a return $R_b$, which is assumed to be higher than the standard deposit account return. Let $b_t$ denote assets in the DC account after contribution $d_t y_t$. The law of motion for $b_t$ is

$$b_t = R_b n_{t-1} + d_t y_t + h(y_t d_t), \ \ \forall t = 1, \dots, T_{ret} - 1.$$

$b_t$ is the amount of savings after the contribution is made; thus it as a *post-decision* variable, whereas retirement account at the beginning of the period is denoted with $n_t$ (*pre-decision* variable). Different timing notation connects the numerical solution method to Druedahl (2020).

Setting up an account does not yield any costs and may be postponed to a later point in the work life. The minimum contribution rate is set out to be $0\%$, thus equals the minimum rate for 401(k) in the U.S. The maximum contribution rate is fixed throughout the work life following the U.S. regulation, and corresponds to a specified dollar amount each year:

$$d_t y_t \leq m.$$

Participation in the DC account is voluntary, allowing all employees to catch up with their contributions as they approach retirement. Workers cannot, however, opt-out and take the resources once they have created an account. [16]. As a result, when optimizing, the worker chooses between consuming out of assets when retired and being able to tap into the liquid account in case of an income shock. The pessimistic outlook of low-income workers incentivizes delays and low

[14] Abbreviated as DC account.

[15] Model estimates in section 6 take the smooth approximation of the step function used in a standard 401(k) employer-employee matching schedule.

[16] In the U.S., the worker can take money from the 401(k), albeit with a penalty of $10\%$ of all illiquid savings. Correspondingly, SCF data shows an insignificant amount of withdrawals.

contributions in retirement plans.

### 5.2 Liquid savings account

A standard saving instrument is a liquid account with the return the $R_a < R_b$ on accumulated liquid assets $a_t$. Lower return of liquid assets encompasses the fact that retirement accounts are tax-deferred. Liquid savings fund current spending and can be accessed at any time. The volatility misperception has the same effect on optimists, up to a point where the income level bias effect outweighs the volatility bias effect. Following Druedahl and Jørgensen (2020), the model solution separates *pre-decision* liquid assets $m_t$ (cash on hand) from the *post-decision* variable $a_t$.

### 5.3 Retiree's problem

Retirement age is deterministic $T_{\text{ret}}$. A defined contribution account is an additional liquid resource throughout retirement. That is, accumulated savings in the DC account are paid out as annuity payments $y_{an}$. The amount left after the annuity payment does not exhibit a return. Because not all workers invest in retirement accounts, a retirement benefit is provided, $\bar{b}$. The retiree's problem boils down to a standard consumption saving problem, conditional on having assets in the DC account:

$$\max_{\{c_t, a_t \geq 0\}} u(c_t) \text{ s.t. } c_t \leq m_t - a_t$$
$$m_{t+1} = R_a a_t + \mathbf{1}_{\{DC\}} y_{\text{an}} + (1 - \mathbf{1}_{\{DC\}})\bar{b}.$$

### 5.4 Worker's problem

While employed, workers receive labor income $y_t$ at the beginning of the employment year and choose the allocation of liquid savings $a_t$, retirement contribution $d_t$ and consumption $c_t$, bringing utility

$$u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}.$$

Workers face a persistent and a transitory shock each period. Subjective workers do not fully understand their income process and extrapolate from previous income realizations, whereas rational workers perceive their income correctly. Problem equations hold for both rational and subjective expectations, commonly denoted with $\mathbf{E}_t$.

State variables are current labor income, cash on hand at the beginning of period and accumulated retirement savings $(p_t, \xi_t, m_t, n_t)$. Cash-on-hand consists of accumulated liquid savings and current labor income $y_t = \alpha_i + g_{i,t} + \lambda p_{i,t} + \gamma_{i,t}$ together:

$$m_{i,t} = R_a a_{i,t-1} + y_{i,t}.$$

Throughout the rest of the paper, subscript $i$ is omitted.

The indicator function tracks DC account participation:

$$\forall t = 1, \dots, T_{ret} - 1, z_t = \begin{cases} 1; & \text{has DC acc} \\ 0; & \text{no DC acc} \end{cases}.$$

Opening the retirement account is a one-time decision, i.e.:

$$\mathcal{Z}(z_{t-1}) = \begin{cases} 1 & , z_{t-1} = 1 \\ \{0, 1\} & , z_{t-1} = 0. \end{cases}$$

Each worker maximizes the value function that is the maximum of two conditional value functions. If she did not start contributing, for $z_{t-1} = 0$

$$V(0, p_t, \zeta_t, \xi_t, m_t, n_t) = \max_{z \in \{0,1\}} \begin{cases} v_t(1, p_t, \zeta_t, \xi_t, m_t, n_t), \\ v_t(0, p_t, \zeta_t, \xi_t, m_t, 0), \end{cases}$$

or the worker already contributes, so $z_t = 1$ and

$$v_t(1, p_t, \zeta_t, \xi_t, m_t, n_t) = \max_{0 < d_t \leq 1, c_t \geq 0} u(c_t, d_t) + \beta \mathbb{E}_t \Big[ V_t(1, p_{t+1}, \zeta_{t+1}, \xi_{t+1}, m_{t+1}, n_{t+1}) \Big]$$

such that

$$
\begin{aligned}
c_t + y_t d_t &\leq m_t - a_t \\
b_t &= R_b b_{t-1} + y_t d_t + h(y_t d_t) \\
m_{t+1} &= R_a a_t + y_{t+1}.
\end{aligned}
$$

If the worker postpones DC participation, she chooses her consumption and liquid assets to transfer to the next period, earning return $R_a$ and reconsiders adding to retirement savings in the next period.

$$
v_t(0, p_t, \zeta_t, \xi, t, m_t, 0) = \max_{c_t, a_t \geq 0} u(c_t) + \beta \mathbb{E}_t \big[ V_t(0, p_{t+1}, \zeta_{t+1}, m_{t+1}, 0) \big]
$$

such that

$$
\begin{aligned}
c_t &\leq m_t - a_t \\
m_{t+1} &= R_a a_t + y_{t+1}
\end{aligned}
$$

Interior solution to the DC participant problem satisfies

$$
\begin{aligned}
\frac{c_t^{1-\gamma}}{1-\gamma} &= \beta \mathbb{E}_t \big[ v_{m,t+1} \big] \\
d_t y_t &= \frac{\chi}{R_a \mathbb{E}_t \big[ v_{m,t+1} \big] - R_b \mathbb{E}_t \big[ v_{n,t+1} \big]}.
\end{aligned} \tag{6}
$$

Expression 6 represents the trade-off workers face each period. Contribution rate increases with benefits $\chi$, and decreases whenever the current marginal value of liquidity exceeds the marginal value of saving in retirement. Pessimistic expectations overstate low income realizations and thus increase the marginal value of liquidity, which drives the difference between rational and subjective savings paths. However, as the worker approaches retirement, the value of adding to the retirement plan increases.

Table 6: Maximum contribution limits - calibrated thresholds.

| Dollar terms | Amount | Percentile |
|---|---|---|
| 2015 $ | 18 000 | 0.08 |
| | 24 000 | 0.135 |

# 6 Estimation

The model estimation follows two steps. The first step uses sample estimates of the income process in the PSID data (Cocco et al., 2005), and calibrates the income growth bias using the MSC forecast errors at every income quantile.

The benefit function matches the amount added to the retirement savings account, approximating the matching schedule in the employer-employee level data (Parker et al., 2022; Choukhmane, 2021; Beshears et al., 2020). Yearly contribution thresholds are calibrated to match the income process in the MSC data.

## 6.1 Contribution match schedule

Contribution limits correspond to the cap given by U.S. law in $2015 terms. In 2015 the maximum contribution amount was $18000 for workers under 50 and $24000 after. Each threshold corresponds to the sample percentile in the MSC income data. Both percentiles from the simulated income sample of 10000 agents using the parametrized income process serve as the constraint in the model (Table 6).

Finally, the contribution function $h(y_t d_t)$ corresponds to the matching schedule among U.S. employers. Most employers match their employee contribution rate up to $6\%$ of wage. That is, as long as the worker contributes less than $3\%$ of her wage, her employer will match with the same amount. If the contribution rate is higher than $3\%$, her employer matches with $3\%$ of the employee's pre-tax wage. In the baseline case, $3\%$ is **the default rate** and is a subject of the policy change in this paper. Since the benefit function is a smooth approximation for the employer matching schedule, the parameters are pinned down via curve fitting (Table 7).

$$h(d_t, y_t) = \chi \log(a d_t y_t + b)$$

Table 7: Benefit function parameters approximation.

| $\chi$ | $a$ | $b$ |
|---|---|---|
| 0.34 | 5.63 | 1 |

### 6.2 Other model parameters

The retirement savings return corresponds to the average return of a standard life-cycle fund, which is known to be the default and most popular choice among 401(k) contributors (Mitchell et al., 2006; Parker et al., 2022). The return on liquid assets incorporates a tax differential, since gains on 401(k) savings are tax-deferred. The rest of the model parameters correspond to standard values found in the literature.

Once simulated, consumption and savings paths define the calibration objective for the risk aversion parameter, as preferences are independent of the income expectations formation (Table 8).

Table 8: Fixed parameters in the model.

| Fixed parameter | Source | Value |
|---|---|---|
| $\beta$ | Cocco et al. (2005) | 0.98 |
| $T_{ret}$ | Love (2006) | 70 |
| $T$ | Love (2006) | 90 |
| $R_a$ | *exogenous parameter* | 1.02 |
| $R_b$ | *target-date fund performance average* | 1.04 |
| $\gamma$ | *calibrated to match illiquid-to-liquid ratio in the SCF* | 3.7 |

## 7 Solution method

The worker's function is non-convex due to opting into the retirement account. I solve the problem by utilizing the model's upper-envelope property, where the consumption choice is solved independently of the contribution choice. The upper envelope defines values comparable across workers' DC participation choices. The solution algorithm uses the endogenous grid over assets;[17] thus, it is computationally faster. For a fixed contribution rate, the worker consumes out

[17] The model solution builds on Druedahl (2020) and allows for both the persistent and transitory components in the income process.

of her net labor income and current liquid savings.

The next section shows the shape and differences in policy functions across the income distribution. The key finding - the consumption and savings plans differences between subjective and rational workers are outlined with life-cycle simulation comparisons after.

### 7.1 Policy functions

Consumption and savings are functions of current lliquid and illiquid account balances. In figures 5 and 6, the left axis represents current liquid savings, while the right axis represents current retirement savings. For a fixed level of income persistence, each plane represents consumption and contribution rates. Different income quantiles are represented as different figures and are labeled as low- and high-income, respectively. The difference between rational and subjective worker policies lies in the shape and level of the policy plane, which is shown in the appendix.

As shown in figure 5, young low-income workers (left) consume less than their high-income workers (right).

