Heterogenous Beliefs about Stock Returns and Wealth Inequality

Kwok Yan Chiu*

December 16, 2024

Most Recent Version

Abstract

This paper investigates how differences in subjective beliefs about stock returns contribute to wealth inequality through portfolio choice. I argue that this channel significantly contributes to wealth inequality in the US. Using the Michigan Survey of Consumers, I find that (1) subjective beliefs about future stock returns are widely dispersed, (2) optimistic households are more likely to participate in the stock market and invest more in stock, and (3) subjective beliefs are persistent over time. Motivated by these findings, I develop and calibrate a heterogenous-agent model that matches data on income inequality, beliefs distribution and portfolio choice. Compared to a model without dispersion in beliefs, the model with belief heterogeneity generates an additional 0.12 in the Gini coefficient of wealth inequality and 10 percent more wealth owned by the top 15% of the households. This suggests that the dispersion in subjective beliefs about stock returns could be an important factor in explaining the size of wealth inequality in the US.

Keywords: Subjective Beliefs, Wealth Inequality, Portfolio Choice, Heterogeneous-Agent

*Northwestern University; KwokYan.Chiu@u.northwestern.edu.
I am grateful to my advisors Matthias Doepke, Marios Angeletos, Matt Rognlie and Alireza Tahbaz-Salehi for their guidance and support. I thank Michael Cai, Jana Obradovic, Pablo Sanchez, Dalton Zhang, Evan Majic, Pooya Molavi, Cristoforo Pizzimenti, Diego Kanzig, Giorgio Primeceri, Nemanja Antic, and Xiaoyun Qiu for helpful comments. I also thank seminar participants at Macro Lunch Seminar of Northwestern University. All errors are my own.

1 Introduction

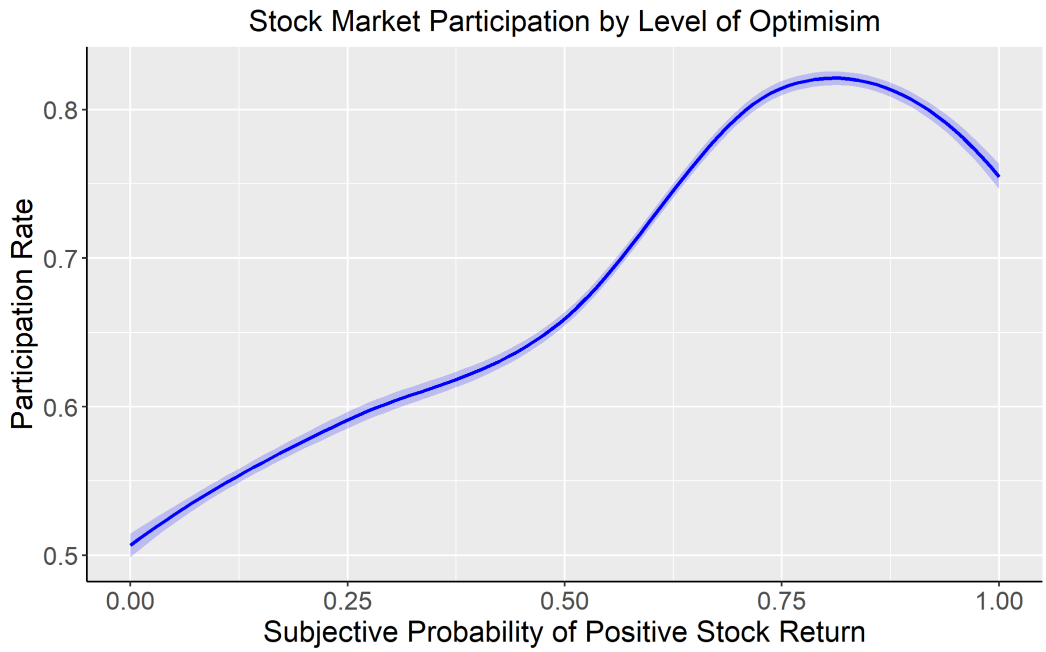
Wealth is highly concentrated in the US, with the top 10% of households owning 70% of the total wealth. Understanding the sources of wealth inequality has been a central question in economics, as it sheds light on household saving behavior and guides policy recommendations. As pointed out by De Nardi (2015), wealthier households save more than less wealthy households, leading to a higher growth rate of wealth among the wealthy. This feature cannot be explained by income differences alone. Standard incomplete market models, such as Huggett (1996), Bewley (1977), and Aiyagari (1994), which assume households save primarily to insure against uninsurable labor income risk, fail to provide sufficient saving incentives for wealthier households. Consequently, the literature has turned to other sources of heterogeneity to explain the variation in wealth accumulation rates.

