Learning to Be Rich:
How Expectations Amplify Wealth Inequality*

Pau Belda   Janko Heineken   Adrian Ifrim

November 2025
(Latest version here)

Abstract

We document systematic heterogeneity in stock-market beliefs across the wealth distribution: low-wealth households are persistently too pessimistic, while high-wealth households' expectations are approximately unbiased. We develop a heterogeneous-agent model in which belief dispersion arises endogenously through learning from experience, generating a feedback loop: good past returns foster optimism, induce higher equity shares, and lead to higher future returns that further reinforce optimistic beliefs. The calibrated model matches key features of the joint distribution of expectations, portfolio returns, and wealth. Heterogeneous beliefs increase the top 1% wealth share by 50% relative to homogeneous expectations. Methodologically, we show that Internal Rationality—where households learn directly about the law of motion for prices rather than forecasting entire distributions—makes heterogeneous-agent models with aggregate risk both more realistic and tractable.

Keywords: Heterogeneous Expectations, Wealth Inequality, Internal Rationality.

*Preliminary version. First version: May 2023. Previous versions have circulated under the title *"Heterogeneous Expectations and Wealth Inequality"*. Belda: Bank of England (Pau.Belda@bankofengland.co.uk), Heineken: European Comission (Janko.Heineken@ec.europa.eu), Ifrim: European Comission (Adrian.Ifrim@ec.europa.eu). We have benefited from comments from Albert Marcet, Juan Rubio-Ramírez, Woulder Den Haan, Olivier Coibion, and seminar participants at the Federal Reserve Board, CREi, BSE - Summer Forum, Bellaterra Macroeconomics Workshop, ERMAS. The views expressed here are solely of the authors and do not represent, in any way, those of the Bank of England, the European Commission or any of its committees. This research received funding from the European Research Council under the European Union's Horizon2020, research and innovation program GA project number 788547 (APMPAL-HET).

1. Introduction

The rise in wealth inequality is one of the central economic developments of recent decades. Across advanced economies, wealth has become increasingly concentrated at the top. In the United States, for example, the wealth share of the top 1% has risen from about 23% in 1980 to over 35% today.[1] While modern macroeconomics has responded by placing inequality at center stage, standard models often struggle to generate the extreme concentration of wealth observed in the data —especially at the very top.[2]

This paper takes a step toward closing that gap by relaxing one of the core assumptions shared by most of these models: homogeneous expectations.[3] Despite the central role of beliefs in saving and portfolio decisions, the vast majority of heterogeneous-agent macro models retain what Sargent termed a "Communism of Beliefs," under which all agents share the same subjective distribution over future variables, typically identified with full-information rational expectations. This stands in contrast to a growing empirical literature documenting substantial and systematic heterogeneity in beliefs about key macro-financial variables, including stock returns, inflation, and house prices, across households and along the wealth distribution (e.g., Vissing-Jorgensen (2003), Malmendier and Nagel (2011), Giglio et al. (2021)). Our contribution to this debate is to show that such heterogeneous expectations can themselves be a powerful amplification mechanism for wealth inequality, helping heterogeneous-agent models better match the observed concentration of wealth at the top.[4]

Among the various dimensions of belief heterogeneity, we focus on stock return

[1]Data from `realtimeinequality.org/`

[2]Typically, matching the extreme concentration of wealth in the data requires strong assumptions about bequest motives, exogenous return heterogeneity, or rare entrepreneurial fortunes (e.g., Castaneda, Diaz-Gimenez, and Rios-Rull (2003); De Nardi (2004); Benhabib, Bisin, and Zhu (2011); Kaymak and Poschke (2016); Hubmer, Krusell, and Smith Jr (2021); Kaymak, Leung, and Poschke (2022); Benhabib, Cui, and Miao (2024)).

[3]Previous research has identified various drivers of wealth concentration, including bequests (e.g., De Nardi (2004)), preference heterogeneity (e.g., Krusell and Smith (1998)), earnings risk (e.g., De Nardi, Fella, and Pardo (2016)), return heterogeneity (e.g., Fagereng et al. (2020)), entrepreneurship (e.g., Quadrini (2000)), and tax policy (e.g., Hubmer, Krusell, and Smith Jr (2021)). We contribute to this literature by providing a novel microfoundation for heterogeneous returns and portfolio choices, showing how differences in beliefs can generate and amplify wealth inequality.

[4]Recent work by Tobias et al. (2022) is a key exception, incorporating heterogeneous beliefs about inflation, unemployment, and house prices in a heterogeneous-agent incomplete-markets model. We share their emphasis on belief heterogeneity but focus on expectations about stock returns and their interaction with portfolio choice, and we propose an alternative information structure that generalizes rational expectations and reduces computational complexity.

expectations for both theoretical and empirical reasons. Stock market beliefs directly determine portfolio allocation between safe and risky assets, creating persistent differences in portfolio composition that compound over time. As Figure 1 shows, survey data reveal substantial disagreement about stock returns: rather than clustering around a consensus forecast, households' expectations display a wide cross-sectional distribution. This systematic disagreement about the stock market's future performance, combined with the high returns to equity, suggests that heterogeneous stock market beliefs could be a powerful driver of wealth concentration.

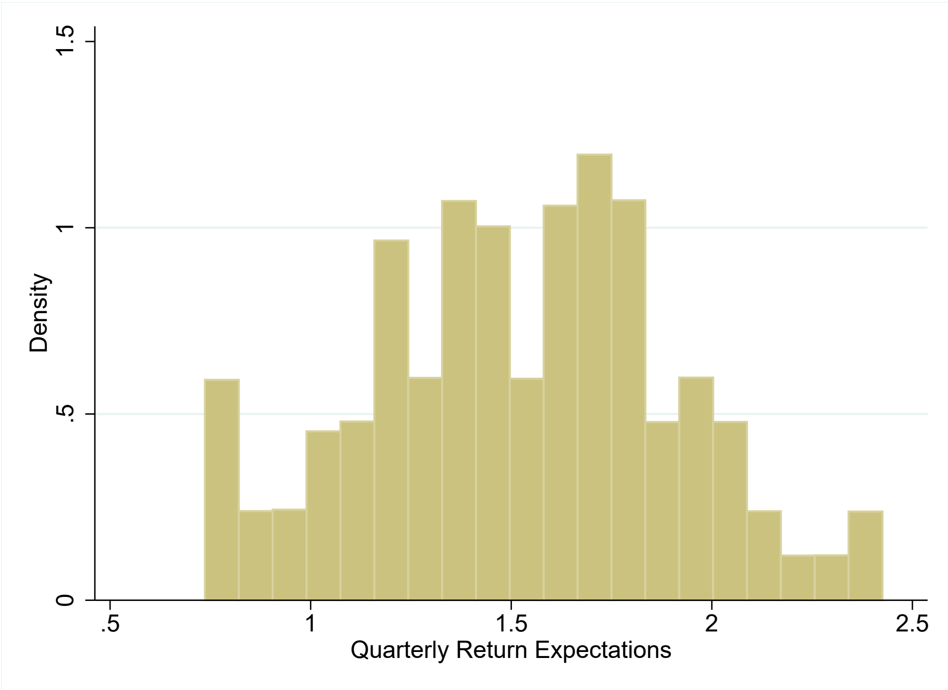


Figure 1: The Distribution of Expected Returns. *The histogram plots the quarterly stock return expectations elicited from all respondents in all the waves 2008-2017 of the RAND - American Life Panel dataset.*

We begin by documenting key facts about the cross-sectional distribution of stock return expectations. Our main dataset is the RAND–American Life Panel, a representative panel of U.S. households from 2008 to 2017 with detailed information on wealth and subjective return expectations. We complement RAND with the Survey of Consumer Finances (SCF) and an integrated RAND–SCF dataset to characterize the joint distribution of expected returns, portfolio composition, and wealth across the entire distribution, including the very top.

RAND data display familiar deviations from full-information rational expectations (FIRE). First, forecast errors are predictable in the time series: respondents extrapolate from recent realizations—high current returns lead to overly optimistic forecasts—mirroring the patterns documented by Adam, Marcet, and Beutel (2017) for capital gains and Kohlhas and Walther (2021) for output growth, and they overreact to new information in line with Bordalo et al. (2020). Second, as in Giglio et al. (2021), individual fixed effects dominate panel variation in beliefs: optimists remain optimistic and pessimists remain pessimistic, generating persistent disagreement about stock market performance.

While forecast-error predictability has been extensively studied, the dominance of individual fixed effects remains unexplained. Giglio et al. (2021) find these fixed effects orthogonal to observables –including wealth and own past returns– in their Vanguard investor panel. Using the representative RAND panel, we find otherwise: beliefs—and the individual-specific component that summarizes them—are strongly related to wealth.[5] The relationship is monotonically positive: Low-wealth households systematically expect lower returns than high-wealth households—a 400 basis point gap between the bottom 50% and top 1%. Forecast errors show the reverse pattern, declining with wealth, so the rich hold more accurate expectations. Combined with SCF evidence that portfolios tilt increasingly toward equity at higher wealth levels, this reveals that wealthier households are both more exposed to and better informed about stock market performance.

We propose a model of expectations formation that accounts for these patterns. The basic setup builds on the notion of Internal Rationality in Adam and Marcet (2011): agents make optimal decisions given their subjective beliefs, but do not possess full knowledge of the underlying market environment. They treat asset prices as a non-degenerate stochastic process and form beliefs by filtering the observed price sequence. Under imperfect knowledge, the subjective law of motion for prices becomes a primitive of the model, needed to complete the probability measure used in optimization. We use this primitive to microfound the positive association between beliefs, portfolio composition, and wealth.

We begin with a benchmark where agents hold heterogeneous long-run views about fundamental asset values—primitive parameters in their subjective price process. Under optimal forecasting, expected returns become weighted averages of past market data and each agent's long-run view, naturally generating dominant individual fixed effects: beliefs fluctuate around agent-specific constants without

[5]This difference likely reflects sample composition: the Vanguard panel consists of active investors, while RAND–ALP represents the full U.S. population, revealing systematic patterns across the entire wealth distribution.

converging.

Taken literally, however, this benchmark implies a theory of inequality in which the rich are rich because they were born optimistic. While this is a logical possibility, the strong correlation between beliefs and wealth in the data is at least as consistent with the reverse causality: agents may be optimistic because they have grown rich. Motivated by this, we therefore extend the framework to make optimism endogenous, shaped by investment experiences rather than exogenous initial views.

Drawing on the literature showing that agents overweight personally experienced outcomes when forming expectations (e.g., Malmendier and Nagel (2011), Kuchler and Zafar (2019), D'Acunto et al. (2021)), we assume investors extrapolate from their own portfolio returns to market expectations. All households observe the same aggregate history but experience different returns based on their portfolios. Investors treat their own portfolio performance as a signal about the underlying stock-return process, so those heavily invested in equity during booms become durably optimistic and maintain high equity shares, while those with limited stock exposure experience lower returns and remain pessimistic and underinvested. This creates a feedback loop between wealth and beliefs. Using UBS–Gallup data, we verify that past portfolio returns strongly predict expected market returns even controlling for demographics and wealth—patterns consistent with our mechanism but not with age-based experience models.

We embed this expectation formation model in a heterogeneous agent framework in the Aiyagari-Bewley-Huggett tradition, with idiosyncratic wage risk, incomplete markets, and aggregate risk as in Krusell and Smith (1998), but following Fernández-Villaverde and Levintal (2024) we include portfolio choice while abstracting from production. Our calibrated model matches the joint distribution of income and expectations and endogenously generates wealth and return distributions. We find that belief heterogeneity substantially amplifies wealth concentration: optimistic households maintain persistently higher equity shares, compounding their wealth advantage over time. Relative to homogeneous expectations, heterogeneous beliefs increase the top 10% wealth share by 20% and the top 1% share by 50%.

While heterogeneous expectations might seem to add complexity to already demanding heterogeneous-agent models, Internal Rationality actually simplifies computation. Rather than forecasting entire future distributions of wealth and income to predict prices—the infinite-dimensional problem that Moll (2024) called "nonsensical"—agents simply forecast prices directly using their subjective models. This preserves theoretical discipline while sidestepping computational intractability: agents optimize given a well-specified subjective law of motion for prices rather than implicitly forecasting it via forecast of the entire cross-sectional distribution of states.[6] Thus we match micro-level heterogeneity in beliefs and portfolios while achieving greater tractability than standard rational expectations approaches (e.g., Fernández-Villaverde and Levintal (2024)).[7]

The rest of the paper is organized as follows. Section 2 documents facts about the joint distribution of expected returns and wealth, using a novel dataset that integrates RAND–ALP with SCF data. Section 3 presents the baseline model and solution method. Section 4 describes the calibration, evaluates the model's fit, and quantifies the impact of heterogeneous beliefs on wealth concentration. Section 5 develops the portfolio-based learning mechanism, provides empirical evidence in its support, and assesses its quantitative role in the model. Section 6 concludes.

2. Facts about Expected Returns and Wealth

Our analysis uses data from the RAND American Life Panel (ALP), a regular household survey with 61 waves conducted between 2008 and 2017. The RAND panel is uniquely suited for studying the relationship between wealth and expectations, as it is the only public panel dataset that simultaneously tracks both household wealth and return expectations. The unbalanced panel includes approximately 3,000 participants, providing a comprehensive view of household belief formation over time.

To elicit return expectations, the survey asks participants three questions about their beliefs regarding the one-year-ahead performance of mutual funds invested in blue-chip stocks. Specifically, respondents provide probability assessments for: (1) positive returns of any magnitude, (2) returns exceeding 20 percent, and (3) losses exceeding 20 percent.[8] Let p^i_+, $p^i_{>20}$, and $p^i_{<-20}$ denote respondent i's answers to these three questions (in probability units). These answers imply a subjective distribution over four mutually exclusive return ranges for the gross stock-market return R_{t+1} $P^i(R_{t+1} < 0.8) = p^i_{<-20}$, $P^i(0.8 < R_{t+1} < 1) = 1 - p^i_+ - p^i_{<-20}$,

[6] Moll (2024) argues that standard heterogeneous-agent models with aggregate risk force households to forecast infinite-dimensional objects and calls for alternative formulations. Our framework provides one such alternative.