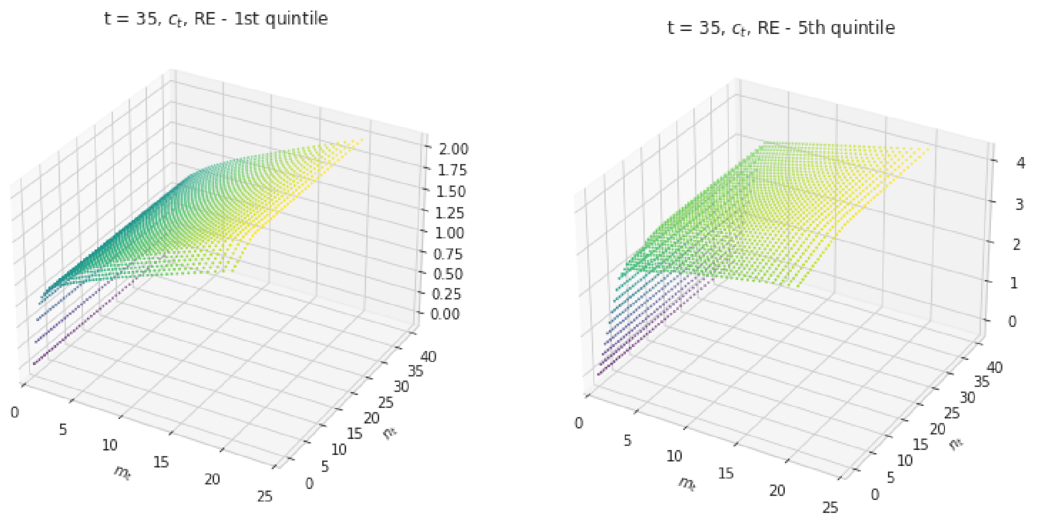


Figure 5: Rational expectations consumption policies for first (left) and fifth (right) income quintile.

Figure 6 depicts differences in contribution rates between low- and high-income young workers. Low-income workers are only incentivized to contribute more to retirement accounts if they

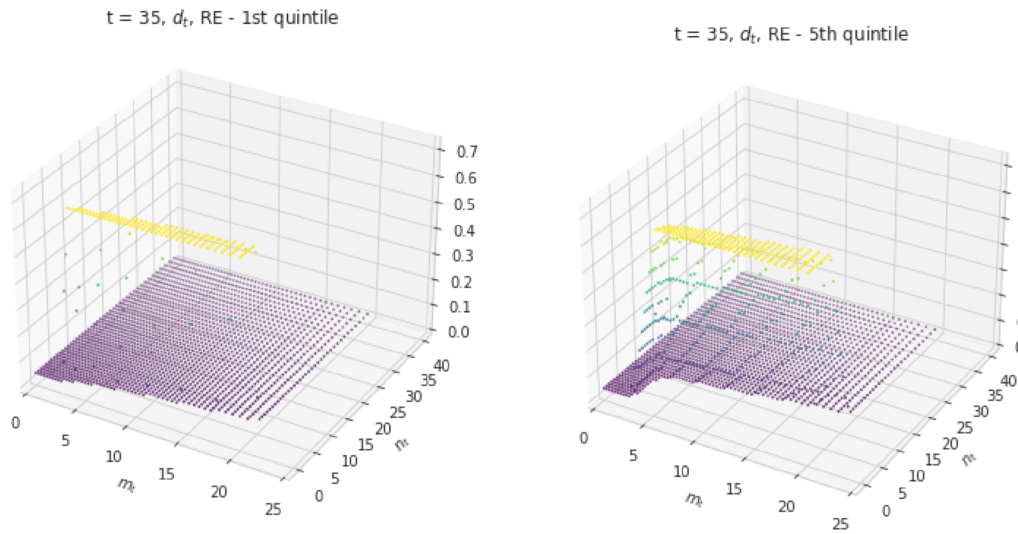


Figure 6: Rational expectations, contribution rate policies, for first (left) and fifth (right) income quintile.

are low on retirement savings, whereas high-income workers contribute at higher rates even for substantial retirement saving levels.

## 7.2 Rational and subjective worker - comparison

In the remaining part of the paper, all simulations compare the subjective expectations solution ($\mathbb{E}^*$) to the rational expectations ($\mathbb{E}$). Consumption and savings policy differences between rational and subjective workers accumulate over the life-cycle, and overall have a different effects on retirement savings. On average, subjective workers save less in retirement accounts, and rely on liquid savings instead (Figure (7), left). In addition, subjective workers consume less than their rational counterparts only to consume more once their income level bias overcomes the uncertainty misperception (Figure (7), right).

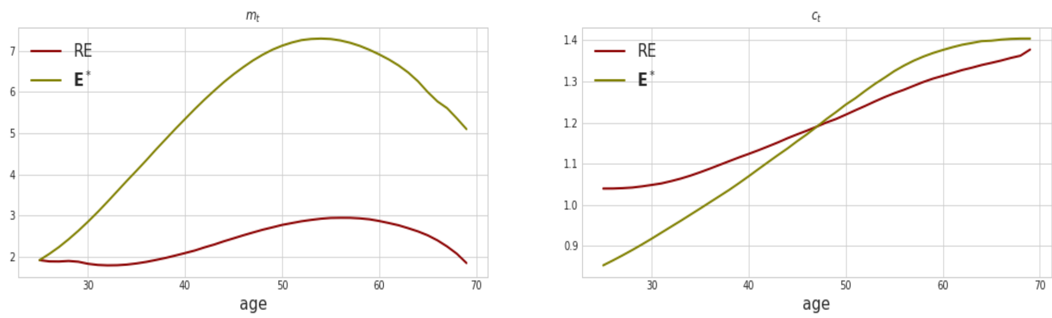


Figure 7: Liquid savings (left) and consumption (right) paths comparison for bottom 25%-income workers.

Even though the calibration targets liquid-to-illiquid assets ratios, subjective expectations solution aligns with the data on contribution rates over the work life (Choukhmane, 2021; Parker et al., 2022). Contribution rates increase steadily only to decrease just before retirement; Figure 8 contains dotted yearly contribution rates, averaged on a smaller sample of "middle class investors" from financial institution data in Parker et al. (2022). In general, subjective expectations substantiate a slow increase in contribution rates over the working period, whereas rational workers decrease their contributions over the work life. Further analysis shows that, corresponding to Parker et al. (2022), the steady increase in contributions show up regardless of the initial income.

At the bottom of the income distribution, rational workers keep their contribution rates lower and do not change over the work life (Figure 9, red line). In contrast, pessimists delay their contributions at the start of the work life, only to increase their contributions after (Figure 9, green

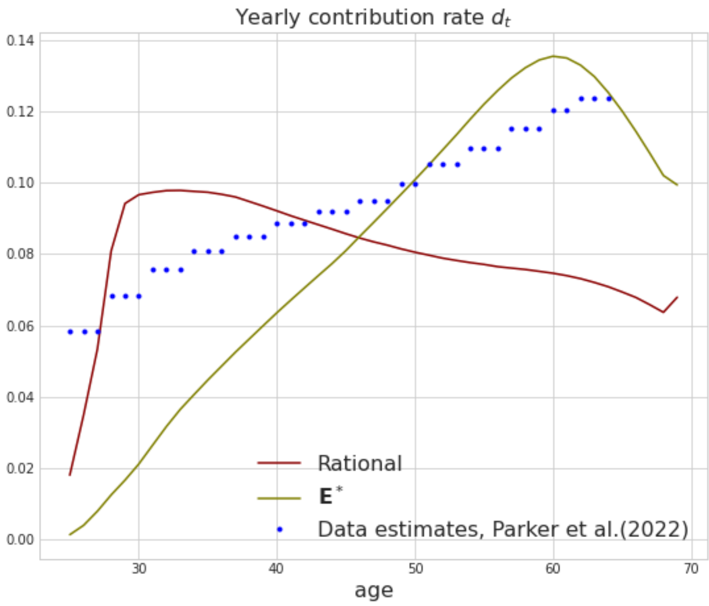


Figure 8: Contribution rates at the mean of the income distribution, rational (red) and subjective (green) life cycle paths.

line). That is, subjective workers slowly increase their contribution towards the end of work life. Due to the data unavailability on contribution rates for low-income workers[18], subjective expectations solution establishes important facts for low income workers - even though there is a delay in contributions, this delay is offset by increased contributions later on in the work life. A few years before retirement, workers decrease their contributions due to lower incentives once the retirement year is closer, corresponding the bottom tercile estimates in Parker et al. (2022).

[18] Parker et al. (2022) outline their estimates for middle class workers using the data from a financial institution. Their analysis for the bottom tercile supports the findings in this paper.

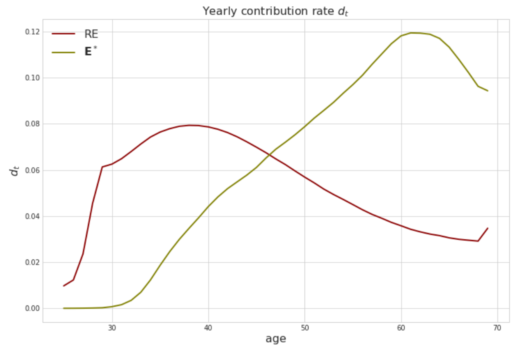


Figure 9: Rational (red) and subjective (green) lifecycle contribution paths for bottom 25%.

#### 7.2.1 Liquid-to-illiquid savings ratio across the work life

In addition to contribution rates, subjective expectation model performs well in fitting liquid-to-illiquid savings ratios to SCF data estimates throughout workers' careers. In contrast, rational expectations model understates liquid savings across the life cycle. Throughout this section I compare savings ratios to SCF data estimates from the first part of the paper (Figure 4).

Figure 10 depicts savings ratios across wage percentiles for the two models, for workers age 45-54. Subjective expectations capture the shape and slightly overstate savings ratios in comparison to SCF data estimates for the same age group (Figure 4, top panel, right graph). Moving one cohort up (Figure 11), subjective expectations capture the ratios even better, both with shape and size (Figure 4, bottom panel, left graph).

#### Subjective expectations and saving rates implications

The effects of extrapolation are also attributable to saving rates findings in the data. Specifically, Fagereng et al. (2019) find that the net saving rate (i.e., liquid savings from income) remain flat from the $20^{th}$ wealth percentile onward. Subjective expectations simulations support this empirical fact, whereas rational expectations imply increasing the net saving rate across the wealth distribution (Figure 12, left). Extrapolative expectations generate wealthy workers with incentives to save, due to the mispercieved volatility of income.

Including gross saving rates (i.e., savings that include retirement accounts), extrapolation implies a larger difference between the two saving rates across the wealth distribution, (Figure 12, right), which is consistent with empirical findings on capital gains differences across the

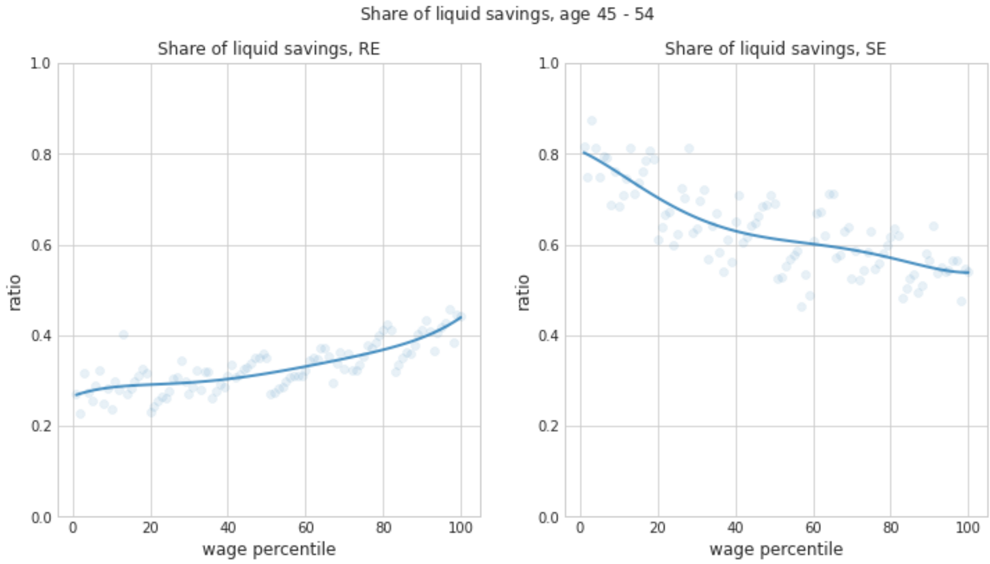


Figure 10: Savings ratios for workers age 45-54, model simulations. Rational expectations; left, and subjective expectations; right.