While many factors have been studied to explain wealth inequality, most assume that households form full-information rational expectations (FIRE) about future returns, meaning all households share the same outlook on the stock market. Yet, survey evidence consistently challenges this assumption, revealing wide variation in households' subjective beliefs about macroeconomic and financial outcomes. Empirical studies using surveys to elicit households' beliefs, such as Vissing-Jorgensen (2003), Manski (2004), Dominitz and Manski (2007), Ameriks *et al.* (2020), Giglio *et al.* (2021) and Beutel and Weber (2023) consistently document sizable substantial cross-sectional heterogeneity in beliefs. This dispersion in subjective beliefs about future stock returns may serve as an important source of heterogeneity in wealth accumulation rates. Hubmer *et al.* (2021) demonstrates that incorporating empirical portfolio return heterogeneity from administrative tax data such as Bach *et al.* (2020) and Fagereng *et al.* (2020), generates wealth inequality levels consistent with observed wealth inequality data. A key driver of this heterogeneity in portfolio returns could be differences in households' subjective beliefs.

An optimistic household, expecting higher future returns, is likely to invest more in equities than a pessimistic household with similar characteristics. To support this, Figure 1 reveals a positive correlation between stock market participation rates and households' subjective beliefs about future stock returns. This variation in stock market participation can translate into differing wealth accumulation rates, with optimistic households accumulating wealth more rapidly than their pessimistic counterparts.

Assessing the quantitative importance of subjective beliefs in shaping wealth inequality is crucial for informing policy. Observed portfolio choice data can be rationalized by infinitely many models of

Figure 1. Stock Market Participation by Subjective Beliefs about Stock Returns



Note: Data from the *Michigan Survey of Consumers*. Subjective Probability of Positive Stock Return is measured by respondents' prediction of the probability that a diversified stock mutual fund would have a positive return in the next year. Stock Market Participation is measured by the percentage of respondents who own stocks. The solid line is the estimated conditional expectation and the shaded area represents the 95% confidence interval.

beliefs and preferences.[1] However, the policy implications differ depending on the underlying model of wealth inequality. If portfolio choices are driven solely by risk preferences heterogeneity, there is little scope for policy intervention. In contrast, if belief heterogeneity is the main driver, policies such as financial literacy programs could reshape households' expectations and potentially reduce wealth inequality.

In this paper, I argue that differences in subjective beliefs about stock returns significantly contribute to wealth inequality. I document significant dispersion of subjective beliefs and show that optimistic households are more likely to invest in the stock market. Motivated by the empirical evidence, I then develop and calibrate a heterogeneous-agent model that incorporates endogenous portfolio choice and heterogeneous subjective beliefs. Risk preferences are calibrated to match observed equity shares, while subjective beliefs are disciplined by the empirical dispersion. Compared to a model with homogeneous beliefs, my model generates an additional 0.12 in the Gini

[1] See Manski (2004) for a detailed discussion. Without expectation data, revealed preference approaches cannot distinguish between full-information rational expectations and partial information.

coefficient of wealth inequality, accounting for roughly 15% in the Gini coefficient of wealth inequality in the US.

Empirically, I show that subjective beliefs may explain large heterogeneity in investment decisions. Using data from the *Michigan Survey of Consumers*, I measure households' subjective beliefs with their reported subjective probability of a positive market return and analyze how this correlates with their investment decisions. I find that beliefs about future stock returns are widely dispersed, with some households believing future returns to be 16% above or below than the historical average. Households who reported a subjective probability one standard deviation above the mean are 10 percent more likely to participate in the stock market. This correlation holds even after controlling for education and income. Additionally, subjective beliefs are persistent over time: optimism today predicts optimism in six months. This suggests that subjective beliefs may help explain the long-term heterogeneity in investment decisions.