[7] We hope this computational simplicity will broaden the use of rich heterogeneous-agent models and open new research avenues.

[8] The exact framing of one question is (the others are similar): *By next year at this time, what are the chances that mutual fund shares invested in blue-chip stocks like those in the Dow Jones Industrial Average will have increased in value by more than 20 percent compared to what they are worth today?*

$P^i(1 < R_{t+1} < 1.2) = p^i_+ - p^i_{>20}$, $P^i(1.2 < R_{t+1}) = p^i_{>20}$.

For instance, a response pattern of $(p^i_+, p^i_{>20}, p^i_{<-20}) = (0.65, 0.15, 0.10)$ implies a 10% probability of returns below −20%, a 25% probability of returns between −20% and 0%, a 50% probability of returns between 0% and 20% and a 15% probability of returns above 20%. We exclude observations that violate these consistency conditions (e.g., imply negative probabilities for any of the four ranges).

To convert this coarse distribution into an expected return, we follow Gaudecker and Wogrolly (2022) and assign each bin a representative gross return x_j, collected in the vector $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$. For each individual i, we then compute the subjective expected gross return as

$$\begin{aligned} \mathbb{E}_t^{\mathcal{S}^i}(R_{t+1}) = P^i(R_{t+1} < 0.8)\, x_1 + P^i(0.8 < R_{t+1} < 1)\, x_2 \\ + P^i(1 < R_{t+1} < 1.2)\, x_3 + P^i(1.2 < R_{t+1})\, x_4, \end{aligned} \tag{1}$$

where the values in $\boldsymbol{x}$ are chosen so that, when aggregating across respondents, the cross-sectional distribution of $\mathbb{E}_t^{\mathcal{S}^i}(R_{t+1})$ matches the historical mean and variance of stock-market returns.

To ensure sufficient time-series variation for our analysis of belief persistence, we restrict our sample to participants with at least 10 successive quarterly observations. This filtering reduces our sample to 2,321 households, though our results are robust to alternative minimum observation thresholds. Table 1 illustrates the sample size under different minimum observation requirements, ranging from 3,027 households with at least three observations to 303 households with fifty or more observations.

Table 1: *Sample Size by Minimum Number of Successive Observations. This table shows the number of households in our sample that meet different minimum observation requirements. Data from RAND American Life Panel, 2008-2017.*

Minimum Observations	3	4	5	6	10	30	50
Number of Households	3,027	2,969	2,882	2,764	2,321	785	303

We proceed with this data in two stages. First, we conduct tests for three main deviations from full information rational expectations to understand the key features of the survey across time and individuals. These tests enable us to benchmark our survey against existing ones; observing similar patterns across surveys strengthens confidence in both survey methodology generally and our data specifically. Second, we use the data to characterize the joint distribution of beliefs and wealth.

2.1. Facts about subjective expected returns

2.1.1. Forecast Error Predictability

A robust empirical regularity documented across diverse surveys and economic variables is that agents' forecast errors exhibit systematic predictability (e.g., Coibion and Gorodnichenko (2015), Greenwood and Shleifer (2014), Adam, Marcet, and Beutel (2017), Bordalo et al. (2020), Kohlhas and Walther (2021)). This pattern directly contradicts the Full Information Rational Expectations (FIRE) hypothesis, under which forecast errors should be orthogonal to any information available at time t. Consistent with this broader literature, our analysis of the RAND American Life Panel reveals similarly predictable forecast errors.

Following Kohlhas and Walther (2021), we estimate two regressions. The first one is a projection of forecast errors on current returns:

$$R_{t+k} - \mathbb{E}_t^{\mathcal{S}^i}(R_{t+k}) = \alpha^i + bR_t + \varepsilon_t^i \tag{2}$$

where R_t are stock returns at time t, $\mathbb{E}_t^{\mathcal{S}^i}(R_{t+k})$ represents household i's return expectation k periods ahead at time t, α^i is an individual fixed effect and ε_t^i is an error term. The left hand side variable shows the forecast error of agent i at time t; if it is negative, it means agent i was too optimistic at time t.

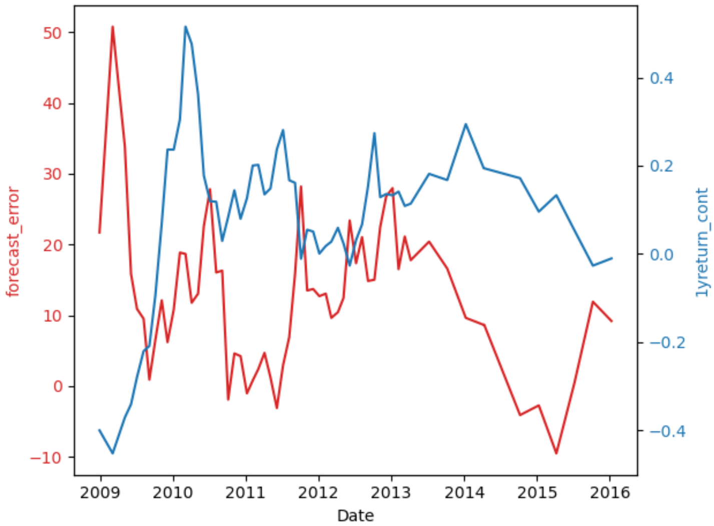
The second one is a projection of forecast errors on average forecast revisions:

$$R_{t+k} - \mathbb{E}_t^{\mathcal{S}^i}(R_{t+k}) = \alpha^i + d(\bar{\mathbb{E}}_t^{\mathcal{S}}(R_{t+k}) - \bar{\mathbb{E}}_{t-1}^{\mathcal{S}}(R_{t+k})) + \varepsilon_t^i \tag{3}$$

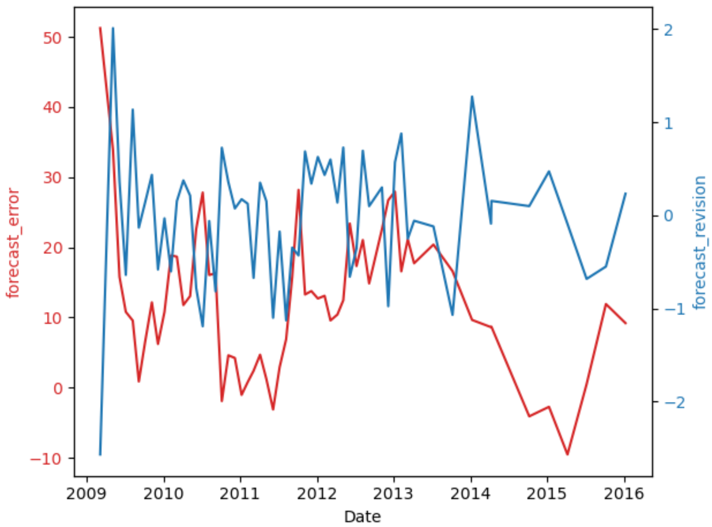
where $\bar{\mathbb{E}}_t^{\mathcal{S}}(R_{t+k})$ is the average forecast across respondents and $\bar{\mathbb{E}}_t^{\mathcal{S}}(R_{t+k}) - \bar{\mathbb{E}}_{t-1}^{\mathcal{S}}(R_{t+k})$ is the average change in forecasts about the same object in subsequent quarters. A negative revision possibly signals bad news released between $t-1$ and t.

The prediction of FIRE is that the coefficients b and d in (2)-(3) should be zero if returns and revisions are observed at time t. That is the null hypothesis of the joint test of full information and rationality.

The raw data suggest deviations from this prediction. Figures 7a and 7b show the comovements of forecast errors, forecast revisions, and realized returns. Specifically, figure 7a plots the average one-year ahead forecast errors and the realized return. A pattern can be grasped: negative returns, as in 2009, lead to overly pessimistic beliefs and large positive forecast errors; positive returns, as 2014, tend to occur with optimistic beliefs and negative forecast errors. Figure 7b also suggests some negative association, although visually less clear, that might point toward over-reaction.



(a) *Realized return and average forecast error in subsequent prediction.*



(b) *Average forecast error and forecast revision.*

Figure 2: *Average relationships between returns, forecast errors, and forecast revisions.*

Table 2 confirms this intuition using both individual and average forecast errors as dependent variables in equations (2)-(3). At the individual level, we estimate $\hat{b} = -0.13$ (statistically significant), indicating that agents extrapolate from past returns when forming expectations—a pattern of overreaction consistent with Adam, Marcet, and Beutel (2017). When using average forecast errors, the coefficient remains negative ($\hat{b} = 0.08$) but loses statistical precision, suggesting heterogeneity in individual forecasting behavior that attenuates in the aggregate.

Turning to forecast revisions in equation (3), we find strong evidence of over-

*Table 2: **Deviations from Full Information Rational Expectations.** The top panel reports the estimation of the slope (b, d) in equations (2)-(3). The two subpanels differ in the left-hand side variable: the top subpanel uses individual forecast errors $R_{t+k} - \mathbb{E}_t^{\mathcal{S}^i}(R_{t+k})$; the bottom subpanel uses an average forecast error $R_{t+k} - \bar{\mathbb{E}}_t^{\mathcal{S}^i}(R_{t+k})$. Standard errors are double-clustered and robust. We use $k = 4$. $***$ is p-value<0.01. The bottom panel reports the coefficient of determination of regression (4) when including only time fixed effects, individual fixed effects and both.*

	1. Forecast Error Predictability		
Individual Forecast Error	(1)	(2)	(3)
R_t	-0.13***	-	-0.14***
$\bar{\mathbb{E}}_t^{\mathcal{S}}(R_{t+k}) - \bar{\mathbb{E}}_{t-1}^{\mathcal{S}}(R_{t+k})$	-	-0.50***	-0.51***
R^2	0.05	0.11	0.07
F-Stat	2757	8368	2020
Obs.	64749	64749	64749
Average Forecast Error			
R_t	-0.18	-	-0.18
$\bar{\mathbb{E}}_t^{\mathcal{S}}(R_{t+k}) - \bar{\mathbb{E}}_{t-1}^{\mathcal{S}}(R_{t+k})$		-3.22	-3.20
R^2	0.1	0.05	0.14
F-Stat	2.1	0.56	0.99
Obs.	59	59	59
	2. Persistent Disagreement		
Fixed Effects	μ_t	α^i	μ_t, α^i
R^2	0.01	0.50	0.52

reaction to new information, with $\hat{d} = -0.50$ (significant at the 1% level). This result aligns with Bordalo et al. (2020), who document similar overreaction patterns, though notably we find this effect persists even when using average forecast errors on the right-hand side. This contrasts with Kohlhas and Walther (2021), who find under-reaction when using average forecast errors, which suggests that reaction patterns are largely variable-specific rather than a universal phenomenon across economic domains.

2.1.2. Persistent Heterogeneity

A second fundamental departure from FIRE emerges in the cross-section of beliefs rather than in time-series patterns. Under rational expectations with full information, all agents should form identical forecasts; yet substantial and persistent disagreement is a defining feature of survey expectations (Mankiw, Reis, and Wolfers

(2003), Patton and Timmermann (2010), Coibion and Gorodnichenko (2012), Giglio et al. (2021)). This heterogeneity cannot be explained by differential information alone, as disagreement persists even among professional forecasters with similar information access (Andrade et al. (2016)).

To quantify the structure of belief heterogeneity, we follow Giglio et al. (2021) and decompose individual expectations using the following specification:

$$\mathbb{E}_t^{\mathcal{S}^i}(R_{t+k}) = \alpha^i + \mu_t + v_t^i \tag{4}$$

where α^i represents individual fixed effects capturing persistent differences in optimism or pessimism across agents, μ_t denotes time fixed effects absorbing common time-series variation in beliefs, and ε_t^i is an idiosyncratic component.

Table 2 reports the results from three specifications. First, time fixed effects alone explain merely 1% of the total variation in beliefs, indicating minimal common movement in expectations. Second, individual fixed effects alone capture 50% of the variation, revealing substantial persistent heterogeneity across agents. Third, the full specification with both sets of fixed effects yields an R^2 of 0.52, with the modest increase from 0.50 to 0.52 confirming that individual heterogeneity and time-series variation are largely orthogonal dimensions of belief dispersion. These findings underscore that cross-sectional heterogeneity, rather than common time variation, constitutes the dominant source of disagreement in return expectations.

To further investigate the persistence of belief heterogeneity, we implement an alternative classification scheme that tracks agents' relative optimism over time. We assign each respondent to a quintile based on their initial expected return forecast when they first enter the survey, then track the average beliefs within these fixed cohorts over the subsequent periods. If individual beliefs were primarily driven by transitory shocks or common time-series factors, we would expect substantial mean reversion—the initially optimistic and pessimistic cohorts would converge toward similar expectations over time, causing quintile means to overlap. Instead, Figure 3 reveals a pattern of persistent separation: respondents classified as optimistic (pessimistic) in their first survey response remain systematically more optimistic (pessimistic) throughout the sample period. The quintile means maintain their rank ordering without crossing, with the spread between the most optimistic (quintile 5) and most pessimistic (quintile 1) cohorts remaining approximately 10 percentage points over seven years. While this might be explained by the short time series dimensions of the survey, the pattern is robust to the exclusion of investors with less

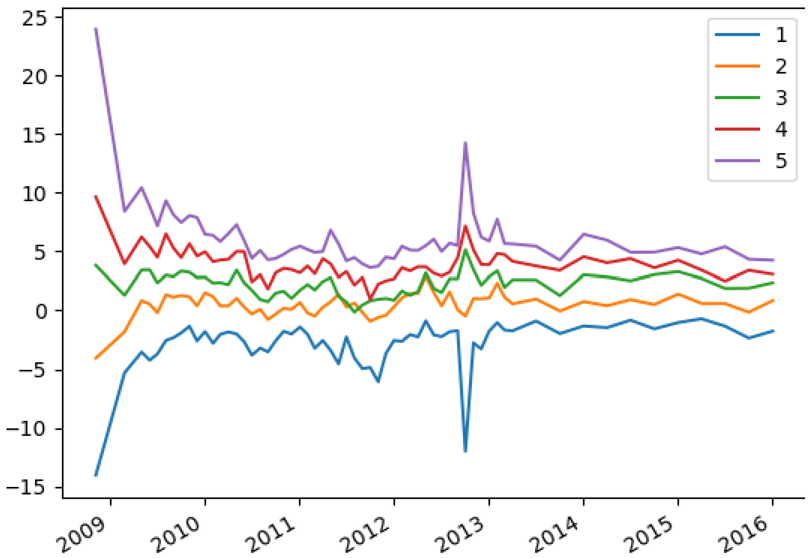


Figure 3: *Mean expected return by quintiles of participants, according to their first answer in the dataset.*

observations.