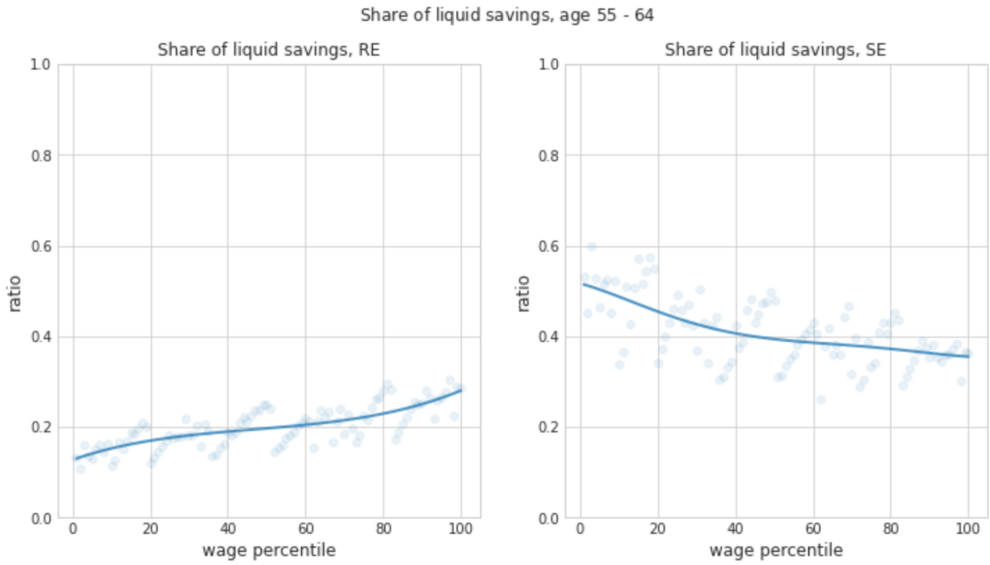


Figure 11: Savings ratios for workers age 55-64, model simulations. Rational expectations; left, and subjective expectations; right.

wealth distribution. In contrast, rational expectations solution exhibits stark differences even for the bottom 20% of the wealth distribution, which is not supported in the data (Nardi and Fella, 2017; Fagereng et al., 2019)[19]. Even when all workers are eligible to contribute, the disparity in gross saving rates across the wealth distribution highlights the effect of unrealized capital gains in retirement accounts. These gains materialize once workers reach retirement and affect retirement consumption inequality.

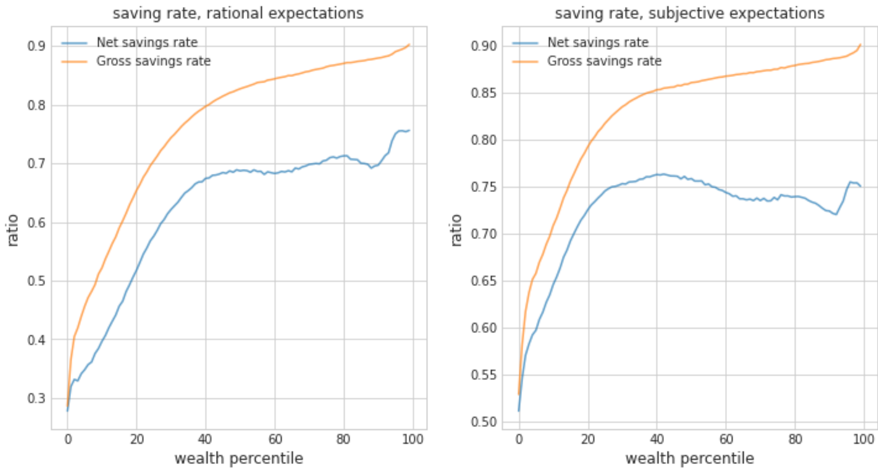


Figure 12: Net and gross saving rates, rational expectations (RE, left) and subjective expectations ($\mathbf{E}^*$, right) simulations.

All agents in the model are eligible to save in employer-matched retirement accounts. Even with the default rate and employer matching, subjective expectations create a lack of incentives to save in retirement accounts. However, the increased eligibility and growing interest in retirement savings incentives provides a foundation for specific policy evaluation that may reflect workers' responses.

# 8 Policy experiment - automatic enrollment

As a way of ensuring retirement security, automatically enrolling workers into their retirement plans has been encouraged by U.S. legislation. The majority of empirical studies estimate differences between two types of enrollment: active enrollment, in which workers actively choose to begin contributing to their 401(k) (benchmark model), and automatic enrollment, which enrolls workers automatically. Employers can thus add 401(k) accounts in their employees' names

[19]Since future spending plans (including mortgage, rent, etc.) draw out of liquid savings, liquid savings rates high.

through automatic enrollment.

The long-term effects of automatic enrollment with default rates cannot be evaluated, simply due to the recentness of the policy introduction, and the resulting findings discuss the short-term effect (5 to 7 years after the enrollment). With the exception of Choukhmane (2021), this paper tests the potential effects of automatic enrollment throughout the work life. The default rate set remains at the standard and is $3\%$, leaving workers to adjust their contributions without any costs.

Only in the first year of employment do employers make the contribution in workers' names. Given that the subjective expectations model recreates the patterns in contributions found in microdata (catching up, increasing contributions, starting with low contributions), I test for policy effects with workers who extrapolate. Policy tests imply that automatic enrollment has an insignificant effect on retirement savings right before retirement.

Figure 13 shows rational and subjective workers' consumption and savings paths across the life cycle. Subjective workers maintain their liquid buffers and thus consume less. Aside from a decrease in consumption due to the exogenous default rate in the first year of tenure, differences between the two expectation solutions remain the same.

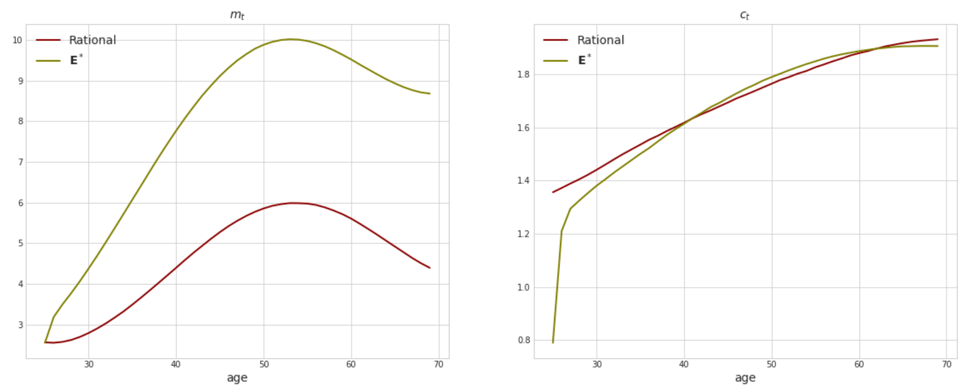


Figure 13: Rational (RE) and subjective ($\mathbf{E}^*$) liquid savings and consumption under auto-enrollment.

Even though consumption and liquid savings initially adjust to the automatic enrollment, retirement contributions in later work life are offset by the initial increase. Knowing that they need to add substantially more than preferred in the first year of tenure, workers offset first-year contributions by delaying their contributions and, ultimately, catching up with a slightly lower contribution rate towards the end of th work life (14, right). On the other hand, rational workers do not change their savings paths because they are already responding to incentives in the voluntary setting.

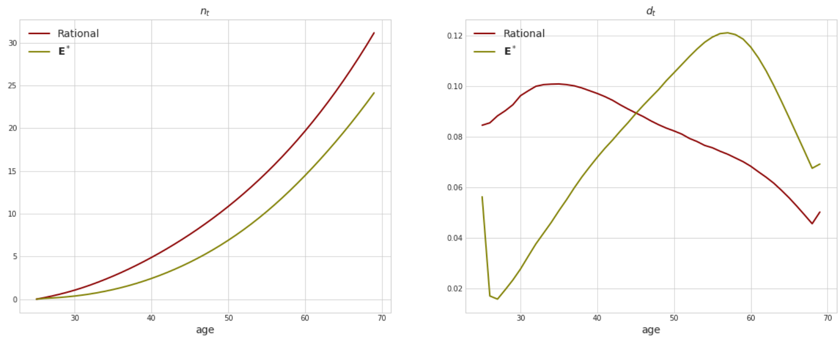


Figure 14: Rational (RE) and subjective ($\mathbf{E}^*$) retirement savings under auto-enrollment,bottom 25%.

## 8.1 Subjective workers under auto-enrollment

Comparing subjective workers' savings paths across voluntary and automatic enrollment renders the effect of auto-enrollment negligible. While voluntary contributions remain flat (non-existent) at the beginning of tenure, automatic contributions decrease right after the initial contribution made in the worker's name (Figure 15, bottom right plot). Therefore, contribution rates increase steadily. Moreover, liquid buffer amounts remain high (Figure 15, top right plot), in line with empirical findings on the insignificant effect of auto-enrollment on other financial decisions (Beshears et al., 2022).

Specifically, even though contribution rates are higher for all levels of liquid and illiquid savings (Figure 16, points in green), they are later offset, and the contribution under voluntary policy prevails (points in red). In sum, automatic enrollment has short-term positive effects, whereas long-term contributions decrease in comparison to voluntary contributors savings rates.

Table 9 shows the effect of automatic enrollment on retirement savings in the last year of tenure, for each quantile of the income distribution. Even though the average effect is negligible, the income quantile-breakdown shows a larger increase in retirement savings for bottom 50% of wage-earners (Table 9). Therefore, the size of the welfare effects varies with social planner

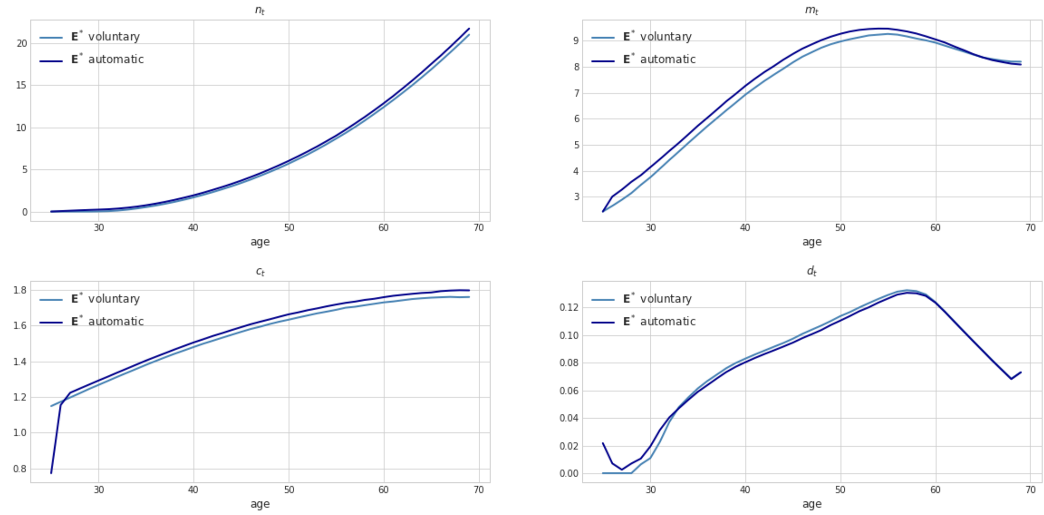


Figure 15: Subjective ($\mathbf{E}^*$) solution under voluntary and auto-enrollment setting.