Quantitatively, I show that dispersion in beliefs about future stock returns significantly amplifies wealth inequality in a heterogeneous-agent model with portfolio choice. I extend a standard heterogeneous-agent model, based on Bewley (1977), which features households saving against uninsurable labor income risk, by incorporating subjective beliefs about the return on the risky asset and allowing for endogenous portfolio choice.

Optimistic households accumulate wealth more rapidly than pessimistic households through two key mechanisms. First, they perceive a higher expected return on the risky asset, leading them to allocate more of their portfolio to equities and earn the risk premium—the average return on stocks over the risk-free rate, which I calibrate at 7 percentage points based on post-war S&P500 data.[2] Second, optimistic households save more due to the higher perceived return on their portfolio, a behavior consistent with an elasticity of intertemporal substitution (EIS) greater than unity, as estimated in Vissing-Jørgensen and Attanasio (2003). Together, these mechanisms create significant disparities in wealth accumulation rates.

In the baseline model, beliefs and labor income are uncorrelated, providing a lower bound on the impact of belief heterogeneity on wealth inequality. This assumption is conservative, as data indicates that high-income households tend to be more optimistic about the stock market. Accounting for this correlation would further amplify wealth inequality between high- and low-income households.

To quantify the impact of belief heterogeneity, I conduct a counterfactual exercise where all

[2] See Mehra (2006) for a discussion of the equity premium range. I picked the moderate value of 7 percentage point, which is also commonly used in asset pricing context.

households are assigned the median belief about stock returns, holding preferences, incomes, and asset returns constant.[3] I compare the stationary wealth distributions between the baseline and counterfactual scenarios to assess the role of beliefs in shaping wealth inequality. The model with belief heterogeneity generates an increase of 0.12 in Gini coefficient of wealth inequality, corresponding to 15% of the Gini coefficient of wealth inequality in the US, and results in 10 percent more wealth owned by the top 15% of households. In a variance decomposition exercise, subjective beliefs account for one-third of the total variance in (log) wealth, highlighting the importance of belief heterogeneity in shaping wealth inequality.

While the model significantly amplifies wealth inequality compared to a standard incomplete market model, the model still falls short of matching the observed wealth inequality in the US. This is unsurprising, as the model abstracts from many factors that could contribute to wealth inequality. The baseline model fails to capture the wealth shares of the top 1% and the bottom 50% of the population. In my first extension, I follow the Krusell and Smith (1998)'s approach to allow additional heterogeneity in discount factors that are uncorrelated to beliefs and match the model to the wealth inequality data. My finding remains robust, with subjective beliefs contributing significantly to wealth inequality and the wealth share of the top 10% households halves in the counterfactual.

In my second extension, I introduce a correlation between beliefs and labor income, motivated by the empirical evidence that high-income households tend to be more optimistic about the stock market. This correlation further amplifies the effect of belief heterogeneity on wealth inequality as high-income households are more insured against the downside risk of risky assets and invest more in risky asset when they anticipate a higher return.

Related Literature

This paper offers a behavioral perspective on wealth inequality. De Nardi (2015) pointed out that understanding the motives behind the savings decisions of high-income households is crucial for explaining wealth inequality in the US. One possible explanation for high savings rates among the wealthy is heterogeneous portfolio returns, which incentivize certain households to save more than others. Apart from the Hubmer *et al.* (2021), a few papers have discussed the possibility heterogeneous portfolio returns, such as Benhabib *et al.* (2019). One paper that offers a similar behavioral perspective is Kacperczyk *et al.* (2019), which demonstrates that heterogeneous information capacity can theoretically produce wealth inequality. They calibrate their model to

[3]This is a partial equilibrium exercise or can be interpreted as an economy with an infinitely elastic supply of assets.

match aggregate moments. My approach differs by using micro-level survey data to calibrate heterogeneity in subjective beliefs, allowing us to assess how much beliefs can explain wealth inequality.

Several studies have documented the dispersion in subjective beliefs about future stock returns. Vissing-Jorgensen (2003) documents dispersion in investors one year expected stocked return. Dominitz and Manski (2007) introduced a question eliciting households' subjective probability of a positive stock market return in the next year. They argue that eliciting subjective beliefs in terms of probabilities can reduce the confusion caused by different interpretations of the term "expected return" and they find a wide dispersion in households' responses. Giglio *et al.* (2021), Beutel and Weber (2023) and Ameriks *et al.* (2020) also document substantial heterogeneity in subjective beliefs about stock returns for retail investors and find that beliefs and investment decisions are correlated. Both Das *et al.* (2020) and Dominitz and Manski (2007) have documented that beliefs are correlated to households characteristics. My paper contributes by studying the implication of this dispersion in a quantitative model of wealth inequality.