Moreover, we construct transition probability matrices that track movements across the belief distribution. Each respondent is assigned to a quintile based on their initial expected return forecast, and we then calculate the probability of transitioning to each quintile in subsequent survey waves. Figure 4 presents these transition probabilities, revealing remarkable stability in relative beliefs. The diagonal elements—representing the probability of remaining in one's initial quintile—range from 0.76 to 0.90, with particularly strong persistence at the extremes of the distribution (0.81 for the most pessimistic and 0.90 for the most optimistic quintile). Off-diagonal transitions are rare and, when they occur, typically involve movement to adjacent quintiles rather than dramatic reversals in relative optimism. Notably, the probability of transitioning from the bottom to the top quintile (or vice versa) is essentially zero, never exceeding 0.004. This near-absorbing state property of initial beliefs provides further evidence that cross-sectional heterogeneity reflects deep-rooted differences in how individuals form expectations.

2.2. Facts about the distribution of wealth

We use the 2016 vintage of the US Survey of Consumer Finance (SCF) for information on the wealth distribution, as it overlaps with our latest data from RAND-ALP. The SCF provides detailed information on household balance sheets, making it par-

	1	2	3	4	5
1	0.81	0.13	0.05	0.006	0.004
2	0.12	0.77	0.1	0.009	0.001
3	0.09	0.12	0.758	0.019	0.013
4	0	0.01	0.05	0.82	0.12
5	0	0.01	0.02	0.07	0.9

Figure 4: *Transition probabilities relative to participants first determined belief quintile.*

ticularly well-suited for analyzing portfolio heterogeneity.[9]

We employ the standard weighting scheme to ensure representativeness at the population level. *Risky assets* are defined as:

Stock Mutual Funds + 0.5 × Combination Mutual Funds + 0.5 × Other Mutual Funds
+ Corporate and Foreign Bonds + Stocks + 0.5 × Quasi-liquid retirement accounts + Businesses.

The 0.5 weights on combination funds and retirement accounts reflect their partial exposure to equity markets. Relatedly, *real estate assets* are defined as:

Primary Residential Houses + Other Residential Property
+ Net Equity in Non-residential Real Estate

Table 3 summarizes the key variables along the distribution of net worth. Several familiar patterns emerge from these data. First, the distribution of wealth exhibits extreme concentration at the top. While the median household has a net worth of \$120,000, households at the 99th percentile hold \$12.8 million—over 100 times the median. This concentration exceeds that of income, where the 99th percentile earns approximately 10 times the median. The negative net worth at the 10th percentile (-\$1,282) indicates that a substantial fraction of households have liabilities exceeding their assets. Second, the sources of income vary dramatically across the wealth distribution. While wage income dominates for most households, capital income

[9] For a detailed overview about the asset classes in the SCF, see: https://www.federalreserve.gov/econres/files/Networth_Flowchart.pdf

becomes increasingly important at higher wealth levels. At the 99th percentile, capital income averages $375,000, representing nearly half of total income.

Perhaps the most crucial pattern for our analysis concerns the composition of household portfolios. Lower-wealth households hold minimal risky assets—just 2.7% at the 10th percentile and 7.9% at the median. In stark contrast, the wealthiest households are almost entirely invested in risky assets, with the 99th percentile holding 98.7% of their wealth in this form. The participation margin reinforces this pattern: only 16.7% of households at the 10th wealth percentile own any risky assets, compared to near-universal ownership (99.9%) at the 99th percentile.

Real estate exhibits a hump-shaped pattern across the wealth distribution. It represents the dominant asset class around the median (63% of wealth at the 50th percentile) but declines in importance for the very wealthy (24% at the 99th percentile). The high homeownership rate even at relatively low wealth percentiles (88% at the median) contrasts sharply with risky asset ownership, suggesting a fundamental difference between these two asset categories.

Table 3: Wealth and Income Distribution and Portfolio Composition in SCF-2016*. Transfer income includes social security, disability, pensions. Wealth/Income/Labor income shares refers to the fraction over the total aggregate. Risky asset/Real state share refers to the fraction in the portfolio.*

Percentile	**10**	**50**	**90**	**95**	**99**
Net Worth	-1282	120k	1.5M	2.9M	12.8M
Income	34.9k	78.4k	187k	270k	811k
Wage income	23.8k	63.8k	116k	134k	363k
Transfer income	9.5k	9.2k	50.2k	52.1k	71.6k
Capital income	1365	5204	23.9k	76.5k	375k
Wealth share	-0.005%	1.1%	77.1%	65.1%	38.5%
Income share	4.3%	20.7%	44.3%	34.3%	18.4%
Labor income share	6%	26.9%	34%	24.7%	11.7%
Risky Asset Share	2.7%	7.9%	26.4%	39.5%	98.7%
Real Estate Share	11.8%	63%	45%	36%	24%
Owns risky	16.7%	69%	93%	99.7%	99.9%
Homeownership Share	13%	88%	97%	96%	99%

Figure 5 provides a visual representation of these portfolio patterns, illustrating how the transition from real estate-dominated portfolios in the middle of the distribution to risky asset-dominated portfolios at the top creates distinct wealth accumulation dynamics across household types.

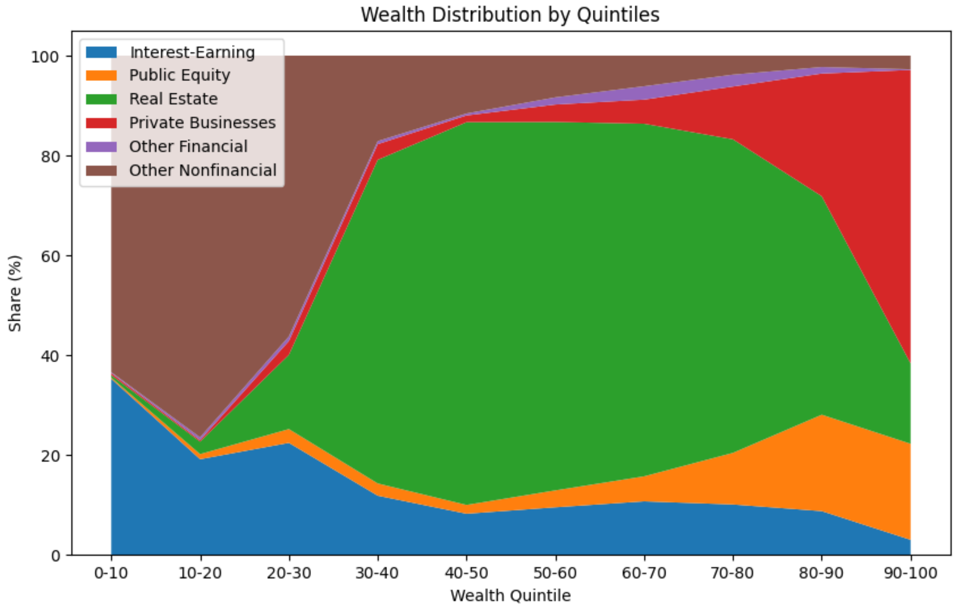


Figure 5: *Asset Holdings by Wealth Percentile*

2.3. The joint distribution of expectations and wealth

2.3.1. A new integrated dataset: RAND - SCF

The persistent heterogeneity in beliefs documented above raises a natural question: how do these differences in expectations translate into actual portfolio choices and wealth accumulation? Addressing this question requires data that combines return expectations with comprehensive portfolio holdings.

While the RAND dataset offers some information about the wealth of the respondents, it suffers from a critical limitation for studying the beliefs-portfolio nexus: it contains only housing wealth and retirement accounts, completely omitting directly-held equities, bonds, and other financial assets. This is particularly problematic because a key implication of heterogeneous beliefs models is that optimistic households should hold larger equity shares in their portfolios. Without observing direct equity holdings—which constitute a substantial fraction of financial wealth for stock market participants—testing whether beliefs actually translate into portfolio tilts is not possible. Furthermore, even for the retirement accounts that RAND does observe, we lack crucial details about asset allocation within these accounts (the equity-bond mix), preventing us from examining how beliefs shape portfolio composition even within observed wealth components.

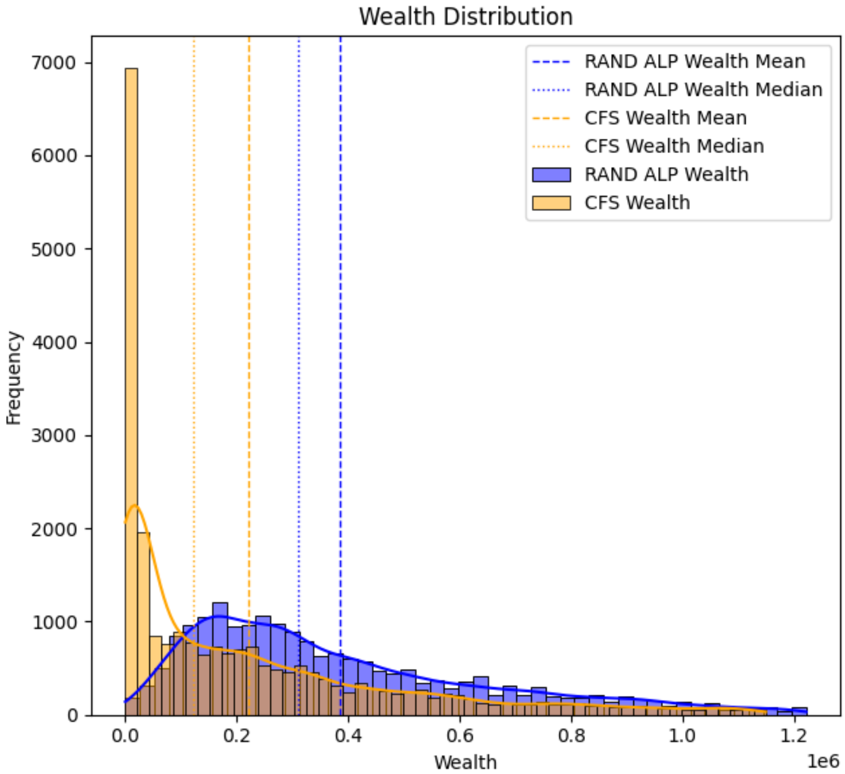


Figure 6: *The Distributions of Wealth in RAND-American Life Panel dataset and the Survey of Consumer Finances. Year 2016.*

To overcome these shortcomings, we use the wealth information in RAND to match it with the Survey of Consumer Finances, which provides comprehensive data on all asset categories including direct equity holdings, mutual funds, bonds, and the composition within retirement accounts, but lacks information on return expectations. The validity of our matching approach hinges on RAND being sufficiently representative of stock market participants. Figure 6 confirms this by comparing the wealth distributions from RAND and SCF in the common categories (housing plus retirement accounts).

While RAND substantially undersamples households below the 33rd percentile, it provides good coverage of middle- and high-wealth households. Crucially, RAND captures the right tail of the wealth distribution reasonably well, validating our approach of imputing beliefs for wealthy households who drive equity demand.

Our approach exploits variables common to both RAND and SCF to impute return expectations for SCF households. The key identifying assumption is that households with identical observable characteristics form similar return expectations—that is, conditional on observables, belief heterogeneity is unsystematic.

To match the surveys, we first identify variables available in both datasets that could predict return expectations: retirement account value (r_i), home value (h_i), homeownership status ($\mathbf{1}_{owner,i}$), wage income (y_i^{wage}), and age (age_i). These variables capture both wealth composition and demographic characteristics that may shape belief formation. Then, using the RAND sample, we estimate the following multivariate regression:

$$\mathbb{E}_t^{\mathcal{S}^i}[R_{t+1}] = \alpha + \beta_1 \mathbf{1}_{owner,i} + \beta_2 \log(r_i) + \beta_3 \log(h_i) + \beta_4 \log(y_i^{wage}) + \beta_5 age_i + \epsilon_i \quad (5)$$

Table 4 reports the results. Wage income emerges as the strongest predictor ($\hat{\beta}_4 = 1.5$, $p < 0.01$), followed by retirement account value ($\hat{\beta}_2 = 0.3$, $p < 0.01$). Notably, homeownership and home value have minimal predictive power, while age shows no significant relationship with expectations. The log specification captures the non-linear relationship between financial variables and beliefs while maintaining tractability. Finally, we apply the estimated coefficients to SCF households to generate imputed expectations $\widehat{\mathbb{E}}_t^{\mathcal{S}^j}[R_{t+1}]$ for each household j.