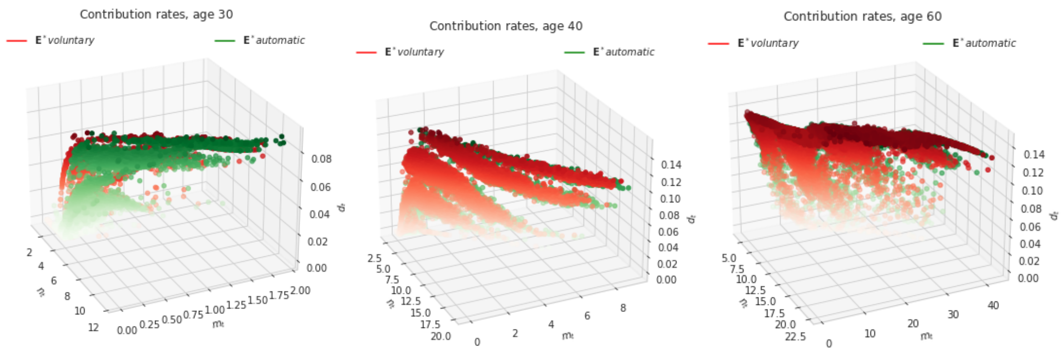


Figure 16: Simulated contribution rates under automatic enrollment, subjective expectations ($\mathbf{E}^*$).

preferences.

| | $q_1$ | $q_2$ | $q_3$ | $q_4$ |
|---|---|---|---|---|
| retirement savings increase (%) | 3,75 | 3,9 | 2,2 | 1,8 |

Table 9: Retirement savings increase under automatic enrollment.

Finally, figure 17 shows consumption differences in retirement for workers of different earnings paths. Based on the median income within last 5 years of tenure, I plot consumption paths under voluntary and automatic enrollment. Across all income quantiles, consumption differences are small. Consumption shifts upwards throughout retirement, owning to a slight increase in annuity payments each year.

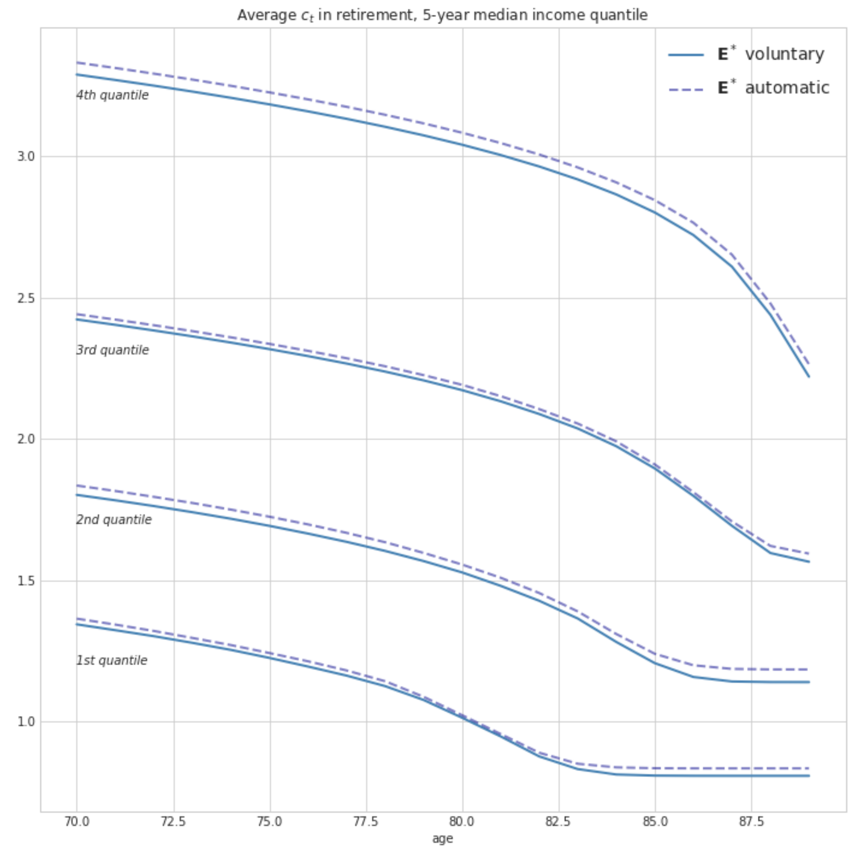


Figure 17: Subjective ($\mathbf{E}^*$) consumption in retirement under the voluntary and auto-enrollment settings.

Ultimately, agents who participate in retirement plans catch up with workers who start adding from the beginning and consume similarly throughout retirement. Including workers in retirement plans right from the beginning yields insignificant effects due to extrapolation and additional precautionary motives. The share of workers who participate increases with age and conforms to the data. In contrast, rational workers start adding from the beginning, utilizing their benefits, and thus leaving automatic enrollment testing redundant.

## 9 Conclusion

This paper introduces a deviation from rational expectations in a life-cycle model with liquid and illiquid savings accounts to explain retirement contribution patterns over the work life. The structural life-cycle model builds on individual income forecast errors found in the Michigan Survey of Consumers data. In the model, agents extrapolate from their past income realization and base their consumption and (illiquid) savings decisions on biased income projections.

The model's expectations incorporate household income forecast biases estimated from the Michigan Survey of Consumers. I expand the data analysis in Schlafmann and Rozsypal (2023) and show that households tend to extrapolate from past income growth to form expectations about their future income. Second, subjective unemployment probabilities imply volatility overstating across all workers. Third, the income forecast bias decreases over the work life.

The three-period stylized model analytically proves that pessimism induces reallocation to liquid savings at the expense of saving for retirement. Consequently, retirees consume out of liquid savings accounts. My findings suggest that pessimists require more significant incentives to save in illiquid accounts. The full life-cycle model connects extrapolation to savings allocation over time, with the presence of transitory and persistent income shocks. Saving for retirement is possible through a private retirement account closely following savings incentives in the employer-employee data.

The extrapolative expectations solution matches the empirical contribution rates pattern in the data. Workers delay participating in retirement accounts only to increase their contribution throughout their careers. Contrary to workers who extrapolate, rational workers contribute at higher rates from the start of the work life and keep their contribution rates flat later on, which is inconsistent with empirical findings (Choukhmane, 2021; Parker et al., 2022).

Even though the benefits of saving in a 401(k) plan include tax deferrals, employer matching the contribution rate, and higher expected returns, retirement studies find that workers do not add to their accounts. Therefore, automatic enrollment remains to be the policy encouraged in U.S. legislation. The recentness of the auto-enrollment policy does not allow testing for long-term effects. Since the extrapolative expectations solution captures contribution patterns across cohorts, I test for automatic enrollment policy effects throughout a worker's career. The effect of auto-enrollment on total retirement savings decreases from 3,75% at the bottom to 1,8% at the top of the income distribution. As bottom quantile continues to add to their liquid accounts, retirement consumption does not increase significantly. That is, throughout their career, workers save in the same way, in line with short-term effects in the U.S. data (Beshears et al., 2022).

This paper is the first to incorporate extrapolation in the life-cycle model to explain retirement contribution patterns. Policy tests call for novel retirement system adjustments (such as auto-escalation policies) that account for the catching up behavior. Retirement plan adjustments that follow the contribution rate patterns reduce differences in unrealized capital gains through private retirement accounts. Neglecting the data-driven bias reinforces retirement policies that have potentially misleading positive effects.

## References


Ameriks, J., Caplin, A., Leahy, J., and Tyler, T. (2007). Measuring Self-Control Problems. *American Economic Review*, 97(3):966–972.

Auclert, A., Rognlie, M., and Straub, L. (2020). Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model. Working Paper 26647, National Bureau of Economic Research.

Benartzi, S. and Thaler, R. (2007). Heuristics and Biases in Retirement Savings Behavior. *Journal of Economic Perspectives*, 21(3):81–104.

Bernheim, B. D., Fradkin, A., and Popov, I. (2015). The Welfare Economics of Default Options in 401(k) Plans. *American Economic Review*, 105(9):2798–2837.

Beshears, J., Choi, J. J., Iwry, J. M., John, D. C., Laibson, D., and Madrian, B. C. (2020). Building Emergency Savings through Employer-Sponsored Rainy-Day Savings Accounts. *Tax Policy and the Economy*, 34(1):43–90.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Skimmyhorn, W. L. (2022). Borrowing to Save? The Impact of Automatic Enrollment on Debt. *The Journal of Finance*, 77(1):403–447.

Bhandari, A., Borovička, J., and Ho, P. (2016). Identifying Ambiguity Shocks in Business Cycle Models Using Survey Data. Working Paper 22225, National Bureau of Economic Research.

Bhutta, N., Blair, J., and Dettling, L. (2022). The Smart Money is in Cash? Financial Literacy and Liquid Savings Among US Families. *Journal of Accounting and Public Policy*.

Blanchett, D., Finke, M. S., and Liu, Z. (2021). The Impact of Employer Defaults and Match Rates on Retirement Saving. Research Paper 3992899, Social Science Research Network.

Blundell, R., Pistaferri, L., and Preston, I. (2008). Consumption Inequality and Partial Insurance. *American Economic Review*, 98(5):1887–1921.

Bordalo, P., Gennaioli, N., Porta, R. L., and Shleifer, A. (2019). Diagnostic Expectations and Stock Returns. *The Journal of Finance*, 74(6):2839–2874.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2018). Diagnostic Expectations and Credit Cycles. *The Journal of Finance*, 73(1):199–227.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2003). Optimal Defaults. *American Economic Review*, 93(2):180–185.

Choi, J. J. and Mertens, T. M. (2019). Extrapolative Expectations and the Equity Premium. Technical Report 3462056, Social Science Research Network.

Choukhmane, T. (2021). Default Options and Retirement Saving Dynamics. In *Working Paper*.

Cocco, J. F., Gomes, F., and Lopes, P. (2022). Evidence on Expectations of Household Finances. Technical Report DP17447, Centre for Economic Policy Research.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *The Review of Financial Studies*, 18(2).

Dahlquist, M., Setty, O., and Vestman, R. (2018). On the Asset Allocation of a Default Pension Fund. *The Journal of Finance*, 73(4):1893–1936.

Das, M. and Van Soest, A. (1999). A Panel Data Model for Subjective Information on Household Income Growth. *Journal of Economic Behavior & Organization*, 40(4):409–426.

De Nardi, M., French, E., and Jones, J. B. (2009). Life Expectancy and Old Age Savings. *American Economic Review*, 99(2):110–15.

DellaVigna, S. (2018). Chapter 7 - structural behavioral economics. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics - Foundations and Applications 1*, volume 1 of *Handbook of Behavioral Economics: Applications and Foundations 1*, pages 613–723. North-Holland.

Devlin-Foltz, S., Henriques, A., and Sabelhaus, J. (2015). The Evolution of Retirement Wealth. *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, 108:1–34.

dH́aultfoeuille, X., Gaillac, C., and Maurel, A. (2021). Rationalizing Rational Expectations: Characterizations and Tests. *Quantitative Economics*, 12(3):817–842.

Druedahl, J. (2020). A Guide On Solving Non-Convex Consumption-Saving Models. *Computational Economics*, pages 1–29.

Druedahl, J. and Jørgensen, T. H. (2020). Can Consumers Distinguish Persistent from Transitory Income Shocks? *The Economic Journal*, 130(632):2410–2437.

Duarte, V., Fonseca, J., Goodman, A. S., and Parker, J. A. (2021). Simple allocation rules and optimal portfolio choice over the lifecycle. Technical Report 29559, National Bureau of Economic Research.

Fagereng, A., Holm, M. B., Moll, B., and Natvik, G. (2019). Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains. Working Paper 26588, National Bureau of Economic Research.