Finally, the paper contributes to the growing literature that uses micro-level survey data on subjective beliefs to calibrate macroeconomic models that depart from full-information rational expectations (e.g. Bhandari *et al.* (2019), Broer *et al.* (2021), Guerreiro (2022), and Velasquez Giraldo (2024)). With the best of my knowledge, this paper is the first quantitative exercise to see how empirical dispersion in subjective beliefs about stock returns can explain wealth inequality in the US.

Organization

In section 2, I describe the data and empirical facts on subjective beliefs and portfolio choice. In section 3, I lay out the quantitative framework for analyzing the effect of subjective beliefs on wealth inequality. In subsection 3.1 and subsection 3.3, I discuss the model and the key calibrations for my result. I present simulation results in section 4. Finally, section 6 concludes.

2 Data and Empirical Facts

In this section, I argue that subjective beliefs about stock returns are widely dispersed and they are correlated with investment decisions of households. I first describe the measurement of subjective beliefs about stock returns in the Michigan Survey of Consumers. Then, I show the heterogeneity in subjective beliefs and its relationship with investment decisions. Finally, I discuss the persistence of subjective beliefs over time that would be relevant for the model calibration.

2.1 Measuring Subjective Beliefs

The Michigan Survey of Consumers elicits subjective beliefs about the stock market of hosuehulds by asking the subjective probability of a positive stock return in the next year. The exact question is as follows:

> *What do you think is the percent chance that a one thousand dollar investment in a diversified stock mutual fund will increase in value in the year ahead, so that it is worth more than one thousand dollars one year from now?*

Responses of this question informs us about households' subjective beliefs about the short-term stock market returns. I consider this subjective probability as an measurement of optimism, the higher the reported probability, the more optimistic the household is about stock returns. This is different from measuring stock-picking skills or risk-aversion, which are both important for the household's investment decision. The question specifically asks for a performance prediction for a "diversified stock mutual fund", which excludes the stock-picking skill of the respondent. The question also asks for a subjective probability of an event, which does not depend on the risk-aversion of the respondent.

2.2 Heterogeneity in Subjective Beliefs

Households disagree on the prospect about the stock market returns. In this section, I show the distribution of subjective beliefs about stock returns and discuss how the dispersion in beliefs could be interpreted as the dispersion in the subjective risk-adjusted return.

As shown in Figure 2, households disagree on the probability of positive stock market return, with considerable shares of respondents reporting 100% or 0% chance of positive stock return. This dispersion is mostly driven by cross-sectional differences. Even after demeaning the data by time average, some hosueholds reported $+50\%$ or -50% probability relative to the average belief at that time. There is a mass of households reporting probability exactly at 50%. One concern is that households may not have a clear idea about the stock market return and they may report 50% as a default answer. Households who answered 50% throughout their presence in the survey constitute around 3% of the sample. This is a small share of the sample and the results are robust to excluding these households[4].

Under additional assumptions, the subjective probability of positive stock return can be interpreted as the subjective risk-adjusted return, which is more accessible for macro-finance models. Suppose households believe that the stock return follows a log-normal distribution[5], with subjective mean return μ_{it} and variance $(\sigma_{it})^2$. Let $R_{t,t+1}$ be the annual stock return

$$R_{t,t+1} = \exp(r_{t,t+1}) \quad \text{with} \quad r_{t,t+1} \sim N(\mu_{it}, (\sigma_{it})^2) \tag{1}$$

The subjective probability of positive stock return is a strictly increasing function of the subjective risk-adjusted return defined as μ_{it}/σ_{it}, subjective mean return divided by standard derivation of return[6]. The exact expression is given by

$$\underbrace{P_{it}(R_{t,t+1} \geq 1)}_{\text{Subjective Probability of Positive Stock Return}} = 1 - \Phi\Big(- \underbrace{\frac{\mu_{it}}{\sigma_{it}}}_{\text{Subjective Risk-adjusted Return}} \Big) \tag{2}$$

where $\Phi(.)$ is the cumulative distribution function of a standard normal random variable.