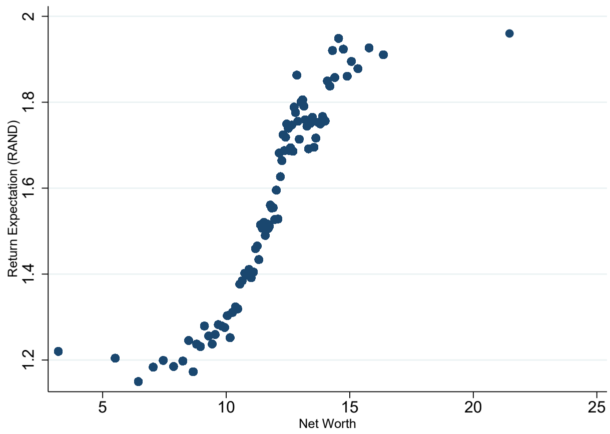
Table 4: Demographics and expected returns. *The table reports the coefficients of equation (5) using expected returns and forecast errors in the left hand side, respectively.* $***$ *signals p-val* $<$ *0.001.*

Predictor	Expected Returns	Forecast Error
Constant	-14.5^{***}	7.5^{***}
Owns home?	-1.9	0.4
Retirement Account	0.3^{***}	-0.1^{***}
Home Value	0.1	0.01
Wage Income	1.5^{***}	-0.4^{***}
Age	0.02	-0.01

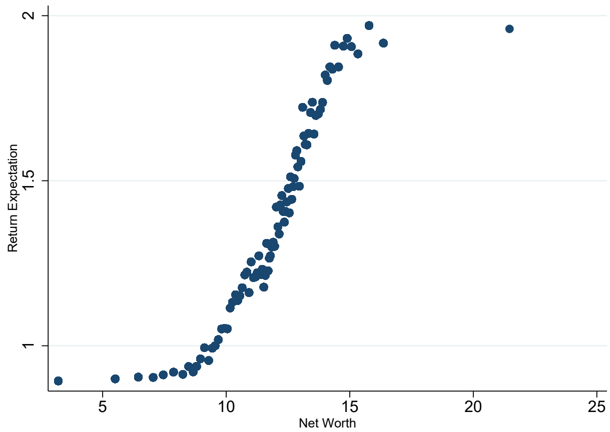
This procedure generates a dataset combining SCF's comprehensive wealth and portfolio data with imputed return expectations, enabling us to examine how beliefs translate into portfolio allocations across the entire wealth distribution.

2.3.2. Expectations and Wealth

We document how expected returns vary systematically with wealth. Figure 7 presents a binned scatterplot with log wealth on the x-axis and quarterly expected returns on the y-axis, where each point represents a percentile mean. A striking pattern emerges: wealth and beliefs are strongly positively correlated. Poor households expect annualized returns of approximately 4%, middle-wealth households expect 6%, while the wealthy expect 8%—a 4 percentage point gap between bottom and top. This pattern holds equally in both datasets, the raw RAND data and the integrated RAND-SCF.



(a) RAND-American Life Panel dataset.



(b) Integrated RAND-SCF dataset.

Figure 7: Binscatter plot: Expected (quarterly) Returns and log Wealth.

The binned scatterplot masks substantial within-percentile variation, as households at similar wealth levels hold diverse beliefs about market returns. While binning reveals a clearer visual pattern, the strong wealth-belief association is not an artifact of this averaging procedure. Figure 8 displays all individual observations from the integrated dataset, showing that the relationship persists at the household level. The correlation coefficient drops from 0.95 in the binned data to 0.67 in the raw data—still remarkably high for cross-sectional survey evidence on expectations.

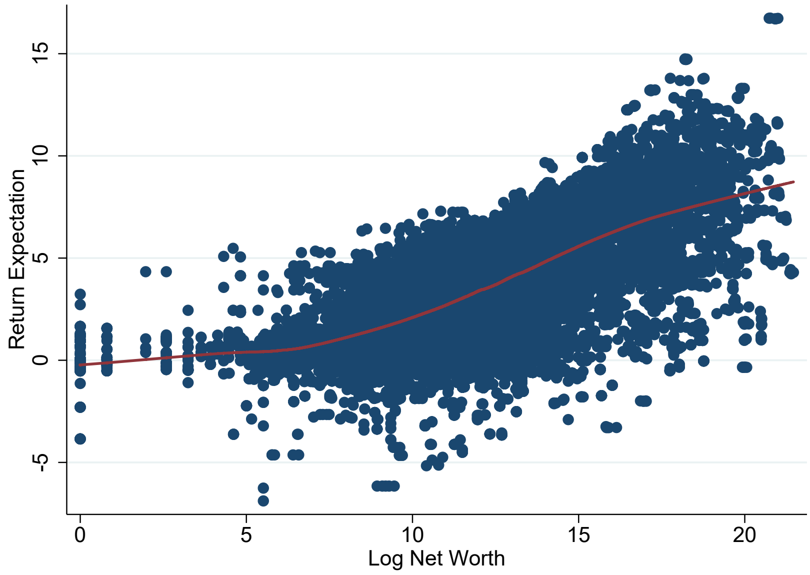


Figure 8: Scatter plot: Expected (annualized) returns and log Wealth in the RAND-SCF integrated dataset.

The previous analysis establishes that wealthy households are more optimistic than poor households. But which group's expectations are more accurate? To answer this, we compute forecast errors by comparing expected returns to the S&P 500's average annual nominal total return. Figure 9 reveals a striking pattern: forecast errors decline monotonically with wealth. While wealthy households' expectations align closely with realized returns, poor households are systematically too pessimistic. The correlation between forecast errors and wealth is -0.60 in the raw data and -0.92 when binned (not shown).

To summarize, RAND expectations exhibit the same deviations from Full Information Rational Expectations found in other surveys: predictable forecast errors and large, persistent disagreement. We also document a strong association between expectations and wealth. Wealthier households hold more risky assets and are more optimistic about returns—and crucially, their optimism is justified by actual market performance. The next section builds a model to replicate these facts and explore their consequences for wealth inequality.

3. A HA Model with Heterogeneous Expectations

The model combines several key features from the heterogeneous agent literature. Following Aiyagari-Bewley-Huggett, we incorporate idiosyncratic wage risk and in-

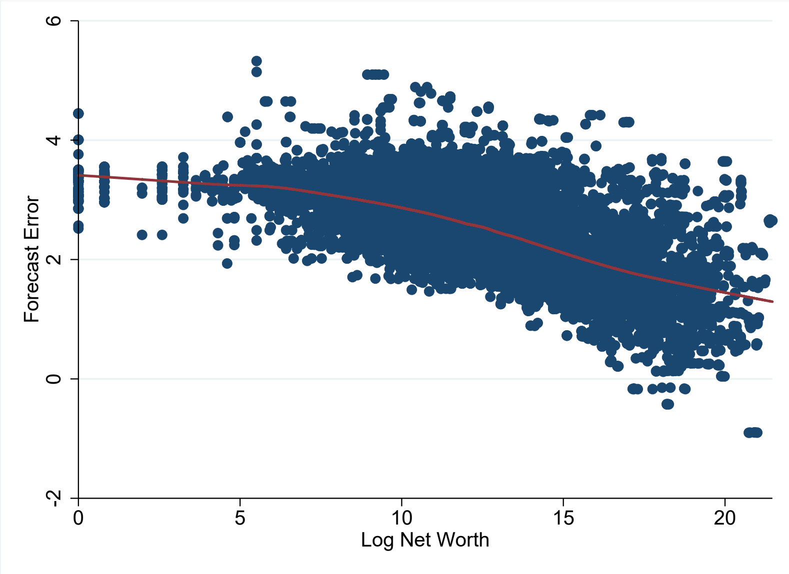


Figure 9: Correlation between wealth and (annualized) forecast errors in the integrated RAND-SCF dataset.

complete markets, preventing households from perfectly insuring against income shocks. We include aggregate risk as in Krusell and Smith (1998), allowing for business cycle fluctuations that affect all agents simultaneously. Rather than modeling production explicitly, we calibrate the wage distribution directly to match US data, simplifying the analysis while maintaining empirical relevance. However, we introduce a portfolio choice problem, allowing households to choose between riskless bonds and risky equity.

Crucially, we depart from rational expectations by adopting Internal Rationality à la Adam and Marcet (2011): agents learn about asset prices directly rather than forming beliefs about the entire stochastic process of the economy. This approach, combined with heterogeneous parameters in agents' subjective models of returns, generates the persistent belief dispersion we observe in the data. For computational tractability, we assume a small open economy where safe interest rate R is constant and determined exogenously, though this assumption is not critical for our main results.

3.1. Model structure

Household problem. Consider an endowment economy with I groups of agents, each of mass μ^i. Agents solve a consumption-savings and portfolio choice problem:

$$\max_{\{C_t^i, S_t^i, B_t^i\}_{t=0}^{\infty}} \mathbb{E}_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t \frac{(C_t^i)^{1-\gamma}}{1-\gamma} \tag{6}$$

subject to the budget constraint

$$C_t^i + P_t S_t^i + B_t^i \leq W_t^i + (D_t + P_t) S_{t-1}^i + R B_{t-1}^i \tag{7}$$

where C_t^i is consumption, S_t^i denotes equity holdings, B_t^i represents bond holdings, W_t^i is wage income, P_t is the equity price, D_t is the dividend, and R is the gross bond return. Agents face position limits:

$$\underline{S} \leq S_t^i \leq \bar{S}, \quad \underline{B} \leq B_t^i \leq \bar{B} \tag{8}$$

Initial endowments are $S_0^i = 1$ and $B_0^i = 0$. The upper bounds ensure a compact feasible space but are set sufficiently large to remain non-binding in equilibrium.

Income processes. This is a small open endowment economy. Aggregate wages are defined in terms of the wage-dividend ratio, with i.i.d. growth around a mean $1 + \overline{WD}$

$$ln\Big(1 + \frac{W_t}{D_t}\Big) = (1-p) ln(1 + \overline{WD}) + p\; ln\Big(1 + \frac{W_{t-1}}{D_{t-1}}\Big) + ln\varepsilon_t^W \tag{9}$$

Dividends growth is i.i.d., fluctuating around an average growth rate a

$$lnD_t = lnD_{t-1} + a + +ln\varepsilon_t^D \tag{10}$$

Aggregate shocks are correlated, reflecting underlying macroeconomic factors that affect the various sources of aggregate income. They follow a multinormal distribution with covariance σ_{DW}

$$\begin{pmatrix} ln\varepsilon_t^D \\ ln\varepsilon_t^W \end{pmatrix} \sim \mathcal{N}\left(-\frac{1}{2}\begin{pmatrix} \sigma_D^2 \\ \sigma_W^2 \end{pmatrix}, \begin{pmatrix} \sigma_D^2 & \sigma_{DW} \\ \sigma_{DW} & \sigma_W^2 \end{pmatrix}\right), \tag{11}$$

People can borrow goods at an exogenous rate R, subject to their borrowing constraints. At rate R, international investors are always willing to lend/borrow from domestic households, as a counterpart for goods imports/exports.

Individual wages are exogenous, expressed as a random share $\tilde{w}$ of aggregate ones and affected by an idiosyncratic income shock $\nu_t^i \sim \mathcal{N}(0, \sigma_\nu^2)$

$$\frac{W_t^i}{D_t} = \tilde{w}^i \frac{W_t}{D_t} \nu_t^i \tag{12}$$

We require $\sum_i \tilde{w}^i = 1$. To that end, let

$$\tilde{w}^i = \frac{w^i}{\sum_j w^j} \tag{13}$$

w^i is simulated from a Mixed Lognormal-Pareto distribution such that

$$w^i \sim \begin{cases} log\mathcal{N}(\mu, \sigma_w^2) & \text{if } i < \underline{w} \\ Pareto(\underline{w}, \alpha) & \text{if } i \geq \underline{w} \end{cases} \tag{14}$$

As in Piketty and Saez (2003), this mixture captures in a parsimonious way the approximately log-Normal shape of the real-world income distribution, while correcting for the particularly right heavy tail with a Pareto distribution.

Agents' Belief System. Agents are endowed with perfect knowledge of the law of motions for dividends and wages given by equations (9)-(10)-(11)-(12)-(13). However, agents have only imperfect knowledge about price formation. Thus, the underlying probability space $(\boldsymbol{\Omega}, \mathcal{B}, \mathcal{P}^i)$ with a typical element $\omega \in \boldsymbol{\Omega}$ with $\omega = \{D_t, W_t^i, P_t\}_{t=0}^{\infty}$. In our model, agents treat prices as another exogenous stochastic process rather than as an equilibrium object as in Adam and Marcet (2011). This imperfect knowledge requires specifying agents' subjective model of prices to complete the probability measure they use for optimization.