Farber, H. S. et al. (2005). What do We Know About Job Loss in the United States? Evidence from the Displaced Workers Survey, 1984–2004. *Economic Perspectives*, 29(2):13–28.

Gelman, A. (2008). Scaling Regression Inputs By Dividing By Two Standard Deviations. *Statistics in Medicine*, 27(15):2865–2873.

Ghilarducci, T., Saad-Lessler, J., and Reznik, G. (2018). Earnings Volatility and 401(k) Contributions. *Journal of Pension Economics and Finance*, 17(4):554–575.

Goda, G. S., Levy, M. R., Manchester, C. F., Sojourner, A., and Tasoff, J. (2020). Who is a Passive Saver Under Opt-in and Auto-enrollment? *Journal of Economic Behavior & Organization*, 173:301–321.

Grevenbrock, N., Groneck, M., Ludwig, A., and Zimper, A. (2021). Cognition, optimism, and the formation of age-dependent survival beliefs. *International Economic Review*, 62(2):887–918.

Krijnen, J. M., Zeelenberg, M., Breugelmans, S. M., and Van Der Schors, A. (2022). Intention and action in retirement preparation. *Behavioural Public Policy*, 6(2):191–212.

Love, D. A. (2006). Buffer Stock Saving in Retirement Accounts. *Journal of Monetary Economics*, 53:1473–1492.

Massenot, B. and Pettinicchi, Y. (2019). Can Households See Into the Future? Survey Evidence from the Netherlands. *Journal of Economic Behavior & Organization*, 164:77–90.

Mitchell, O. S., Mottola, G. R., Utkus, S. P., and Yamaguchi, T. (2006). The Inattentive Participant: Portfolio Trading Behavior in 401 (k) Plans. *Michigan Retirement Research Center Research Paper No. WP*, 115.

Nardi, M. D. and Fella, G. (2017). Saving and Wealth Inequality. *Review of Economic Dynamics*, 26:280–300.

Olafsson, A. and Pagel, M. (2018). The Retirement-Consumption Puzzle: New Evidence from Personal Finances. Working Paper w24405, National Bureau of Economic Research Paper.

Parker, J. A., Schoar, A., Cole, A. T., and Simester, D. (2022). Household Portfolios and Retirement Saving over the Life Cycle. Working Paper 29881, National Bureau of Economic Research.

Schlafmann, K. and Rozsypal, F. (2023). Overpersistence Bias in Individual Income Expectations and its Aggregate Implications. *Forthcoming, AEJ:Macro*.

Storesletten, K., Telmer, C. I., and Yaron, A. (2004). Consumption and Risk Sharing Over the Life cycle. *Journal of Monetary Economics*, 51(3):609–633.

# 10 Appendix

## 10.1 Three-period model, proof

The maximization problem for the worker in the stylized model is

$$\max_{s_1, s_1^R \geq 0} u(c_1) + \mathbb{E}[u(c_2) + u(c_3)] \text{ such that } c_1 + s_1 + s_1^R = y_1,$$
$$c_2 + s_2 = y_2,$$
$$c_3 = Rs_1^R + s_2,$$

where $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ and $2^{nd}$-period income is stochastic and distributed as

$$y_2 \sim \begin{pmatrix} y_L & y_H \\ p & 1-p \end{pmatrix}.$$

The proof builds on two lemmas that ensure that the savings constraint in the second period does not bind in the case of high income realization. This is true for reasonable conditions on the retirement savings account return:

**Lemma 1.** *If $R < (1 + R^{\frac{\gamma-1}{\gamma}})$ and $y_2 = y_H \implies$ borrowing constraint does not bind (i.e., $s_2 > 0$, in the high income state).*

*Proof.* Suppose that the claim is not true. In this case, since the high-income agent is binding ($s_2(y_H) = 0$), then the same must hold for the low-income recipient $\implies s_2(y_L) = 0$. This means that $c_3(y_H) = c_3(y_L) = Rs_1^R$. The optimality of a binding constraint implies that $c_2(y_L)$ and $c_2(y_H)$ are both strictly lower than the $3^{rd}$-period consumption[20]. However, the $1^{st}$-period FOC implies $u'(c_1) = Ru'(c_3) \implies c_3 = R^{\frac{1}{\gamma}} c_1 \implies c_1 = \frac{y_1 - s_1}{R^{\frac{1}{\gamma}-1}}$ and $c_3 = \frac{y_1 - s_1}{1 + R^{\frac{1}{\gamma}-1}} R^{\frac{1}{\gamma}}$. But then $c_3 > c_2^H \iff \frac{R^{\frac{1}{\gamma}}}{1 + R^{\frac{1}{\gamma}-1}}(y - s_1) > y_H + s_1$, which yields contradiction, since $R < (1 + R^{\frac{\gamma-1}{\gamma}})$. □

**Lemma 2.** *If the agent chooses to allocate to both liquid and illiquid savings and $R \neq 1$, the borrowing constraint binds in the low income state $y_L$.*

[20]Otherwise, both agents would be able to smooth their consumption across periods.

*Proof.* Suppose that the claim is not true, i.e. that the low-income agents do not bind. The first period optimality condition states

$$u'(c_1) = R\mathbb{E}u'(c_3),$$

for both high-income and low-income agents. Also, optimality in the second period (taking the first-order condition with respect to liquid savings) yields $u'(c_1) = \mathbb{E}u'(c_2)$. Taking the assumption into account, if the savings constraint does not bind for both types of agents then the consumption smoothing ($c_2^{L,H} = c_3^{L,H}$) implies

$$\mathbb{E}u'(c_3) = R\mathbb{E}u'(c_3),$$

which cannot be true since $R \neq 1$. $\square$

Ultimately, the effect of pessimism regarding illiquid to liquid savings accounts reallocation is implied by the two lemmas and the Implicit Function Theorem.

**Proposition.** *Suppose that retirement savings exhibit greater returns than liquid assets, $R > 1$, but are not too large, satisfying $R < (1 + R^{\frac{\gamma-1}{\gamma}})$. Define $s_1(p)$ an $s_1^R(p)$ as optimal liquid and illiquid savings. Given that the uncertainty in second period income is large enough, $s_1(p) > 0$, the following inequality holds:*

$$\frac{\partial s_1}{\partial p} > \frac{\partial s_1^R}{\partial p}.$$

*That is, an increase in pessimism (assigning $\tilde{p} > p$) implies an increase liquid asset holdings, and a decrease retirement savings.*

*Proof.* Optimal assets allocations are pinned down by the Euler equations

$$\begin{aligned} u'(y - s_1 - s_1^R) - \frac{1-p}{2}u'\left(\frac{y_H + s_1 + Rs_1^R}{2}\right) - pu'\left(y_L + s_1\right) &= 0 \;\; F_1 \\ u'(y - s_1 - s_1^R) - \frac{R(1-p)}{2}u'\left(\frac{y_H + s_1 + Rs_1^R}{2}\right) - Rpu'(Rs_1^R) &= 0 \;\; F_2, \end{aligned}$$

which implicitly define $s_1$ and $s_1^R$ as functions of the probability parameter $p$. The Implicit Function Theorem for $F = (F_1, F_2)$ implies the existence of a function $f(p) = (s_1(p), s_1^R(p))$ such that in the optimum

$$F(p, s_1, s_1^R) = 0 \quad \Longrightarrow \quad F(p, s_1(p), s_1^R(p)) = 0.$$

Now, to determine the effect of a change in the percieved probability of a low-income realization I derive

$$\underbrace{\begin{pmatrix} \partial_2 F_1 & \partial_3 F_1 \\ \partial_2 F_2 & \partial_3 F_2 \end{pmatrix}}_{\mathbf{F}} \begin{pmatrix} \partial_1 f \\ \partial_2 f \end{pmatrix} = \begin{pmatrix} -\partial_1 F_1 \\ -\partial_1 F_2 \end{pmatrix},$$

where

$$\begin{pmatrix} \partial_1 f \\ \partial_2 f \end{pmatrix} = \begin{pmatrix} \frac{\partial s_1}{\partial p} \\ \frac{\partial s_1^R}{\partial p} \end{pmatrix}.$$

The inverse of a 4-dimensional matrix is given with

$$\begin{pmatrix} \partial_2 F_1 & \partial_3 F_1 \\ \partial_2 F_2 & \partial_3 F_2 \end{pmatrix}^{-1} = \frac{1}{\det \mathbf{F}} \begin{pmatrix} \partial_3 F_2 & -\partial_2 F_2 \\ -\partial_3 F_1 & \partial_2 F_1 \end{pmatrix}$$

used to obtain $\frac{\partial s_1}{\partial p}$ and $\frac{\partial s_1^R}{\partial p}$ at the solution.

Since

$$\begin{pmatrix} \partial_1 F_1 \\ \partial_1 F_2 \end{pmatrix} = \begin{pmatrix} \frac{u'(c_2^H)}{2} - u'(c_2^L) \\ \frac{Ru'(c_3^H)}{2} - Ru'(c_3^L) \end{pmatrix}$$

and

$$\begin{aligned} \det \mathbf{F} &= \det \begin{pmatrix} -u''(c_1) - \frac{1-p}{4} u''(c_2^H) - pu''(c_2^L) & -u''(c_1) - \frac{R(1-p)}{4} u''(c_2^H) \\ -u''(c_1) - \frac{R(1-p)}{4} u''(c_3^H) & -u''(c_1) - \frac{R^2(1-p)}{4} u''(c_3^H) - R^2 pu''(c_3^L) \end{pmatrix} \\ &\overset{c_3^H = c_2^H}{=} \left(u''(c_1) + \frac{1-p}{4} u''(c_2^H) + pu''(c_2^L)\right)\left(u''(c_1) + \frac{R^2(1-p)}{4} u''(c_2^H) + R^2 pu''(c_3^L)\right) \\ &- \left(u''(c_1) + \frac{R(1-p)}{4} u''(c_3^H)\right)^2. \end{aligned}$$

Simpifying the expression yields

$$
\begin{aligned}
\det \mathbf{F} = {} & \cancel{u''(c_1^2)} + \left(\frac{R^2(1-p)}{4} + \frac{1-p}{4}\right) u''(c_1)u''(c_2^H) + R^2 p u''(c_1)u''(c_3^L) \\
& + \cancel{\frac{(1-p)^2R^2}{16}\left[u''(c_2^H)\right]^2} + \frac{R^2p(1-p)}{4} u''(c_2^H)u''(c_3^L) + pu''(c_2^L)u''(c_1) \\
& + \frac{R^2(1-p)p}{4} u''(c_2^L)u''(c_2^H) + R^2p^2u''(c_2^L)u''(c_3^L) - \cancel{u''(c_1^2)} - \frac{R(1-p)}{2} u''(c_1)u''(c_2^H) \\
& - \cancel{\frac{R^2(1-p)^2}{16}\left[u''(c_2^H)\right]^2},
\end{aligned}
$$

so

$$
\begin{aligned}
\det \mathbf{F} = {} & \left(\frac{R^2(1-p)}{4} + \frac{1-p}{4} - \frac{R(1-p)}{2}\right) u''(c_1)u''(c_2^H) + R^2 p u''(c_1)u''(c_3^L) \\
& + \frac{R^2p(1-p)}{4} u''(c_2^H)u''(c_3^L) + pu''(c_2^L)u''(c_1) + \frac{R^2(1-p)p}{4} u''(c_2^L)u''(c_2^H) \\
& + R^2p^2u''(c_2^L)u''(c_3^L) > 0,
\end{aligned}
$$

since $p \in \langle 0, 1\rangle$, i.e. the Implicit Function Theorem is applicable in this setting.