As shown in Figure 3, the risk-adjusted return is also widely dispersed, with some households believing a risk-adjusted return between -1.5 to 1.5. Suppose the subjective standard deviation of return is same as the post-war historical standard derivation of the S&P 500 of 0.15, household's subjective beliefs on the mean annual return ranges from -22.5% to 22.5%, reflecting a wide disagreement on the stock market return. This transformation also facilitate comparisons to historical data. The risk-premium and standard derivation for post-war US market is estimated to be around $E[r - r^f] = 7\%$ and 15% respectively[7]. For a risk-free rate of 1%, the risk-adjusted return according to the historical data is around 0.53 and the probability of positive stock return is around

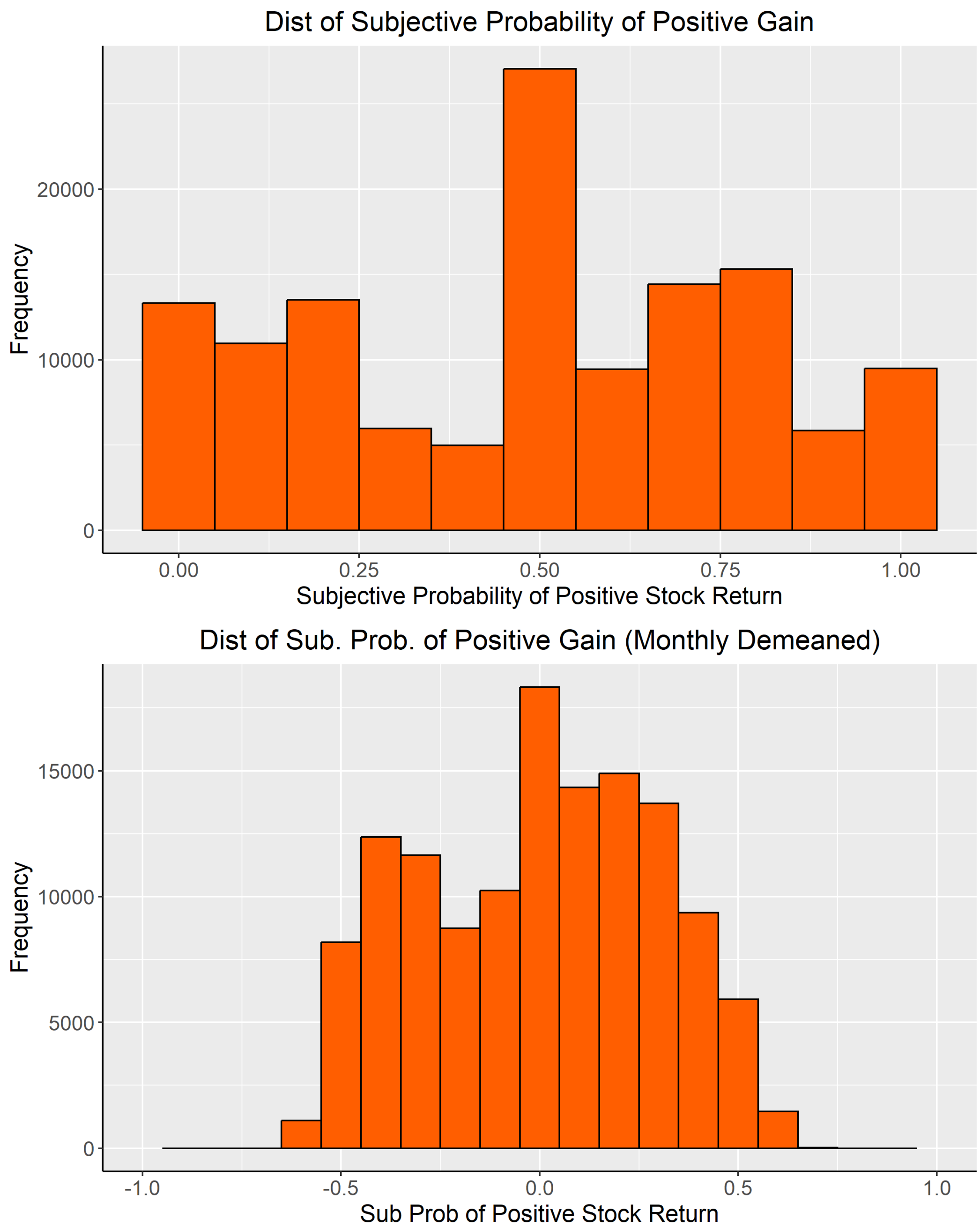
[4]See Appendix D for more robustness checks

[5]A similar transformation is suggested in Dominitz and Manski (2007). They proposed using normal distribution instead. I chose log-normal distribution to be consistent with most of the macro-finance literature.