We conjecture a state-space model of prices with different layers of heterogeneity. Investors from the sentiment group i possess the following subjective model about stock prices

$$\begin{aligned} lnP_t &= lnP_{t-1} + lnb_t^i + ln\varepsilon_t^{P,i} \\ lnb_t^i &= (1-\rho^i) ln\bar{\beta}^i + \rho_i lnb_{t-1}^i + ln\zeta_t^i \\ ln\varepsilon_t^{P,i} &\sim i.i.\mathcal{N}(-\frac{\sigma_P^2}{2}, \sigma_P^2),\ ln\zeta_t^i \sim i.i.\mathcal{N}(-\frac{\sigma_\zeta^2}{2}, \sigma_\zeta^2) \end{aligned} \tag{15}$$

where b_t^i represents the permanent price growth component, $\varepsilon_t^{P,i}$ a transitory innovation to prices and ζ_t^i an innovation to the permanent component of returns.[10] The permanent component, b_t^i, follows an auto-regressive process with persistence ρ^i and mean $\bar{\beta}^i$. The latter represents the perceived long-term return of sentiment group i and captures the stylized fact that we document in Section 2 related to the presence of fixed effects in the cross section of the distribution of survey expectations. The permanent component of price growth, b_t, is not observed and is optimally estimated using the information available from the price signals. Given their belief system from equation 15, the optimal posterior distribution of the permanent component of prices is

$$ln b_t^i \mid \omega^t \sim \mathcal{N}(ln\beta_t^i, (\sigma^i)^2) \tag{17}$$

where $(\sigma^i)^2$ is the steady state variance of the posterior given by

$$(\sigma^i)^2 = \frac{(\sigma_\zeta^i)^2 + \sqrt{((\sigma_\zeta^i)^4 + 4(\sigma_\zeta^i)^2\sigma_R^2}}{2} \tag{18}$$

β_t^i is the conditional mean, which evolves according to the Kalman updating equation

$$ln\beta_t^i = (1-\rho^i)(1-g^i)ln\bar{\beta}^i + \rho^i ln\beta_{t-1}^i + g^i\left(ln\frac{P_{t-1}}{P_{t-2}} - \rho^i ln\beta_{t-1}^i\right) + g^i\varepsilon_t^{P1,i} \tag{19}$$

where $g^i = \frac{(\sigma^i)^2}{(\sigma^i)^2+\sigma_R^2}$ represents the steady state Kalman gain, entailing different views on the signal-to-noise ratio of the price signals. The shock $\varepsilon_t^{P1,i}$ is an information shock to the beliefs of agents from group i. Altogether,

$$\mathbb{E}_t^{\mathcal{P}^i}\left[\frac{P_{t+1}}{P_t}\right] = \kappa^i\bar{\beta}^{i(1-\rho)}(\beta_t^i)^\rho \tag{20}$$

Qualitatively, equation (19) contains elements that might replicate the key observations from surveys: the heterogeneous long-run views about the fundamental value of the asset can be linked to the individual fixed-effects and the perpetual disagreement; the different views about the signal-to-noise ratio of the price signals can

[10] The noisy price component is comprised of two independent zero-mean normal components

$$\varepsilon_t^{R,i} = \varepsilon_{t+1}^{R1,i} + \varepsilon_t^{R2,i} \tag{16}$$

and we assume as in Adam, Marcet, and Beutel (2017), that only $ln\varepsilon_t^{R1,i}$ is observed at time t, giving rise to the lag-updating equation that is usually found in the learning literature.

lead to different degrees of extrapolation; the persistence parameter can be directly linked to the persistence from surveys.

Based on the empirical evidence from Section 2, we set $\rho^i = \rho$ for all agents. Three parameters capture heterogeneity across agents in income and expectations: the Kalman gain g, the long-run capital gains belief $\bar{\beta}$, and the individual wage share w. Let $\boldsymbol{g} = [\underline{g}, \bar{g}]$ and $\bar{\boldsymbol{\beta}} = [\underline{\bar{\beta}}, \bar{\bar{\beta}}]$ denote the supports for the Kalman gain and long-term price growth, respectively.

To capture the documented comovement between $(w, g, \bar{\beta})$, we proceed as follows. First, we standardize the stationary income shares w^i to obtain w^i_z. Next, we generate individual-level noise components: a common shock z^i and specific shocks for beliefs $\varepsilon^i_{\bar{\beta}}$ and gains ε^i_g, all drawn from standard normal distributions. We then generate correlated parameters:

$$\tilde{\beta}^i = \rho_{\bar{\beta},w} w^i_z + \rho_{\bar{\beta},z} z^i + (1 - \rho^2_{\bar{\beta},w} - \rho^2_{\bar{\beta},z})^{0.5} \varepsilon^i_{\bar{\beta}} \tag{21}$$

$$\tilde{g}^i = \rho_{g,w} w^i_z + \rho_{g,z} z^i + (1 - \rho^2_{g,w} - \rho^2_{g,z})^{0.5} \varepsilon^i_g \tag{22}$$

where ρ_{xy} denotes the correlation between variables x and y. Finally, we normalize $\tilde{\beta}^i$ and $\tilde{g}^i$ using min-max scaling to ensure they remain within their respective supports $\bar{\boldsymbol{\beta}}$ and $\boldsymbol{g}$.

3.2. Competitive Equilibrium

Definition 1 (Sequential Competitive Equilibrium)**.** *Given the exogenous processes, agents' probability measures, initial wealth holdings and the international interest rate R, a sequential competitive equilibrium is a set of stochastic sequences for all quantities $\{C^i_t, B^i_t, S^i_t\}_{i,t}$ and prices $\{P_t\}_t$ such that:*

1. *Allocations satisfy households optimality conditions:*

$$(C^i_t)^{-\gamma} = \delta\, \mathbb{E}^{\mathcal{P}^i}_t\big[(C^i_{t+1})^{-\gamma} R\big] \qquad \text{if } B^i_t > \underline{B}, \tag{23}$$
$$(C^i_t)^{-\gamma} > \delta\, \mathbb{E}^{\mathcal{P}^i}_t\big[(C^i_{t+1})^{-\gamma} R\big] \qquad \text{if } B^i_t = \underline{B}, \tag{24}$$

$$(C^i_t)^{-\gamma} = \delta\, \mathbb{E}^{\mathcal{P}^i}_t\big[(C^i_{t+1})^{-\gamma} R^s_{t+1}\big] \qquad \text{if } S^i_t > \underline{S}, \tag{25}$$
$$(C^i_t)^{-\gamma} > \delta\, \mathbb{E}^{\mathcal{P}^i}_t\big[(C^i_{t+1})^{-\gamma} R^s_{t+1}\big] \qquad \text{if } S^i_t = \underline{S}, \tag{26}$$

where $R^s_{t+1} \equiv \dfrac{D_{t+1} + P_{t+1}}{P_t}$.

2. ***Market clearing conditions:***

   (a) *Goods market:*

$$\sum_{i=1}^{I} \mu^i C_t^i = D_t + W_t - B_t + RB_{t-1}, \qquad B_t \equiv \sum_{i=1}^{I} \mu^i B_t^i, \quad W_t \equiv \sum_{i=1}^{I} \mu^i W_t^i.$$

   (b) *Stock market:*

$$\sum_{i=1}^{I} \mu^i S_t^i = 1.$$

Recursive Formulation. To recast the problem in recursive form, we need to specify the aggregate and individual state variables. Recall the budget constraint:

$$C_t^i + P_t S_t^i + B_t^i \leq W_t^i + (D_t + P_t) S_{t-1}^i + RB_{t-1}^i$$

In each period, household resources depend on the portfolio hold at the beginning of the period (S_{t-1}^i, B_{t-1}^i), her wages W_t^i, the realization of dividends which depends on the aggregate shock ε_t^D, and stock prices P_t. Then, a natural way to summarize household income position at the beginning of each period is by all these variables, which involve both individual and aggregate states

$$\boldsymbol{x}_t^i = (S_{t-1}^i, B_{t-1}^i, \nu_t^i, \varepsilon_t^W, \varepsilon_t^D, P_t) \tag{27}$$

Thus, household's problem can be recast as (without superscript i to save notation):

$$V(\boldsymbol{x}) = \max_{c,s',b'} u(c) + \delta \mathbb{E}^{\mathcal{P}}[V(\boldsymbol{x}')] \tag{28}$$

subject to the budget constraint and the asset holdings limits.

States involve past individual choices, exogenous variables and prices, which are an endogenous aggregate variable. In equilibrium, there will be a pricing function that maps resources to prices. Let $h^s(x_t^i)$ be the policy function for stocks. Equilibrium prices are the ones clearing the stock market, that is, they satisfy

$$\sum_i \mu^i h^s(S_{t-1}^i, B_{t-1}^i, \nu_t^i, \varepsilon_t^W, \varepsilon_t^D, P_t) = 1 \tag{29}$$

That equation can be solved for P_t such that the equilibrium pricing function p depends on the distribution of asset holdings, wages and the aggregate shocks, that is,

$$P_t = p(\bar{\Gamma}_t, \varepsilon_t^W, \varepsilon_t^D) \tag{30}$$

where $\bar{\Gamma}_t$ is the joint cumulative distribution function of asset holdings and wages, measuring how many households are below particular combinations of stock, bond holdings and idiosyncratic wage shocks, that is,

$$\bar{\Gamma}_t \equiv \bar{\Gamma}(s, b, \nu) = \sum_{i=1}^{I} \mathbf{1}\left\{ S_{t-1}^i \leq s, B_{t-1}^i \leq b, \nu_t^i \leq \nu \right\} \tag{31}$$

where $\mathbf{1}(\cdot)$ is the indicator function.

Models with Rational Expectations imply that agents know the pricing function p. Using that information, the vector of states becomes

$$(\boldsymbol{x}_t^i)^{RE} = (S_{t-1}^i, B_{t-1}^i, \nu_t^i, \varepsilon_t^W, \varepsilon_t^D, p(\bar{\Gamma}_t, \varepsilon_t^W, \varepsilon_t^D)) = (S_{t-1}^i, B_{t-1}^i, \nu_t^i, \varepsilon_t^W, \varepsilon_t^D, \bar{\Gamma}_t)$$

Thus, knowledge of the current distribution and the expected future distributions is needed for agents to make optimal choices. Since the distribution is a highly dimensional object, this represents a severe complication in these models.

Instead, we relax the information assumptions placed on the households' information set. We conjecture that there is imperfect information about the distribution of wealth, income and expectations in the economy, so that despite agents being rational, they cannot use optimality conditions to derive equilibrium prices. Instead, they consider current prices as a state variable and forecast future prices directly using the subjective model of prices set up above, which is summarized by their current price growth expectations β_t^i. Agents are, then, Internally Rational, ignoring the equilibrium process characterizing prices, but being fully rational conditional on that. In this case, the vector of states become

$$(\boldsymbol{x}_t^i)^{IR} = (S_{t-1}^i, B_{t-1}^i, \nu_t^i, \varepsilon_t^W, \varepsilon_t^D, P_t, \beta_t^i) \tag{32}$$

Equilibrium prices are determined by the stock market clearing condition:

$$\sum_i \mu^i h^s(S_{t-1}^i, B_{t-1}^i, \nu_t^i, \varepsilon_t^W, \varepsilon_t^D, P_t, \beta_t^i) = 1 \tag{33}$$

Let Γ_t be the current cumulative distribution over asset holdings, wages and beliefs, that is,

$$\Gamma_t \equiv \Gamma(s, b, \nu, \beta) = \sum_{i=1}^{I} \mathbf{1}\Big\{S_{t-1}^i \leq s, B_{t-1}^i \leq b, \nu_t^i \leq \nu, \beta_t^i \leq \beta\Big\} \tag{34}$$

With imperfect market knowledge, the equilibrium pricing function reads as

$$P_t = p(\Gamma_t, \varepsilon_t^W, \varepsilon_t^D) \tag{35}$$

Thus, this approach makes prices dependent on a higher-dimensional distribution with respect to rational expectations, as Γ_t includes also agents' beliefs. However, while the distribution Γ_t is needed to compute aggregate endogenous variables, it is not needed for solving the households' problem.

Definition 2 (Recursive Competitive Equilibrium under Internal Rationality)**.** *Given the exogenous processes, agents' probability measures, initial wealth holdings, and the international interest rate R, a recursive competitive equilibrium is a collection of:*

- *Individual policy functions for consumption and portfolio choice in the risky asset,*

$$\begin{pmatrix} C_t^i \\ v_t^i \end{pmatrix} = h^i\big((\boldsymbol{x}_t^i)^{IR}\big), \qquad (\boldsymbol{x}_t^i)^{IR} = (S_{t-1}^i, B_{t-1}^i, \nu_t^i, \varepsilon_t^W, \varepsilon_t^D, P_t, \beta_t^i), \tag{36}$$

  where v_t^i denotes the share of wealth invested in the risky asset and (C_t^i, S_t^i, B_t^i) are obtained from h^i using the budget constraint and asset-holding limits;

- *A pricing function*

$$P_t = p(\Gamma_t, \varepsilon_t^W, \varepsilon_t^D), \tag{37}$$

  where Γ_t is the cumulative distribution of asset holdings, wages and beliefs as in (34);

such that:

1. *Household optimality. For each agent i, the policy h^i solves the Bellman problem (28), subject to the period budget constraint (7), the asset-holding limits, and the law of motion for individual states and beliefs implied by $\mathcal{P}^i$. The associated first-order (Euler) conditions and complementary slackness conditions are satisfied.*

2. *Market clearing and pricing consistency.*

   (a) Stock market:

$$\sum_{i=1}^{I} \mu^i S_t^i = 1, \tag{38}$$

   where S_t^i is the stock position implied by h^i and the individual state $(\boldsymbol{x}_t^i)^{IR}$.

   (b) Goods market: *the aggregate resource constraint holds,*

$$\sum_{i=1}^{I} \mu^i C_t^i = D_t + W_t - B_t + RB_{t-1}, \qquad B_t \equiv \sum_{i=1}^{I} \mu^i B_t^i, \quad W_t \equiv \sum_{i=1}^{I} \mu^i W_t^i. \tag{39}$$

Note that in the Rational Expectations version one must also specify the law of motion for the distribution,

$$\bar{\Gamma}_{t+1} = H\big(\bar{\Gamma}_t, \varepsilon_t^W, \varepsilon_t^D, \varepsilon_{t+1}^W, \varepsilon_{t+1}^D\big), \tag{40}$$

since households need H to forecast future distributions, and thus future prices, when making current choices. Under Internal Rationality, by contrast, households do not track or use the distributional law H: they rely on their subjective model of prices and treat current prices as a state variable, directly forecasting future prices without reference to current or future distributions. Given the exogenous processes, agents' probability measures, initial wealth holdings and the international interest rate R, an equilibrium path under Internal Rationality can be simulated using only the household policy functions and the equilibrium pricing function. At each date, the beginning-of-period distribution and aggregate shocks pin down prices; prices, shocks and predetermined variables then determine allocations and belief updates, which in turn endogenously generate the next-period distribution.

3.3. Solution method

To solve the model, we need to characterize the functions h^c, h^v and p. We approximate these objects with parametric functions $\boldsymbol{\psi}$ of the state vector, following the Parametric Expectations Approach (PEA) Marcet (1988). The choice of functional form for $\boldsymbol{\psi}$ is neither obvious nor unique: common bases include polynomials, splines, and neural networks. Our strategy is to let economic theory guide the specification, in the spirit of homotopy methods: we start from policy functions that are exact solutions in simpler benchmark environments and preserve their structure as a first approximation in the richer model.