Altogether

$$
\begin{pmatrix} \frac{\partial s_1}{\partial p} \\ \frac{\partial s_1^R}{\partial p} \end{pmatrix} = \frac{1}{\det \mathbf{F}} \begin{pmatrix} \partial_3 F_2\big(\frac{u'(c_2^H)}{2} - u'(c_2^L)\big) - R\partial_2 F_2\big(\frac{u'(c_2^H)}{2} - u'(c_3^L)\big) \\ -\partial_3 F_1\big(\frac{u'(c_2^H)}{2} - u'(c_2^L)\big) + R\partial_2 F_1\big(\frac{u'(c_2^H)}{2} - u'(c_3^L)\big) \end{pmatrix}.
$$

Simultaneously

$$
\frac{\partial s_1^*}{\partial p} > 0 \text{ and } \frac{\partial s_1^R *}{\partial p} < 0
$$

hold under two conditions.

First, the two lemmas imply

$$
c_2^L < c_3^L \overset{u''<0}{\Longrightarrow} u'(c_3^L) < u'(c_2^L),
$$

so

$$-\partial_3 F_1\big(\frac{u'(c_2^H)}{2} - u'(c_2^L)\big) + R\partial_2 F_1\big(\frac{u'(c_2^H)}{2} - u'(c_3^L)\big) < \underbrace{(-\partial_3 F_1 + R\partial_2 F_1)}_{<0}\big(\frac{u'(c_2^H)}{2} - u'(c_3^L)\big).$$

Under the assumption that "the agent is not too hungry in the low-income case" $u'(c_3^L) < \frac{u'(c_3^H)}{2}$, the retirement savings decrease once the probability of the low-income realization increases. That is, pessimistic expectations $\tilde{p} > p$ the retirement savings are decreased.

Under the same assumption it has to hold

$$\partial_3 F_2\big(\frac{u'(c_2^H)}{2} - u'(c_2^L)\big) - R\partial_2 F_2\big(\frac{u'(c_2^H)}{2} - u'(c_3^L)\big) > (\partial_3 F_2 - R\partial_2 F_2)\underbrace{\big(\frac{u'(c_2^H)}{2} - u'(c_3^L)\big)}_{<0} > 0.$$

$$\partial_3 F_2 - R\partial_2 F_2 < 0 \Leftrightarrow (R-1)(u''(c_1) - R^2 p u''(c_3^L)) \overset{c_3^L > c_1}{<} (R - 1 - R^2 p)u''(c_3^L) \overset{p \in \langle 0, \frac{1}{4} \rangle}{<} 0.$$

Altogether, we have

$$\frac{\partial s_1^*}{\partial p} > 0 > \frac{\partial s_1^R *}{\partial p}.$$

$\square$

## 10.2 Additional estimates

### Age coefficients

Forecast error density estimates in the text reveal quantile-based differences and the transition from pessimism to optimism. Given that the regression coefficients with age polynomial are significant, albeit of different signs, this serves as another argument that age does affect the income growth bias.

Kernel density estimates consider only the working-age population and reveals that error distributions differ among across age groups. During the work life, the mode of the errors distribution is positive. This finding indicates that even experienced workers do not completely correct their forecasts. The distribution changes shape over age groups, owing to income volatility for younger cohorts. Correspondingly, in the model, agents start to add to their retirement accounts as the effect of misperception in income volatility decreases.

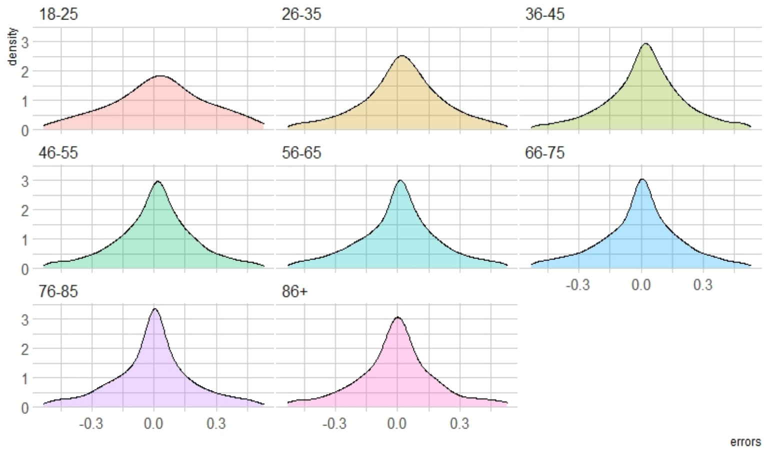


Figure 18: Income growth bias density by age group, MSC data

### 10.2.1 Regression checks

In addition to the linear regression model in the text, I estimate the model with HH who had their first interview in the second half of the year. Their responses are not sensitive to the imperfect time overlap between the period of expectations and realizations. Again, standard errors are clustered at the region level. Signs of all coefficients remain the same, while the size of coefficient with the income quantile input increases (Table 10). Again, results indicate that household tend to be overly pessimistic at the left end of the income distribution while their right-end counterparts tend to be overoptimistic.

Table 10: Linear Regression Results

| | *Dependent variable:* |
|---|---|
| | Income Growth Forecast Errors |
| $q_2$ | 0.225*** |
| | (0.008) |
| $q_3$ | 0.306*** |
| | (0.008) |
| $q_4$ | 0.356*** |
| | (0.009) |
| $q_5$ | 0.430*** |
| | (0.014) |
| $male$ | −0.015* |
| | (0.006) |
| no HS | 0.039*** |
| | (0.006) |
| college | −0.054*** |
| | (0.003) |
| $age$ | −0.114*** |
| | (0.026) |
| $age^2$ | 0.155*** |
| | (0.034) |
| $1\ adult$ | 0.106*** |
| | (0.005) |
| $> 2\ adults$ | −0.029 |
| | (0.008) |
| Constant | −0.381*** |
| | (0.012) |
| Observations | 29,414 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

### 10.2.2 Housing as a mean of saving for retirement

Finally, retirement savings may not include only private retirement accounts, as people may be saving in other illiquid savings such as housing. I address the issue of saving for retirement through real estate by checking to what extent home ownership affects retirement confidence. I use the survey question that asks to assign the probability of having a comfortable retirement **only from social security and job pensions**.

I binned subjective probabilities into four separate groups ($< 25\%$, $25 - 50\%$, $50 - 75\%$ and $75 - 100\%$) that translate into groupings from harsh pessimists to enthusiastic optimists. The estimates imply that retirement confidence rises with age and income, whereas owning a home does not have a significant effect. Since the recent income growth forecast error is not significant, I conclude that retirement confidence is based on individual attitudes (persistent pessimism or optimism).

| | *Dependent variable:* |
|---|---|
| | P(comfortable retirement) |
| $male$ | $0.215^{***}$ |
| | (0.026) |
| $no\ HS$ | 0.093 |
| | (0.074) |
| $college$ | 0.035 |
| | (0.028) |
| $age$ | $0.507^{***}$ |
| | (0.029) |
| $1\ adult$ | 0.055 |
| | (0.034) |
| $> 2\ adults$ | $0.081^{*}$ |
| | (0.041) |
| $q_2$ | $0.122^{***}$ |
| | (0.058) |
| $q_3$ | $0.259^{***}$ |
| | (0.056) |
| $q_4$ | $0.416^{***}$ |
| | (0.058) |
| $q_5$ | $0.482^{***}$ |
| | (0.060) |
| $homeowner$ | 0.060 |
| | (0.037) |
| Observations | 20,743 |
| *Note: Controlled for year effects, age is standardized.* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

Table 11: Ordered logistic regression results

In addition to retirement confidence, I check to what extent homeownership affects future income growth forecast. Since my model incorporates illiquid savings in the form of the retirement account, I check if housing assets position affect income growth forecasts. Including homeownership in the regression analysis reduces number of observations to 37,000. The income quantile coefficients remain similar. Moreover, among homeowners, home value has a significant,albeit small, effect (Table 12).

### Job loss predictions

In the main text I argue that the income quintile is the significant predictor for pessimistic job loss predictions. Consequently, once I compare empirical job separation rates to the predicted values I argue that overstating these probabilities remains consistent with how the income growth forecast bias is implemented in the life cycle model. Thus, the fact that the mispercieved persistence parameter implies the mispercieved volatility remains consistent with empirical estimates. The only age group that is significant are workers closer to retirement age.

Table 12: Linear regression results

| | *Dependent variable:* | |
|---|---|---|
| | Income growth forecast errors | |
| | Homeowners only | All |
| Income quantile: | | |
| $q_2$ | $0.176^{***}$ | $0.189^{***}$ |
| | (0.018) | (0.014) |
| $q_3$ | $0.248^{***}$ | $0.273^{***}$ |
| | (0.008) | (0.007) |
| $q_4$ | $0.272^{***}$ | $0.316^{***}$ |
| | (0.010) | (0.009) |
| $q_5$ | $0.355^{***}$ | $0.381^{***}$ |
| | (0.018) | (0.015) |
| male | $-0.008$ | $-0.014^{***}$ |
| | (0.007) | (0.006) |
| HS | $0.030^{*}$ | $0.050^{***}$ |
| | (0.014) | (0.008) |
| College | $-0.031^{***}$ | $-0.049^{***}$ |
| | (0.008) | (0.004) |
| $age$ | $-0.123$ | 0.016 |
| | (0.058) | (0.043) |
| $age^2$ | 0.095 | 0.048 |
| | (0.052) | (0.043) |
| 1 adult | $0.068^{***}$ | $0.081^{***}$ |
| | (0.011) | (0.005) |
| $>2$ adults | $-0.022^{*}$ | $-0.039^{***}$ |
| | (0.009) | (0.011) |
| Home value, quantiles: | | |
| $h_2$ | $-0.037^{***}$ | – |
| | (0.008) | |
| $h_3$ | $-0.031^{***}$ | – |
| | (0.012) | |
| $h_4$ | $-0.067^{***}$ | – |
| | (0.016) | |
| $h_5$ | $-0.088^{***}$ | – |
| | (0.013) | |
| Renter | – | $0.068^{***}$ |
| | | (0.003) |
| Constant | $-0.080$ | $-0.334$ |
| | (0.362) | (0.330) |
| Observations | 11,992 | 36,932 |
| *Note:* | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

Table 13: Ordered logistic regression results

| | *Dependent variable:* |
|---|---|
| | P(job loss within 5 years) |
| $male$ | 0.081*** |
| | (0.029) |
| $no\ HS$ | 0.288*** |
| | (0.080) |
| $college$ | −0.123*** |
| | (0.032) |
| age 25-34 | −0.112 |
| | (0.073) |
| age 35-44 | −0.050 |
| | (0.068) |
| age 45-54 | −0.023 |
| | (0.067) |
| age 55-66 | −0.550*** |
| | (0.069) |
| $1\ adult$ | −0.049 |
| | (0.038) |
| $> 2\ adults$ | 0.226*** |
| | (0.047) |
| $q_2$ | −0.001 |
| | (0.062) |
| $q_3$ | −0.115* |
| | (0.060) |
| $q_4$ | −0.208*** |
| | (0.062) |
| $q_5$ | −0.324*** |
| | (0.064) |
| Observations | 20,395 |
| *Note: Year effects are not reported.* | *p<0.1; **p<0.05; ***p<0.01 |