[6]Notice that this is different from Sharpe ratio because the numerator is the expected rate of return instead of the excess return

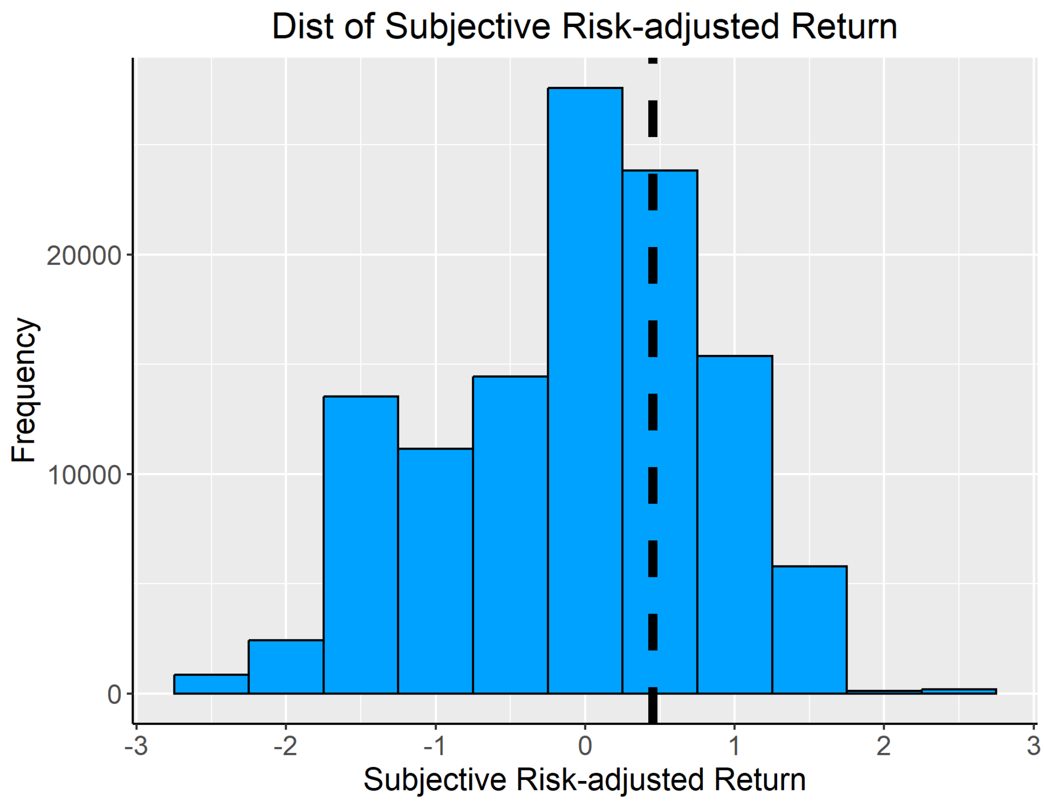
[7]This is in line with Mehra and Prescott (1985)

Figure 2. Distribution of Subjective Beliefs about Stock Returns



Note: Data from *Michigan Survey of Consumers.* Subjective Probability of Positive Stock Return is measured by respondents prediction of the probability that a diversified stock mutual fund would have a positive return in the next year. Raw data (top panel) and Demeaned data (bottom panel) are shown.

Figure 3. Distribution of Subjective Beliefs about Stock Returns



Subjective risk-adjusted return is defined as subjective mean of return divided by subjective standard deviation of return. It is calculated by transforming subjective probability, $\mu_i/\sigma_i = -\Phi^{-1}(1 - P_i(R \geq 1))$ using Equation 2. The histogram shows the distribution of subjective risk-adjusted return and the dashed line shows the risk-adjusted return according to the historical data.