For consumption, we conjecture that the policy function is well approximated by

$$h^{c,i}(\boldsymbol{x}_t^i) \approx \psi_1(\boldsymbol{x}_t^i; \theta_1^i) = m_t^i Z_t^i, \tag{41}$$

where θ_1^i is a vector of type-specific parameters,

$$m_t^i = 1 - \theta_1^i \beta_t^i \tag{42}$$

is the marginal propensity to consume out of wealth (allowed to depend on the agent's belief β_t^i), and

$$Z_t^i = W_t^i + (P_t + D_t) S_{t-1}^i + R B_{t-1}^i \tag{43}$$

is beginning-of-period total wealth (cash-on-hand). Thus, conditional on beliefs, the consumption rule is linear in wealth, in line with the permanent-income hypothesis, where optimal consumption is (exactly or approximately) affine in total wealth (see Benhabib, Cui, and Miao (2024) for a recent example in the context of heterogeneous agents models). However, marginal propensities are idiosyncratic and time-varying, related to idiosyncratic beliefs about future returns; thus, more optimistic agents save more and consume less.[11]

Using ψ_1 in the budget constraint solves the consumption-savings problem. Yet, households still need to decide where to allocate savings. To select the risky share, we conjecture the following policy:

$$h^{v,i}(\boldsymbol{x}_t^i) \approx \psi_2(\boldsymbol{x}_t^i; \boldsymbol{\theta}_2^i) = \max\left\{\theta_{21}^i \theta_{22}^i + (1 - \theta_{21}^i)\, \frac{\beta_t^i - R}{\gamma(\sigma^i)^2}\ ,\ \underline{v}\right\}, \tag{44}$$

where $\underline{v}$ is a minimum desired equity share, β_t^i is agent i's subjective expectation of the gross stock return, σ^i is her perceived return volatility, and γ is risk aversion. Above the lower bound $\underline{v}$, the risky share is a convex combination of a fixed share θ_{22}^i and the Merton portfolio rule,

$$v_t^{M,i} = \frac{\beta_t^i - R}{\gamma(\sigma^i)^2},$$

[11] We simulate the model for the ratio of variables to dividends, but we keep here the non-normalized variables for the clarity of the exposition.

which is the optimal static portfolio choice in the canonical mean–variance / Merton model (with CARA–Normal or, approximately, CRRA–lognormal returns). In our dynamic environment with state variables and changing investment opportunities, the true optimal portfolio generally includes an additional intertemporal hedging component that the simple Merton rule omits. The parameters θ^i_{21} and θ^i_{22} allow the policy to deviate from the pure Merton benchmark and flexibly absorb such hedging motives and other dynamic effects, while keeping the Merton formula as the core building block of the risky-share decision.

Using $h^{c,i}$ and $h^{v,i}$ in the budget constraint, the stock policy function is approximated by

$$h^{s,i}(\boldsymbol{x}^i_t) \approx (1 - m^i_t) v^i_t \frac{Z^i_t}{P_t} \tag{45}$$

and for bonds,

$$h^{b,i}(\boldsymbol{x}^i_t) \approx (1 - m^i_t)(1 - v^i_t) Z^i_t \tag{46}$$

Using $h^{s,i}$ for all the agents in the stock market clearing condition and solving it for prices,

$$P_t = \frac{\sum_{i=1}^{1} (1 - m^i_t) v^i_t \Big(W^i_t + D_t S^i_{t-1} + R B^i_{t-1} \Big)}{1 - \sum_{i=1}^{1} (1 - m^i_t) v^i_t S^i_{t-1}} \tag{47}$$

which depends on the distribution of beliefs, wages and wealth at the beginning of the period. The particular functional-form guess for policies lets us solve for equilibrium prices analytically, avoiding a costly root-finding step. This makes price determination both conceptually transparent and computationally fast. Importantly, the choice of approximation functions is orthogonal to our main methodological point: under Internal Rationality, prices are explicit state variables, so agents condition on prices and their forecasts rather than on the entire cross-sectional distribution. Our specific functional forms simply provide an additional layer of simplification for the solution method; alternative approximations could be used without changing this core insight.

4. Quantitative Analysis

Our quantitative exercise aims to measure the impact of heterogeneous expectations on wealth inequality. The calibration strategy follows three steps. First, we calibrate key distributional parameters using U.S. data, focusing on the wage distribution, aggregate income dynamics, and the empirically observed distribution of expectational

Table 5: Model Parameters Calibration.

Parameter	Symbol	Value	Source
Panel A: Aggregate Parameters			
Discount factor	δ	0.99	Conventional
Mean dividend growth	β^D	2%	US data
Dividends growth standard deviation	σ_D	1.9%	US data
Risk-free rate	$R-1$	2.32%	US data
Risk aversion parameter	γ	2	Conventional
Panel B: Distributional Parameters			
Correlation wages expectations	$\rho_{\bar{\beta}w}$	0.85	RAND panel
Pareto distribution parameter	α	2.5	Match wage distribution
Mean wage shares	μ	10	Match wage distribution
Panel C: Expectation Parameters			
State persistence	ρ	0.90	RAND panel
Kalman gain	g	0.02	AMB(2017)
Fixed effects support	$[\underline{\beta},\ \bar{\beta}]$	[0.005,0.02]	Match expected returns

fixed effects. Second, we ensure that the correlation between wages and expectations matches our empirical findings from the RAND panel. Then, we let the model generate its endogenous wealth distribution and evaluate its performance by comparing it with the data. Finally, to quantify the impact of heterogeneous expectations on wealth inequality, we simulate a counterfactual model where all households share identical beliefs calibrated to match the average expected returns in the data. By comparing this homogeneous-beliefs economy with our baseline model, we can isolate the contribution of expectational heterogeneity to wealth concentration.

4.1. Calibration Strategy

Table 5 presents our parameter choices. The aggregate parameters follow standard values in the macro-finance literature, with a quarterly discount factor of 0.99 and a risk aversion parameter of 2. The dividend process parameters are calibrated to match U.S. data, with mean growth of 2% and standard deviation of 1.9%. The wage-dividend ratio and its dynamics are chosen to match the capital income share and wage dynamics in the data.

4.2. Performance

Tables 6 and 7 assess the model's performance along various dimensions. The model generates aggregates which are reasonable. For instance, the equity premium is 8.57%, the capital income share of 25.79%, and a savings rate of 4.59%. Additionally, the safe-to-total assets ratio of 44.63% aligns well with aggregate portfolio allocations in the data.

Table 6: Model Performance: Aggregate Moments.

		Model
Mean equity return	$\mathbb{E}(r_t^s)$	10.58%
Equity premium	$\mathbb{E}(r_t^s) - R - 1$	8.57%
Mean income growth	g^Y	2.11%
Mean capital income share	α_k	25.79%
Mean safe to total assets	$\mathbb{E}(B_t/(B_t + P_t S))$	44.63%
Mean savings rate	$\mathbb{E}(s_t)$	4.59%

Table 7 compares key distributional moments between our model and U.S. data. The model generates significant inequality in wealth, income, and returns across the wealth distribution, though it understates the degree of wealth concentration at the top. While the model predicts a top 1% wealth share of 10.64% compared to 38.5% in the data, it successfully captures the gradient in both portfolio returns and expectations. The top 1% achieves portfolio returns of 10.41% compared to 4.44% for the bottom 50%, close to the empirical spread of 8.30% to 3.40%. Similarly, expected returns exhibit a monotonic pattern across wealth groups, ranging from 8.63% for the top 1% to 4.41% for the bottom 50%, matching the empirical pattern.

The model also replicates key patterns in capital income. The capital income share increases sharply with wealth, with the top 1% earning 43.96% of their income from capital compared to 15.86% for the bottom 50%. While these numbers understate the empirical concentration (52.4% versus 3.8%), they capture the qualitative relationship between wealth and capital income. The wage distribution in the model aligns well with the data, particularly for the middle of the distribution, though it somewhat understates inequality at both extremes.

Table 8 shows that the model successfully replicates key features of expectation formation documented in the survey data. The model generates extrapolation ($\hat{b} = -0.20$) and overreaction ($\hat{d} = -0.30$), though it somewhat overstates the de-

Table 7: Distributional Moments: Model vs Data. *This table compares distributional moments from the model with their empirical counterparts. All values are in percentages. Wealth, income, and wages show the share of total. Capital income share represents the fraction of total income from capital for each group. Portfolio and expected returns are average percentage returns for each group.*

	Model				Data			
By Wealth percentile	Top 1%	Top 10%	Middle 40%	Bottom 50%	Top 1%	Top 10%	Middle 40%	Bottom 50%
Wealth	10.64	49.80	37.41	12.79	38.5	77.1	21.7	1.2
Income	8.87	45.83	40.15	14.02	18.4	44.2	35	20.8
Wages	6.60	40.66	43.70	15.63	11.7	34	39.1	26.9
Capital income share	43.96	28.01	17.75	15.86	52.40	41.10	9.30	3.80
Portfolio return	10.41	7.24	4.95	4.44	8.30	6.12	3.63	3.40
Expected return	8.63	6.39	4.76	4.41	8.10	7.70	6.10	4.30

gree of extrapolation relative to the data. The model also matches the relative importance of individual fixed effects versus time effects in explaining expectation heterogeneity, with R-squared values of 0.54 and 0.04 respectively, close to their empirical counterparts of 0.50 and 0.01.

Table 8: Model Performance: Expectation Formation. **** indicates statistical significance at the 1% level.*

		Model	Data
Extrapolation coefficient	b	-0.20***	-0.13***
Overreaction coefficient	d	-0.30***	-0.50***
R^2 Individual Fixed Effects	$R^2_{\alpha^i}$	0.54	0.50
R^2 Time Fixed Effects	$R^2_{\mu_t}$	0.04	0.01
Correlation (expectations, wealth)	$\rho_{\beta,W}$	0.88	0.95

Figure 10 shows the joint distribution of wealth and expectations in the model. As in the data, an important fraction of the population is an area with low wealth and pessimistic beliefs and a tiny minority are wealthy and optimistic.

4.3. The Impact of Heterogeneous Expectations on Wealth Inequality

Table 9 presents our key quantitative results, comparing distributional outcomes between our baseline model with heterogeneous expectations and a counterfactual

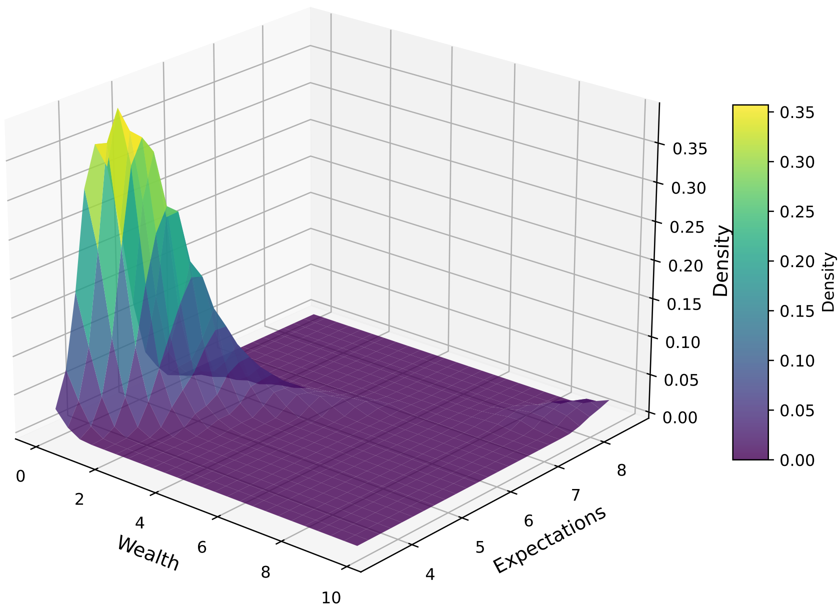


Figure 10: Joint distribution of wealth and expectations in the model.

economy where all households share identical beliefs. The contrast is striking: heterogeneous expectations substantially amplify wealth concentration. The wealth share of the top 1% increases from 7.02% under homogeneous expectations to 10.64% with belief heterogeneity, while the top 10% share rises from 40.71% to 49.80%. Figure 11 illustrates the effect. This amplification occurs through two channels: portfolio returns and capital income shares.

First, heterogeneous expectations generate substantial variation in portfolio returns across the wealth distribution. In the baseline model, the top 1% achieves returns of 10.41% compared to 4.44% for the bottom 50%. In contrast, returns are nearly identical across wealth groups under homogeneous expectations, hovering around 6.80%. Second, belief heterogeneity leads to dramatic differences in capital income shares. While these shares are uniform at roughly 26% under homogeneous expectations, they range from 43.96% for the top 1% to 15.86% for the bottom 50% in the baseline model.

Notably, consistent with Stachurski and Toda (2019), the wealth distribution under homogeneous beliefs inherits the tail behavior of the income distribution under homogeneous expectations. However, heterogeneous expectations generate additional skewness through their effects on portfolio choices and savings behavior, substantially amplifying wealth concentration at the top of the distribution.

Table 9: The Effect of Heterogeneous Expectations. *This table compares distributional moments between our baseline model with heterogeneous expectations and a counterfactual economy with homogeneous expectations. All values are in percentages. The homogeneous expectations case uses the average expected return from the baseline model.*

	Heterogeneous Expectations				Homogeneous Expectations			
By Wealth percentile	Top 1%	Top 10%	Middle 40%	Bottom 50%	Top 1%	Top 10%	Middle 40%	Bottom 50%
Wealth	10.64	49.80	37.41	12.79	7.02	40.71	43.66	15.62
Income	8.87	45.83	40.15	14.02	7.05	40.67	43.71	15.62
Wages	6.60	40.66	43.70	15.63	7.06	40.66	43.72	15.62
Capital income share	43.96	28.01	17.75	15.86	25.76	25.90	25.86	25.88
Portfolio return	10.41	7.24	4.95	4.44	6.80	6.79	6.80	6.79
Expected return	8.63	6.39	4.76	4.41	5.36	5.46	5.41	5.43

5. Learning from Experience

While our baseline model successfully captures the joint distribution of expectations and wealth, it relies on an unsatisfying assumption: individuals have fixed long-run views about stock market performance. In this framework, people are simply born optimists or pessimists, and these fixed beliefs drive wealth inequality. Though this approach provides a transparent first step for understanding the role of heterogeneous beliefs, it imposes a unidirectional causality—beliefs determine wealth—when the relationship could equally run in reverse, with wealth shaping optimism.