## 10.3 Model equations and numerical implementation

The agent's problem can be formulated as the dynamic programming problem, for the state variables mentioned in this paper. The model for subjective expectations satisfies the same equation with different expectations, so the derivations hold for both subjective and RE agents. The model given in the paper satisfies the Bellman equation:

$$v(1, m_t, p_t, \xi_t, \zeta_t, n_t) = \max_{0.03 \leq d_t \leq 1, c_t \geq 0} u(c_t) + \beta \mathbb{E}_t \big[ v(1, m_{t+1}, p_{t+1}, \xi_{t+1}, \zeta_{t+1}, n_{t+1}) \big] \tag{7}$$

$$\text{such that all the equations hold}$$

The value function $V_t$ from the paper is not necessarily concave because of the discrete opting-in decision as an absorbing state. The *Nested endogenous grid method* uses the FOC for consumption[21]

$$c_t \ldots u'(c_t) = \beta R_a v_{m,t+1}(1, m_{t+1}, p_{t+1}, \xi_{t+1}, \zeta_{t+1}, n_{t+1})$$

and the standard approach of the EGM in general - computing the continuation value beforehand. The continuation value is obtained with functions of *post-decision* variables that map the solution into *pre-decision* variables. Following Druedahl (2020)

$$w_t(a_t, b_t, p_t) = \beta \mathbb{E} \big[ v_t(1, m_{t+1}, p_{t+1}, \xi_{t+1}, \zeta_{t+1}, n_{t+1}) \big]$$

for end-of period assets $a_t$ and $b_t$ and the persistent component $p_t$. The agent who does not contribute to DC account in time $t$, faces the standard consumption-savings problem, which is easily solved using EGM. Fixing $d_t \implies b_t = n_t + d_t y_t + \chi \log(1 + d_t y_t)$ and $n_t t + 1 = R_b b_t$, (7) boils down to

$$v(1, m_t, p_t, \xi_t, \zeta_t, n_t | d_t) = \max_{c_t \geq 0} u(c_t) + w_t(p_t, a_t, b_t) \big]$$

$$\text{such that } \; c_t \leq m_t - y_t d_t - a_t$$

$$m_{t+1} = R_a a_t + y_{t+1},$$

[21] Interior solution

which fixes the idea of interpolation of the continuation value for $a_t, b_t$ and consequently using the FOC

$$u_c(c_t|d_t) = w_{a,t}(p_t, a_t, b_t) \implies m_t = a_t + c_t, \tag{8}$$

and using the upper envelope method to interpolate values for all $d_t$ on the exogeneous grid for cash-on-hand. Once consumption is calculated from (8) for each $d_t$, I use the grid search as in Druedahl (2020) and evaluate the optimal $d_t$. Even though interpolating the continuation value via post-decision variables seems cumbersome, I use the modified version of the interpolation, which shortens the grid-search procedure, due to monotonicity of the value function with respect to cash-on-hand generated by the previous period monotonically increasing assets.

If the agent does not contribute, the problem is then similar to the contributor's, i.e. it can be nested into the special case for $d_t = 0$. I solve for the problem using the same "inner" function as above, using the post-decision value function. However, the post-decision function is corrected for $b_t = 0$. Once both consumption choices are computed, I use the upper envelope method that combines the solutions to the common grid for cash-on-hand, so that the values are comparable and defined on a regular grid.

Once retired, each agent gets the annuity payment out of their retirement account, so the retirement consumption depends on both liquid account and the amount saved for retirement through the DC account. Agents who did not contribute to the account get the minimum yearly income[22]. I build on Druedahl (2020) and extend the retirement consumption function solution method to a two-dimensional space.

I build on Druedahl (2020) and implement the solution method for a persistent ($AR(1)$) process. That is, I extend the method to track all the combinations of shocks to both the persistent and transitory components (altogether 30 combinations).

The grids are finer at low values of both savings accounts to take a closer look at the behavior of the poor low-income workers (or workers with low retirement savings). Utility function is a standard CRRA function

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma},$$

[22]I computed solutions for various cases of minimum retirement subsidy. In every scenario upper quintiles end up contributing at some point in their worklife, so the subsidy is set at the lowest value of income grid.

where $\gamma = 2$. I did not resort to high $\gamma$s to isolate the effect of subjective expectations on the percieved variance of the future income.

## Consumption and savings paths

Both correct and biased income forecasts imply consumption and savings paths that follow the usual patterns found in the data. For example, both consumption paths exhibit the decrease in consumption towards the end of life, which is commonly stated as the *retirement consumption puzzle* in retirement studies(De Nardi et al., 2009; Olafsson and Pagel, 2018).

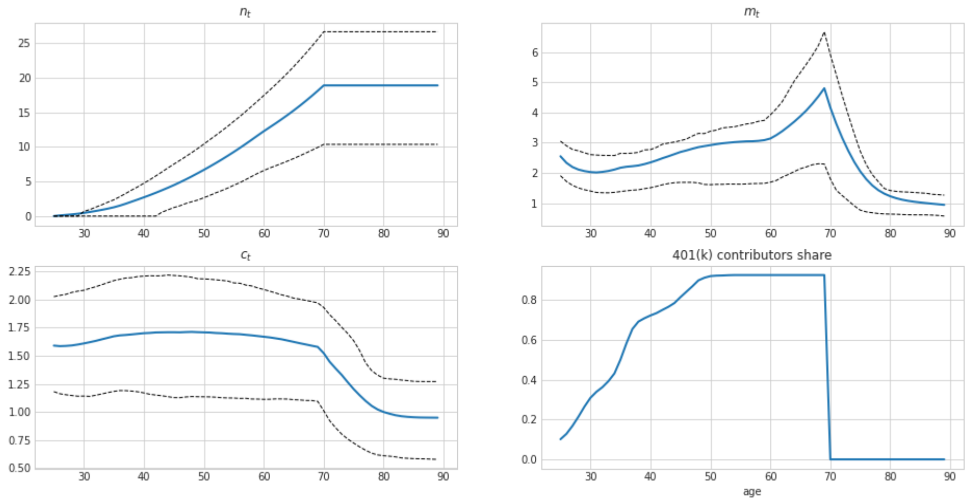


Figure 19: Lifecycle paths - RE

The effect of mispercieved volatility acts across the income distribution, which is shown in the policy function plots (Figure 20). The shares of DC contributors are lower for all income quantiles. Ultimately, all workers start adding and catching up. Ultimately, there is a point in the work life where workers forgo their liquid assets and start adding to illiquid ones.

## Policy functions, 3-d planes

As stated in the paper, differences in life-cycle paths come from differences in period-by-period consumption and savings allocations. Period-policy functions reflect additional precaution with young subjective workers and a slow increase in retirement contributions later on in life. All of the policy functions are depicted as functions of liquid and illiquid savings accounts. I denote the median within the income distribution with a red dot.

Consumption policy differs across income quantiles and, of course, expectations about future income. Figure (20) shows consumption as a function of savings levels for workers in the top 50% of the income distribution at the age of 30. At all savings levels, subjective workers consume less than their rational counterparts ( Figure 20). Moreover, having a large amount of retirement savings does not imply higher consumption in the subjective expectation solution (Figure (20), right). Therefore, precautionary motives generate saving at the beginning of the work life.

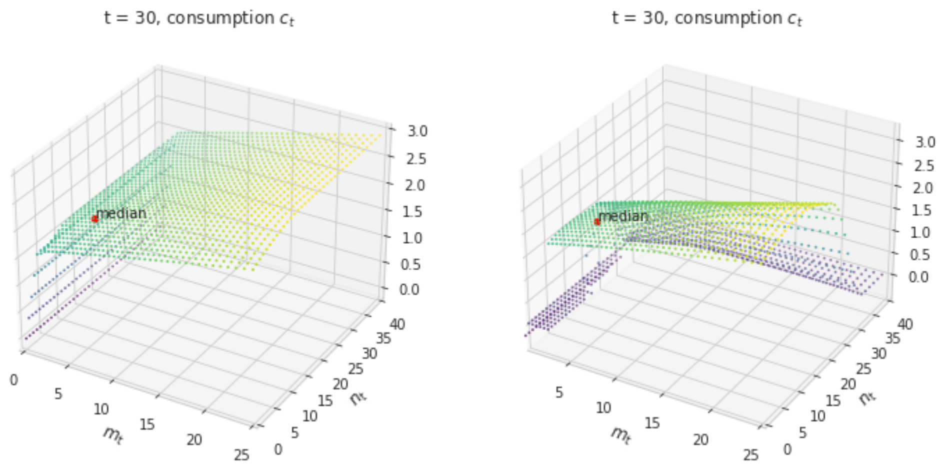


Figure 20: Consumption function, RE (left) and subjective expectations (right) solution.

Corresponding to everything presented in the main part of the paper, contribution rates switch from being lower for subjective workers at the beginning of work life. Figure 21, depicts contribution rates for top 50% of the income distribution - rational contributor adds 12%, whereas subjective one adds 6% of her current income. Later on in the work life, subjective workers start adding more and catch up (Figure 22, depicted for the top 50% of the income distribution). At the median, rational contributor adds 9% out of their wage, whereas subjective contributors add 12% (Figures 22, red dots in respective graphs). As the effects of income volatility overstating fade, workers with significantly low liquid savings contribute at the highest rates possible. Overall, contribution rates switch places for **all income quantiles**, as functions of liquid savings $m_t$ and illiquid retirement account amounts $n_t$. Therefore, the effect of extrapolation does not depend on the amount of workers' savings, owing to the misperception of future income realizations only.

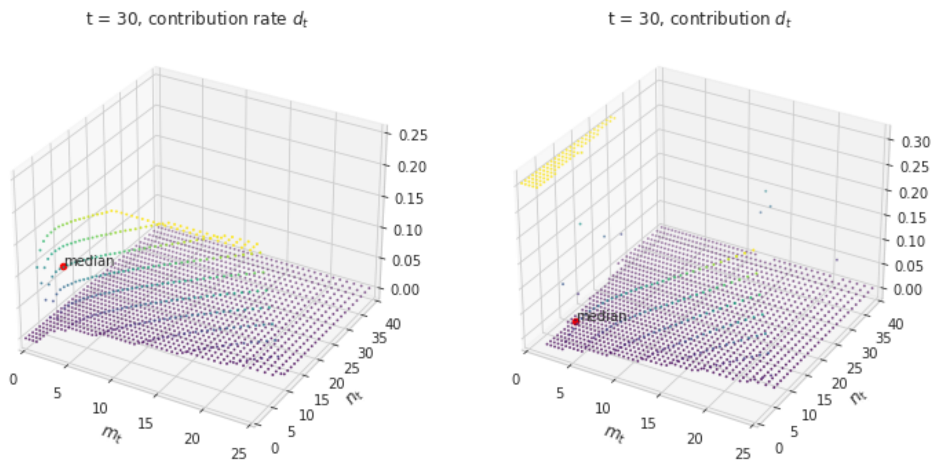


Figure 21: Median contributor at age 30, RE (left) and subjective expectation solution (right).

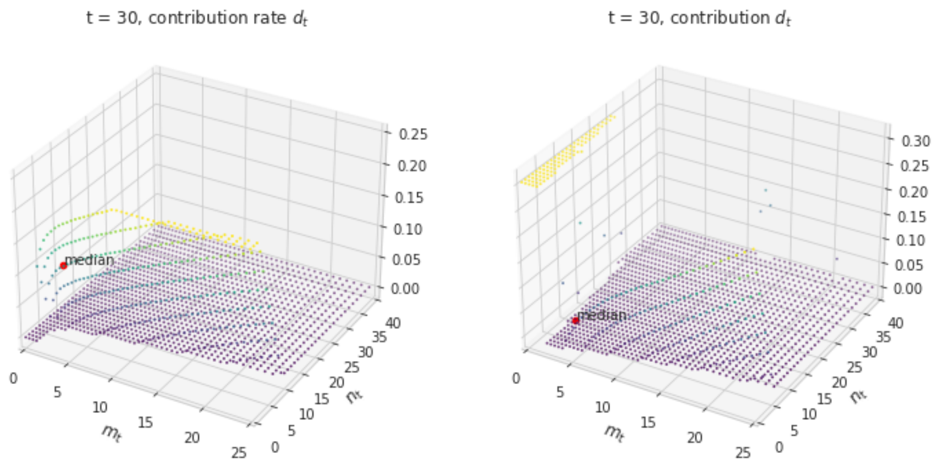


Figure 22: Median contributor at age 50, RE (left) and subjective expectation solution (right).