0.65. Hence, households are slightly more pessimistic than the historical data.

2.3 Subjective Beliefs and Investment Decisions

Beliefs elicited in the survey are not just noise, but they are relevant for the investment decisions of households. Subjective beliefs about stock returns are largely associated with individual investment decisions. Optimistic households are likely stock investors and they invest significantly more in stock. Subjective beliefs seem to have a good explantory power for the investment decisions of households, even after controlling for other household characteristics.

Regression Setup

The following regression extracts the households investment decision associated at each level of subjective beliefs.

$$Y_{it} = \alpha_t + \sum_{i \in g} \alpha_g \times I(i \in g) + \beta \times \text{Belief}_{it} + \epsilon_{it} \tag{3}$$

In this setup, Y_{it} is the outcome investment decision for household i, which is either a binary indicator of *stock market participation* or the *log investment share*, defined as the logarithm of total amount invested in stock divided by current income. Belief_{it} is the subjective probability of positive stock market returns reported by household i at time t. Results are robust to using subjective risk-adjusted returns μ_i/σ_i as the regressor and they are presented in Appendix D[8]. β measures the investment decisions associated with different subjective beliefs. α_t is the time fixed effects to control for movements in aggregate investment and aggregate beliefs over the business cycle. α_g is the group fixed effects to control for the observable characteristics of the households, which include income quintile and whether the respondent has a college degree.

Before presenting the results, it is worth discussing the variations being captured in this regression. After adding time and group fixed effects, the regression captures the cross-sectional variation in investment decisions associated with subjective beliefs, *within* the same time, income quintile and education group. As pointed out by Das *et al.* (2020), optimism is associated with higher income and education attainment. The difference in optimism *across* group may drive the difference in investment decisions across group. Those effects cannot be captured in a regression analysis. In the extension, I consider the correlation between subjective beliefs and other household characteristics in the calibration to account for their effect on wealth inequality.

Empirical Findings

On the extensive margin, stock market participation is associated with subjective beliefs about stock returns. In Table 1, the stock market participation rate increases by 33% from the households who believe in 0% chance of positive stock return to the households who believe in 100% chance of positive stock return. The standard deviation of subjective probability is around 0.3. Thus, one standard deviation increase in optimism is associated with $0.3 \times 33\% \approx 10\%$ increase in the

[8] Response with 0% and 100% for subjective beliefs are removed for the analysis using subjective risk-adjusted returns. This is because they correspond to $-\infty$ and ∞ subjective risk-adjusted returns. A regression could not be performed.

stock market participation rate. The association mostly comes from the cross-sectional differences. Once individual fixed effects are added, the association is much less pronounced because all cross-sectional differences are absorbed by the fixed effects. Education and income cannot fully explain the association between subjective beliefs and stock market participation, as the regression coefficients remains significant after controlling for these variables. One standaard deviation increase in optimism still accounts for roughly 5% increase in the stock market participation rate.

On the intensive margin, stock investment conditioned on investing is also associated with subjective beliefs about stock returns. In Table 2, investment in stock increases by 33% from the households who believe in 0% chance of positive stock return to the households who believe in 100% chance of positive stock return. One standard deviation incrase in optiism (≈ 0.3) is associated with 10% increase in the stock investment conditional on investing. The regression coefficient remains large but less significant after controlling for income and education fixed effect. This is because low-income and non-college educated households invest very little in stock. Subjective beliefs have a very small effect on the intensive margin of these households. In Appendix D, the regression coefficients for each income and education group are presented. The investment share is associated strongly with the subjective beliefs for higher income and college-educated households.

Table 1. Stock Market Participation on Beliefs

Dependent Variable:	Participation in Stock Market					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Sub. Prob of Positive Return	0.331***	0.022***	0.334***	0.182***	0.257**	0.160***
	(0.004)	(0.005)	(0.006)	(0.021)	(0.020)	(0.019)
Individual FE		Yes				
Time FE			Yes	Yes	Yes	Yes
Income Group FE				Yes		Yes
College FE					Yes	Yes
Observations	129,841	129,841	129,841	122,586	129,274	122,172
Adjusted R^2	0.04546	0.75055	0.04940	0.26458	0.12880	0.28105

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: Participation in Stock Market is a binary indicator of whether the individual has stock market investment. The subjective probability of positive stock return is the individual's subjective probability of positive stock market return. Income group fixed effects are indicators for income quintiles. Education fixed effects are indicators for college-educated or not.