This section develops a microfoundation for these individual fixed effects that captures the two-way interaction between wealth and beliefs. The key insight is that initial differences in portfolio choices generate heterogeneous realized returns. When agents learn about aggregate returns from their own portfolio performance, these different experiences create persistent belief divergence through a feedback mechanism: successful investors become more optimistic and choose riskier portfolios, leading to higher mean returns that further reinforce their beliefs.

5.1. The mechanism

A large literature on learning from experience shows that economic agents do not use the full available history of aggregate data, but instead over-weight personally lived realizations when forming expectations. One influential strand, following Malmendier and Nagel (2011), models experience as cohort-specific exposure to aggregate realizations: agents of different ages have observed different subsets of the

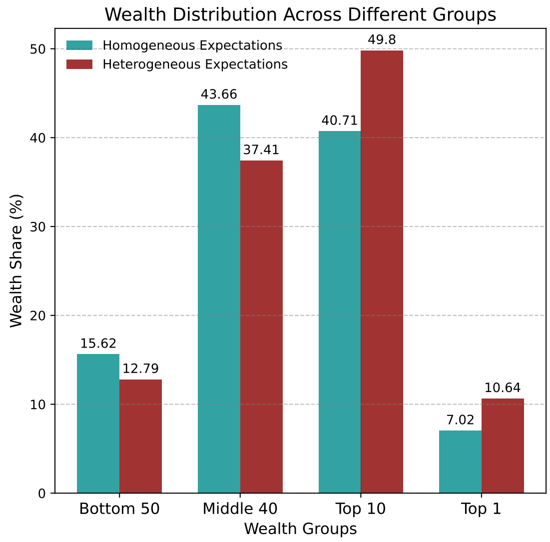


Figure 11: Wealth distribution with and without heterogeneous expectations.

aggregate time series (inflation, stock returns, etc.) and therefore use different effective samples when estimating moments of the underlying process. A complementary strand, more recently, emphasizes that idiosyncratic or local experiences can spill over into beliefs about aggregate conditions and even other domains: for example, individuals extrapolating from local house price or labor market conditions to macro expectations, or from personal credit and lending experiences (such as loan rejections or tighter credit conditions) to beliefs about nationwide credit markets and the macroeconomy (see, among others, Kuchler and Zafar (2019), D'Acunto et al. (2021) or Fidrmuc, Hainz, and Hölzl (2024)). Our mechanism belongs to this broader individual-to-aggregate extrapolation tradition, but implements it in a portfolio-choice environment and links it directly to wealth dynamics.

In our setting, all households observe the same aggregate price and dividend history, so a pure "cohort-sample" story would generate limited disagreement for agents of similar age. Instead, we exploit the fact that in an economy with heterogeneous portfolios, individual portfolio returns can differ substantially and persistently from the market return. A highly exposed household that holds mostly equities experiences a much more volatile and, on average, higher return process than a conservative household holding mostly safe assets. If investors treat their own portfolio performance as an informative signal about the underlying return process, then heterogeneous portfolio choices generate heterogeneous signals and, consequently, heterogeneous beliefs, even for observationally similar agents who live through exactly the same aggregate history. This closes the loop between beliefs and wealth, in a way that parallels the broader evidence on extrapolation from personal experiences to macro beliefs, but here operates specifically through realized portfolio returns.

To capture this idea, we model subjective returns as having both a common and an idiosyncratic component:

$$\begin{aligned} \ln R_t &= \mu \ln b_t^i + (1-\mu)\,\ln R_t^i + \ln \varepsilon_t^{R,i}, \\ \ln b_t^i &= \ln b_{t-1}^i + \ln \varepsilon_t^{b,i}, \end{aligned} \tag{48}$$

where R_t denotes the (gross) market return and R_t^i is household i's realized portfolio return. The latent component b_t^i plays the role of a persistent individual trend or long-run view about returns, and $\mu \in [0,1]$ governs the weight agents assign to this common long-run component relative to the idiosyncratic portfolio component. The shocks $\varepsilon_t^{R,i}$ and $\varepsilon_t^{b,i}$ are i.i.d. innovations capturing transitory noise in perceived returns and the evolution of the long-run component, respectively. When μ is high, investors put more weight on the slow-moving long-run belief b_t^i; when μ is low, they rely more strongly on their recent portfolio experience R_t^i.

Households do not observe b_t^i directly; instead, they filter it from observed returns. Let β_t^i denote investor i's time-t estimate of the persistent component b_t^i. Given the perceived state-space structure above, the optimal updating rule is a Kalman filter of the form

$$\ln \beta_t^i = \ln \beta_{t-1}^i + g^i \Big(\underbrace{\ln R_{t-1} - \ln R_{t-1}^i}_{\text{common vs. own return}} + \underbrace{\mu\big(R_t^i - \beta_t^i\big)}_{\text{return surprise vs. trend}} \Big) + g^i \varepsilon_t^{R1,i}, \tag{49}$$

where $g^i \in (0,1)$ is a (possibly heterogeneous) Kalman gain summarizing how aggressively investor i reacts to new information, and $\varepsilon_t^{R1,i}$ is the innovation in the observed return signal.

This updating equation highlights two distinct channels through which experience matters. First, the term $\ln R_{t-1} - \ln R_{t-1}^i$ compares the aggregate return to the household's own portfolio return. If R_{t-1}^i has systematically exceeded the market return—e.g., because the household has been heavily invested in equities during a boom—then this term is negative on average, leading the investor to infer a more favourable underlying process and gradually revise β_t^i upward.

Second, the term $\mu(R_t^i - \beta_t^i)$ captures the surprise in the household's realized return relative to its own long-run trend. Positive return surprises push beliefs up;

negative surprises push them down. The parameter μ modulates how much these surprises feed back into the permanent component as opposed to being treated as transitory noise.

Together, these channels implement a natural notion of experience-based learning that is portfolio-specific: investors learn about the common return process through the lens of their own realized returns. An investor who has repeatedly "ridden the boom" with a risky portfolio experiences a sequence of high R_t^i realizations and, through the updating rule above, becomes persistently more optimistic. A conservative investor with low equity exposure experiences a much flatter and less volatile return process and therefore revises beliefs far less, remaining comparatively pessimistic.

Given the filtered belief β_t^i, investor i's subjective expectation of next period's gross return takes the multiplicative form

$$\mathbb{E}_t^{\mathcal{P}^i}(R_{t+1}) = \kappa^i \left(\beta_t^i\right)^{\mu} \left(R_t^i\right)^{1-\mu}, \tag{50}$$

where κ^i is a log-normal correction factor that ensures consistency between the perceived log-linear process and level returns. This specification nests several useful benchmarks. If $\mu = 1$, expectations depend only on the filtered long-run component β_t^i, and we are back to a standard model of adaptive learning. If $\mu = 0$, expectations are fully extrapolative: investors simply project their latest portfolio return R_t^i forward. Intermediate values of μ generate a convex combination of long-run beliefs and recent portfolio experience.

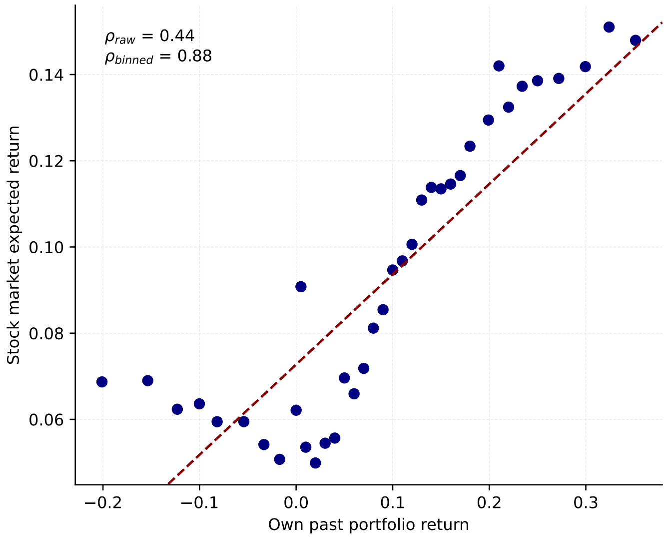
Relative to age-based experience models, our mechanism shifts the focus from *cohort-driven* differences in exposure to aggregate histories to *portfolio-driven* differences in exposure to risky returns, while remaining conceptually close to the broader literature where individual experiences shape beliefs about aggregate outcomes. In our framework, two agents of the same age and information set can hold very different beliefs because they chose different portfolios and hence experienced different return paths. As a result, investors who happen to experience good early outcomes become increasingly optimistic, move further into risky assets, and accumulate wealth at a much faster rate, while those with poor early experiences become more pessimistic and remain under-exposed to equity. This portfolio-experience channel provides a natural microfoundation for the persistent individual fixed effects in expectations documented in the data and substantially amplifies wealth inequality relative to models in which beliefs are homogeneous or exogenously fixed.

5.2. Evidence about the mechanism

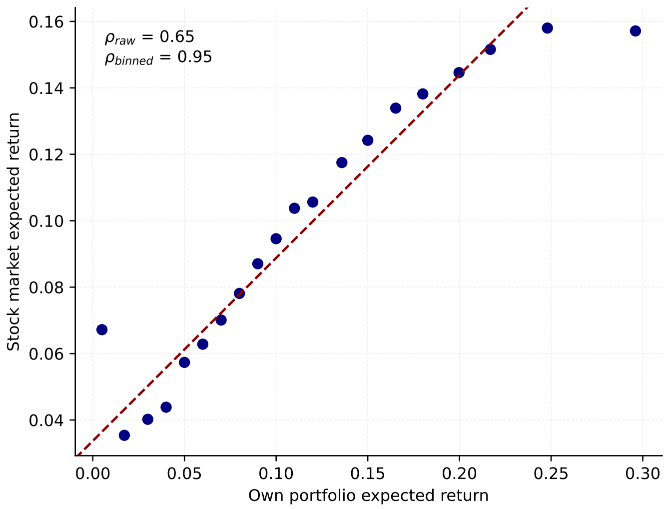
The experience-based learning mechanism developed above implies a tight link between investors' own portfolio returns and their expectations about aggregate market returns. Unfortunately, the RAND dataset does not contain information on individual portfolio performance. By contrast, the UBS–Gallup data include (i) realized past portfolio returns R^i_{t-1}, (ii) survey expectations about future *market* returns $\mathbb{E}^i_t(R^m_{t+1})$, and (iii) survey expectations about future *own portfolio* returns $\mathbb{E}^i_t(R^i_{t+1})$. This richer information set allows us to directly examine whether the patterns in the cross-section of beliefs are consistent with experience-based learning: we can test whether higher own past returns are systematically associated with more optimistic expectations about both market and own portfolio returns, and whether expectations about own and market returns co-move in the way implied by the model.

We begin by documenting the raw relationships in the data. Figure 12 presents binned scatter plots that reveal a strong positive association between portfolio returns (both realized and expected) and expected market returns. Conditional on positive realized returns, higher past portfolio returns are monotonically associated with higher expected stock market returns (Figure 12a). This pattern is exactly what a simple learning-from-experience mechanism would predict: investors who have recently done well in the market extrapolate their good performance into more optimistic beliefs about future aggregate returns. For negative portfolio returns, by contrast, the relationship becomes non-monotonic. Respondents who have recently suffered losses still report low but *positive* expected market returns, suggesting that bad experiences dampen optimism but do not typically lead to outright pessimism about the equity premium.

Figure 12b compares expectations about own portfolio versus market returns. The correlation between these two variables is remarkably strong: about 0.65 in the raw micro data and 0.95 when using binned averages. The raw correlation of 0.65 indicates that investors' beliefs about their own portfolio and the overall market are tightly linked, but still contain a significant idiosyncratic component at the individual level (e.g., measurement error, idiosyncratic risk, or investor-specific narratives about stock-picking ability). Once we average beliefs within percentiles of the distribution, most of this idiosyncratic noise is washed out, and the correlation rises to 0.95, very close to a one-to-one relationship. This near-perfect binned correlation is consistent with a strong common component in beliefs: investors largely move their expectations about own and market returns together, with deviations around that common factor being small and largely unsystematic.



(a) Own past portfolio return vs. stock market expected return.



(b) Own portfolio vs. market expected return.

Figure 12: Binscatter plot: Own past and expected returns vs. stock market expected return in UBS-Gallup survey. *The data sample is 1998-2007. Outliers have been removed.*

Complementarily, we can derive more structured predictions from the model in Section 5.1. In the model, expectations about the market satisfy

$$\mathbb{E}_t^{\mathcal{P}^i}\left(R_{t+1}^m\right) = \kappa_m^i \left(\beta_t^i\right)^{\mu_m} \left(R_t^i\right)^{1-\mu_m}, \tag{51}$$

so that

$$\ln \mathbb{E}_t^{\mathcal{P}^i}(R_{t+1}^m) = \ln \kappa_m^i + \mu_m \ln \beta_t^i + (1-\mu_m) \ln R_t^i. \tag{52}$$

Holding the long-run component β_t^i fixed, the model predicts

$$\frac{\partial \ln \mathbb{E}_t^{\mathcal{P}^i}(R_{t+1}^m)}{\partial \ln R_t^i} = 1 - \mu_m > 0,$$

i.e. investors extrapolate from their own past portfolio performance to beliefs about aggregate returns. A corresponding empirical specification is

$$\ln \mathbb{E}_t^i(R_{t+1}^m) = \alpha_m + \gamma_m \ln R_{t-1}^i + \delta_m \widehat{\ln \beta_t^i} + X_t^{i\prime} \lambda_m + u_{m,t}^i, \tag{53}$$

where $\widehat{\ln \beta_t^i}$ is a proxy for the long-run belief about market returns and X_t^i are controls. The key testable implication is

$$\mathcal{H}_0 : \gamma_m = 0 \quad \text{vs.} \quad \mathcal{H}_1 : \gamma_m > 0.$$

Standard age-based experience models (where beliefs depend only on aggregate histories) imply $\gamma_m = 0$ once $\widehat{\ln \beta_t^i}$ and demographics are controlled for. Our portfolio-based learning mechanism predicts $\gamma_m > 0$: conditional on experienced market history, investors with higher past portfolio returns should report higher expected market returns.