Rational expectations solutions overstate the share of contributors in the economy, when compared to empirical studies. On the other hand, contributor share is increasing for the extrapolative solution; as agents age they decide to participate in the DC account (Figure 23).

Quantile based comparisons show the presence of both optimism and pessimism on each

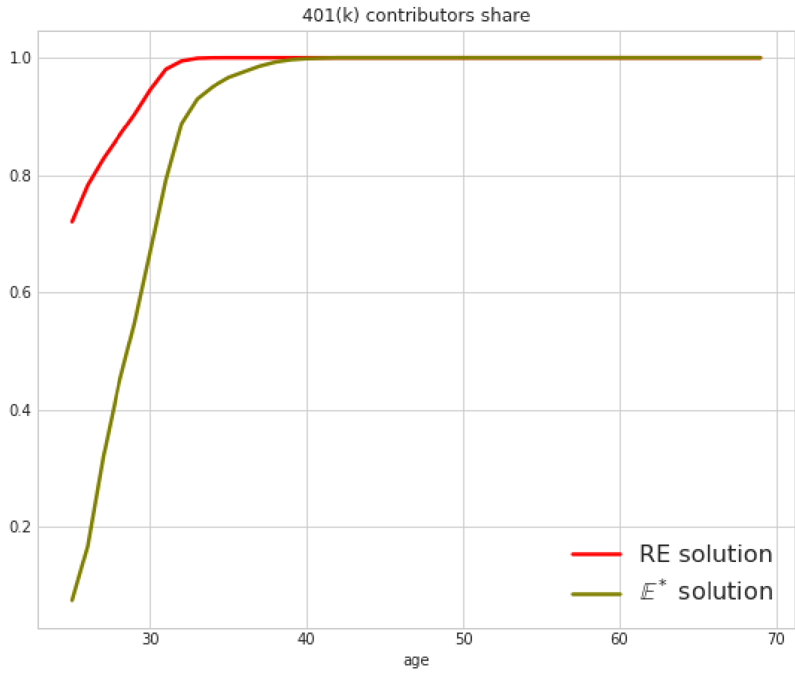


Figure 23: DC contributors share, rational (red) and extrapolative (green) expectations solution.

side of the income distribution. In contrast to the rational expectations solution (Figure 24), workers who extrapolate start participating later (Figure 25).

Contribution rates differences vary by income quantile. In each part of the income distribution, subjective expectations capture the slow increase in contributions over the tenure, whereas the rational solution fails in this respect. This is not the case with rational workers - low income workers even decrease their contribution rates (Figure 26). In this regard, including extrapolative expectations shows that eligibility may be enough for low income workers to contribute in an auto-escalating manner. On the other hand, top-income workers who extrapolate contribute at significantly higher rates later on in work life, matching empirical patterns (Figure 27).

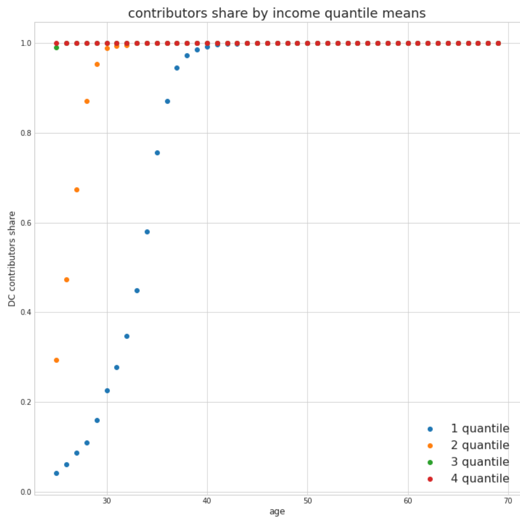


Figure 24: Rational workers, DC contributors' share by income quantile over the work life.

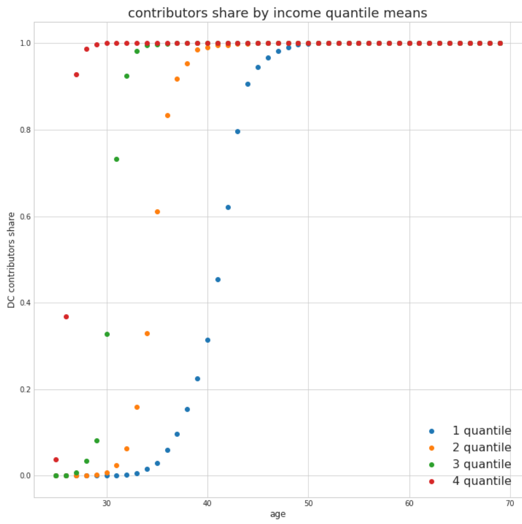


Figure 25: Biased workers, DC share lifts off gradually over the work life.

## 10.4 Savings ratios for the youngest and oldest workers

As previously mentioned, even though subjective expectations solution overstates liquid-to-illiquid savings ratios for the youngest cohort (Figure 28, right graph), model simulations show that the shape of the savings ratio across wage percentiles matches the empirical estimates from

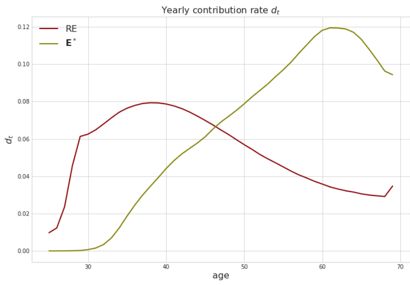


Figure 26: Contribution rates over the tenure, bottom 25%.

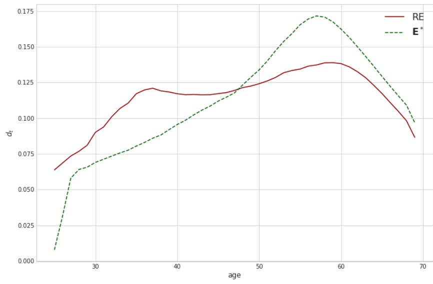


Figure 27: Contribution rates over the tenure, top 25%.

the SCF data. In contrast, rational expectations do not match the shape or the size (Figure 28, left graph). Moreover, just before retirement, the savings ratios of subjective workers are matched in shape, but slightly understated when compared to SCF workers (Figure 29, right graph and Figure 4 from the main part of the paper).

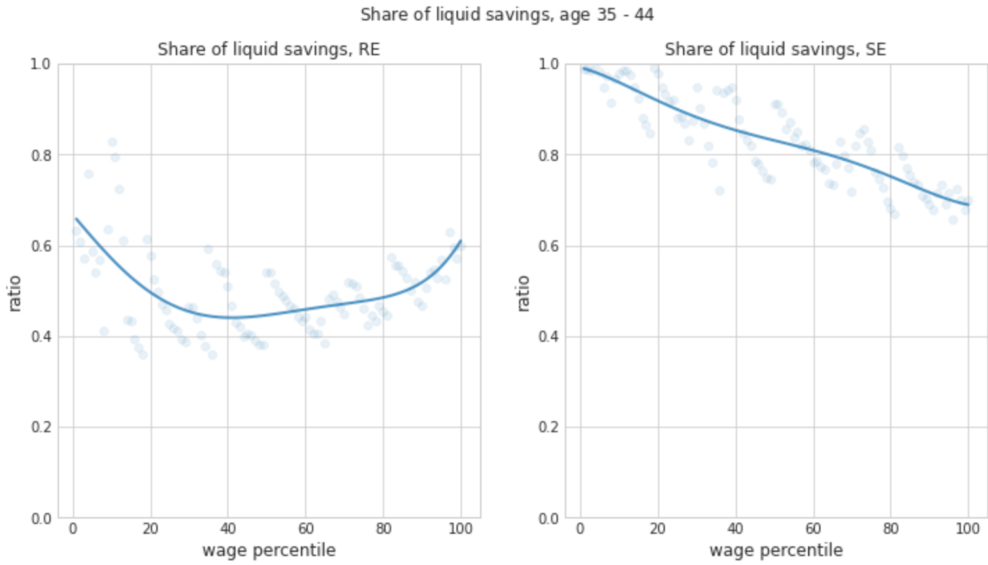


Figure 28: Savings ratios across wage percentiles, model simulations for workers aged 35-44. Rational expectations; left and subjective expectations; right.

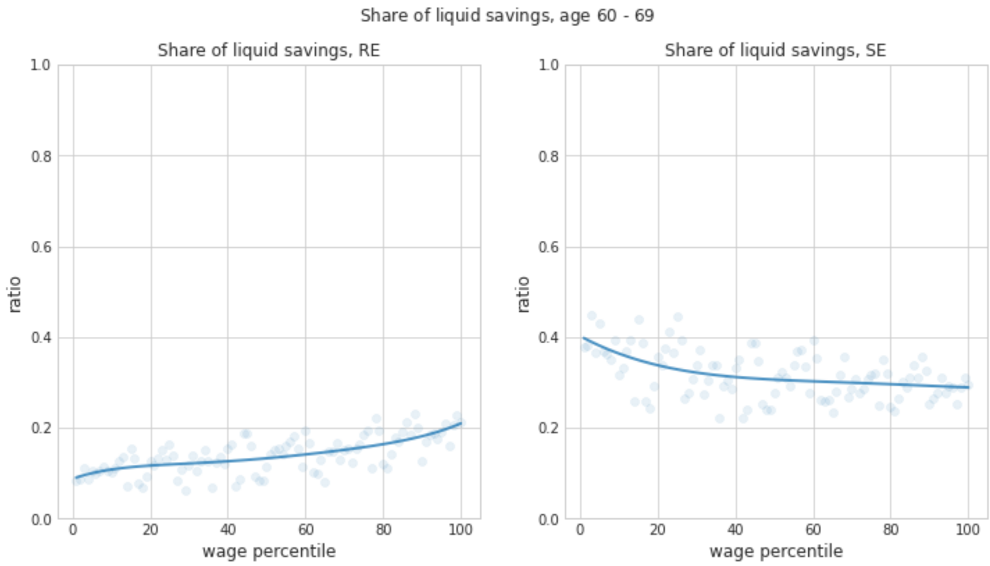


Figure 29: Savings ratios across wage percentiles, model simulations for workers aged 60-70. Rational expectations; left and subjective expectations; right.

# Abstrakt

Proč se penzijní příspěvky zaměstnanců během jejich kariéry postupně zvyšují? Tento článek používá strukturální model životního cyklu založený na údajích o očekáváních domácností k vysvětlení rozhodnutí pracovníků o penzijních příspěvcích. Data Michigan Survey of Consumers ukazují, že mladé domácnosti extrapolují ze svých nedávných příjmů a nadhodnocují persistenci a volatilitu svých budoucích příjmů. Strukturální model životního cyklu s extrapolativními očekáváními kvantifikuje rozdíl v důchodových příspěvcích ve srovnání s racionálními očekáváními. Na rozdíl od racionálních pracovníků se příspěvky extrapolativních pracovníků shodují s daty o penzijních příspěvcích v průběhu životního cyklu. V důsledku toho má zavedení povinné účasti v penzijním spoření jen zanedbatelné dopady na výši příspěvků.

Working Paper Series
ISSN 2788-0443

Individual researchers, as well as the on-line version of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the CAS, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.



Published by
Charles University, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the CAS, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: https://www.cerge-ei.cz/

Editor: Byeongju Jeong

The paper is available online at https://www.cerge-ei.cz/working-papers/.

ISBN 978-80-7343-558-5 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-675-8 (Národohospodářský ústav AV ČR, v. v. i.)