To bring equation (53) to the data, we proxy the long-run belief about market returns $\widehat{\ln \beta_t^i}$ with the log of the respondent's expected own portfolio return, $\ln \mathbb{E}_t^i(R_{t+1}^i)$. We estimate (53) in the UBS–Gallup cross-section using survey weights and heteroskedasticity-robust standard errors, first in an unconstrained specification and then imposing the model restriction $\delta_m = 1 - \gamma_m$.

Table 10 reports estimates of equation (53) using UBS–Gallup survey data and expected own portfolio returns as a proxy for long-run beliefs, $\widehat{\ln \beta_t^i} \approx \ln \mathbb{E}_t^i(R_{t+1}^i)$. Column (1) includes only the experience term and the proxy for long-run beliefs and column (2) adds controls for age, income, and total financial assets; column (3) and (4) are equivalent but restricting $\delta_m = 1 - \gamma_m$. In all specifications, the coefficient on past own portfolio returns, γ_m, is positive and precisely estimated: $\hat{\gamma}_m \approx 0.13$ without controls and $\hat{\gamma}_m \approx 0.12$ with controls, with p-values well below 1% when unconstrained; and $\hat{\gamma}_m \approx 0.38$ with parameter restrictions. Thus, we strongly reject the null $\mathcal{H}_0 : \gamma_m = 0$ in favor of $\mathcal{H}_1 : \gamma_m > 0$.

The proxy for long-run beliefs, also enters with a large, highly significant coefficient around 0.62, indicating that investors who are more optimistic about their own portfolio are also more optimistic about the market. Adding age, income, and savings has only a modest effect on the fit of the regression and does not attenuate the experience coefficient. Overall, these results are difficult to reconcile with pure age-based experience models, which predict $\gamma_m = 0$ once long-run beliefs and demographics are controlled for, and instead support the portfolio-based learning mechanism whereby higher own past returns are associated with higher expected market returns.

Table 10: Experience-based learning: survey evidence. *Robust standard errors in parentheses.* $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. *Data from the UBS - Gallup Survey.*

	(1)	(2)	(3)	(4)
	$\ln \mathbb{E}_t^i(R_{t+1}^m)$			
$\ln(R_{t-1}^i)$	0.1304***	0.1206***	0.3834***	0.3852***
	(0.025)	(0.025)	(0.007)	(0.008)
$\ln \mathbb{E}_t^i(R_{t+1}^i)$	0.6183***	0.6152***	0.6166***	0.6148***
	(0.007)	(0.007)	(0.007)	(0.008)
Controls (income, savings, age)	-	✓	-	✓
Model restriction $\delta_m = 1 - \gamma_m$	-	-	✓	✓
R^2	0.438	0.441	0.201	0.203
Observations	26,607	26,607	26,607	26,607

5.3. Quantification of the mechanism

In this section, we simulate the model with learning from experience. Crucially, this mechanism amplifies rather than originates inequality. Starting from perfect equality, the mechanism alone generates no inequality. However, any source of initial heterogeneity—whether idiosyncratic portfolio shocks or differences in fundamental characteristics like risk aversion—activates the learning channel.

For comparability with the literature, we modify the model from Section 3: the subjective return model is given by (48), incorporating past own returns but excluding idiosyncratic long-term views, and risk aversion parameters γ^i are agent-specific. We first simulate an economy with heterogeneous risk aversion but homogeneous beliefs, where agents learn symmetrically about returns. Heterogeneous risk aversion generates some inequality—more risk-tolerant investors hold more stocks and become wealthier—but this model fails to reproduce the empirical correlation between beliefs and wealth. Column 1 of Table 11 reports the resulting distributions. While the gradients are qualitatively correct (higher capital income share, realized returns, and expected returns for the wealthy), the magnitudes are far too small. Expected returns are virtually identical across the wealth distribution, contradicting the data.

Column 2 introduces learning from own portfolio returns, dramatically improving model fit. The belief distribution now aligns with empirical evidence: the top 1% expect 9% returns while the bottom 50% expect below 4%, despite observing identical market returns. Without any predetermined views, systematic disagreement emerges endogenously across the wealth distribution. The mechanism also substantially amplifies inequality: the top 1% wealth share increases from 9% to 14% (a 50% increase), while the bottom 90% share falls from 54% to 47%. Capital income comprises over 80% of top 1% income, up from 50%. Thus, learning from experience endogenously generates the wealth-belief correlation and substantially amplifies inequality.

Table 11: The Amplification Effect of Learning from Experience. *This table compares distributional outcomes between an economy with only heterogeneous risk aversion and one that adds learning from own portfolio returns. All values are in percentages.*

	Heterogeneous Risk Aversion				+ Learning from experience			
By Wealth percentile	Top 1%	Top 10%	Middle 40%	Bottom 50%	Top 1%	Top 10%	Middle 40%	Bottom 50%
Wealth	9.34	45.96	39.93	14.11	14.36	52.47	35.38	12.15
Capital income share	48.85	39.55	30.54	28.45	83.01	34.62	21.14	18.46
Portfolio return	5.13	4.03	3.12	2.95	8.84	4.72	3.31	2.96
Expected return	9.70	9.69	9.49	9.25	9.06	5.40	4.09	3.79

Crucially, learning generates a feedback loop that amplifies initial differences: heterogeneous portfolios lead to different realized returns, which through learning create heterogeneous expectations, further reinforcing portfolio heterogeneity. This mechanism transforms initially similar expectations (varying only from 9.25% to 9.70% across wealth groups under pure risk aversion heterogeneity) into highly dispersed beliefs (ranging from 3.79% to 9.06% with learning).

These results suggest that learning from experience can provide a unified explanation for the joint distribution of returns, expectations, and wealth observed in the data. Moreover, it identifies a novel amplification mechanism: the interaction between portfolio heterogeneity and belief formation through learning can substantially magnify wealth inequality beyond what standard theories would predict.

6. Conclusion

This paper documents substantial heterogeneity in households' stock-return expectations. Using survey data, we show that poorer households are systematically too pessimistic about stock market performance, while richer households' expectations are approximately aligned with realized returns. As a result, richer households are both more heavily invested in equities and better informed about the stock market.

We develop a model in which households learn about returns by treating their own portfolio performance as a signal about the underlying market process. This portfolio-based experience mechanism generates persistent belief heterogeneity and, when disciplined with survey evidence, replicates key micro facts: the stability of individual belief differences over time, the positive association between optimism and wealth, and between past own returns and expected market returns.

Embedding this expectations framework into a heterogeneous-agent model with incomplete markets, aggregate risk, and portfolio choice, we quantify the role of belief heterogeneity for wealth concentration. Heterogeneous expectations emerge as a powerful amplification mechanism: when we shut down belief dispersion in the calibrated model, the wealth share of the top 1% falls by roughly 50%, and the cross-sectional dispersion in portfolio returns nearly disappears. These findings suggest that accounting for heterogeneity in beliefs is crucial for understanding the dynamics of wealth inequality.

Methodologically, our use of Internal Rationality makes the model consistent with survey evidence while simplifying the solution of heterogeneous-agent models with aggregate risk. By treating prices as state variables that agents forecast directly, we avoid the need for households to form expectations over infinite-dimensional distributions, addressing recent concerns about the computational tractability of these frameworks.

Looking ahead, our results open several avenues for future research. The pronounced excess pessimism of the bottom ninety percent of households calls for deeper work on the role of information frictions, financial literacy, and market participation costs. More broadly, the complexity reduction delivered by Internal Rationality could be applied to other forms of heterogeneity and to nonlinear macroeconomic phenomena as economic fluctuations. Finally, our findings invite a reconsideration of policy debates—such as about the design of sovereign wealth funds—as potential tools to democratize risk-taking and return, in particular if the excessive pessimism of the poor is rooted in policy-invariant primitives.

References

Adam, Klaus and Albert Marcet (2011). "Internal rationality, imperfect market knowledge and asset prices". In: *Journal of Economic Theory* 146.3, pp. 1224–1252.

Adam, Klaus, Albert Marcet, and Johannes Beutel (2017). "Stock price booms and expected capital gains". In: *American Economic Review* 107.8, pp. 2352–2408.

Andrade, Philippe, Richard K Crump, Stefano Eusepi, and Emanuel Moench (2016). "Fundamental disagreement". In: *Journal of Monetary Economics* 83, pp. 106–128.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2011). "The distribution of wealth and fiscal policy in economies with finitely lived agents". In: *Econometrica* 79.1, pp. 123–157.

Benhabib, Jess, Wei Cui, and Jianjun Miao (2024). "Capital income jumps and wealth distribution". In: *Quantitative Economics* 15.4, pp. 1197–1247.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020). "Overreaction in macroeconomic expectations". In: *American Economic Review* 110.9, pp. 2748–2782.

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull (2003). "Accounting for the US earnings and wealth inequality". In: *Journal of political economy* 111.4, pp. 818–857.

Coibion, Olivier and Yuriy Gorodnichenko (2012). "What can survey forecasts tell us about information rigidities?" In: *Journal of Political Economy* 120.1, pp. 116–159.

— (2015). "Information rigidity and the expectations formation process: A simple framework and new facts". In: *American Economic Review* 105.8, pp. 2644–2678.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber (2021). "Exposure to grocery prices and inflation expectations". In: *Journal of Political Economy* 129.5, pp. 1615–1639.

De Nardi, Mariacristina (2004). "Wealth inequality and intergenerational links". In: *The Review of Economic Studies* 71.3, pp. 743–768.

De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz Pardo (2016). *The implications of richer earnings dynamics for consumption and wealth*. Tech. rep. National Bureau of Economic Research.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri (2020). "Heterogeneity and persistence in returns to wealth". In: *Econometrica* 88.1, pp. 115–170.

Fernández-Villaverde, Jesús and Oren Levintal (2024). "The Distributional Effects of Asset Returns". In.

Fidrmuc, Jarko, Christa Hainz, and Werner Hölzl (2024). "Individual credit market experience and beliefs about bank lending policy: evidence from a firm survey". In: *The Scandinavian Journal of Economics* 126.2, pp. 387–414.

Gaudecker, Hans-Martin von and Axel Wogrolly (2022). "Heterogeneity in households' stock market beliefs: Levels, dynamics, and epistemic uncertainty". In: *Journal of Econometrics* 231.1, pp. 232–247.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus (2021). "Five facts about beliefs and portfolios". In: *American Economic Review* 111.5, pp. 1481–1522.

Greenwood, Robin and Andrei Shleifer (2014). "Expectations of returns and expected returns". In: *The Review of Financial Studies* 27.3, pp. 714–746.

Hubmer, Joachim, Per Krusell, and Anthony A Smith Jr (2021). "Sources of US wealth inequality: Past, present, and future". In: *NBER Macroeconomics Annual* 35.1, pp. 391–455.

Kaymak, Barış, David Leung, and Markus Poschke (2022). *Accounting for wealth concentration in the United States*. Tech. rep. Federal Reserve Bank of Cleveland.

Kaymak, Barış and Markus Poschke (2016). "The evolution of wealth inequality over half a century: The role of taxes, transfers and technology". In: *Journal of Monetary Economics* 77, pp. 1–25.

Kohlhas, Alexandre N and Ansgar Walther (2021). "Asymmetric attention". In: *American Economic Review* 111.9, pp. 2879–2925.

Krusell, Per and Anthony A Smith Jr (1998). "Income and wealth heterogeneity in the macroeconomy". In: *Journal of political Economy* 106.5, pp. 867–896.

Kuchler, Theresa and Basit Zafar (2019). "Personal experiences and expectations about aggregate outcomes". In: *The Journal of Finance* 74.5, pp. 2491–2542.

Malmendier, Ulrike and Stefan Nagel (2011). "Depression babies: Do macroeconomic experiences affect risk taking?" In: *The quarterly journal of economics* 126.1, pp. 373–416.

Mankiw, N Gregory, Ricardo Reis, and Justin Wolfers (2003). "Disagreement about inflation expectations". In: *NBER macroeconomics annual* 18, pp. 209–248.

Marcet, Albert (1988). “Solving nonlinear stochastic growth models by parametrizing expectations.” In.

Moll, Benjamin (2024). “The Trouble with Rational Expectations in Heterogeneous Agent Models: A Challenge for Macroeconomics”. In: *CEPR working paper*.

Patton, Andrew J and Allan Timmermann (2010). “Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion”. In: *Journal of Monetary Economics* 57.7, pp. 803–820.

Piketty, Thomas and Emmanuel Saez (2003). “Income inequality in the United States, 1913–1998”. In: *The Quarterly journal of economics* 118.1, pp. 1–41.

Quadrini, Vincenzo (2000). “Entrepreneurship, saving, and social mobility”. In: *Review of economic dynamics* 3.1, pp. 1–40.

Stachurski, John and Alexis Akira Toda (2019). “An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity”. In: *Journal of Economic Theory* 182, pp. 1–24.

Tobias, Broer, Alexandre Kohlhas, Kurt Mitman, and Kathrin Schlafmann (2022). “Expectation and Wealth Heterogeneity in the Macroeconomy”. In.

Vissing-Jorgensen, Annette (2003). “Perspectives on behavioral finance: Does" irrationality" disappear with wealth? Evidence from expectations and actions”. In: *NBER macroeconomics annual* 18, pp. 139–194.