The regression analysis so far focuses on the correlation of subjective beliefs and investment

Table 2. Investment Share on Beliefs

Dependent Variable:	Log Investment to Income					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Sub. Prob of Positive Return	0.335***	0.100***	0.303***	0.263**	0.247	0.227*
	(0.019)	(0.019)	(0.020)	(0.086)	(0.093)	(0.082)
Individual FE		Yes				
Time FE			Yes	Yes	Yes	Yes
Income Group FE				Yes		Yes
College FE					Yes	Yes
Observations	71,845	71,845	71,845	71,845	71,723	71,723
Adjusted R^2	0.00425	0.83148	0.01680	0.02665	0.02445	0.03184

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

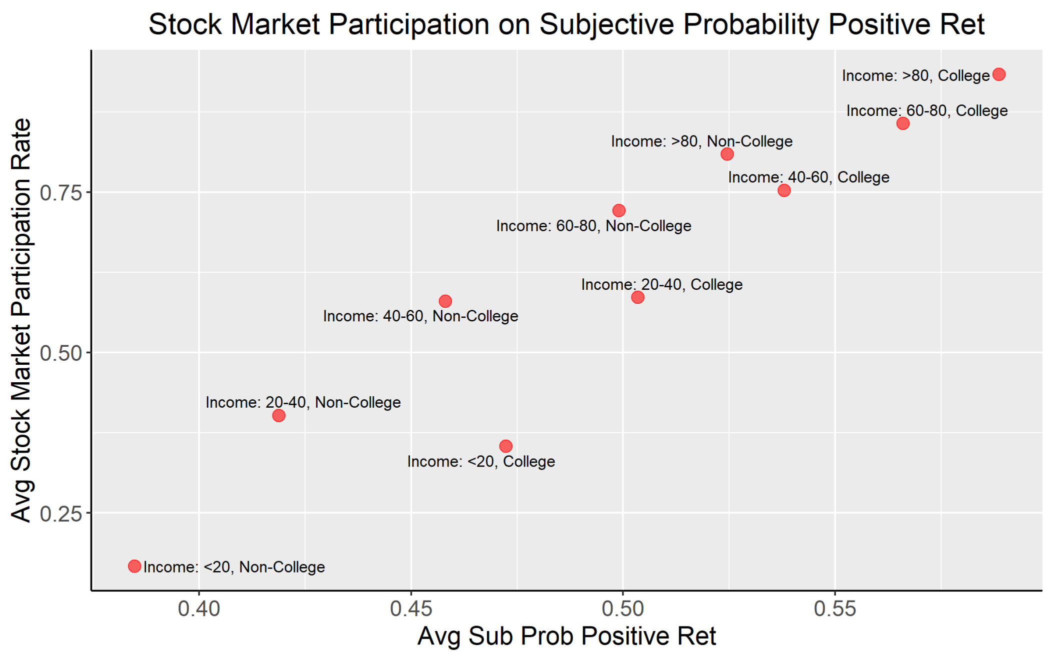
Note: Invest Share is defined as the logarithm of stock investment amount divided by current income, log(Stock Amt Invested / Income) for households with stock investment only. The subjective probability of positive stock return is the individual's subjective probability of positive stock market return. Income group fixed effects are indicators for income quintiles. Education fixed effects are indicators for college-educated or not.

decisions *within* the same time, income quintile and education group. High-income college-educated households are more likely to invest in stock and they are more optimistic about the stock market on average. In Figure 4, the differnce in subjective probability between group is about 20 percentage points, while the average stock market participation varies from 25% to 80%. The effect of subjective beliefs on investment decisions could not be analyzed in the regression analysis. I account for this correlation in the calibration in the extension.

2.4 Persistent Difference of Subjective Beliefs

Households tend to hold their subjective beliefs about stock returns over time. In this section, I estimate the autocorrelation of the subjective beliefs and show that households who are optimistic today are likely to be optimistic in the future. This suggests that subjective beliefs may explain the long-term differences in portfolio choices across households.

Figure 4. Stock Market Participation and Subjective Beliefs Accross Groups



Note: Avg Sub PRob Positive Ret is the average subjective probability of positive stock return of the households in the group. Stock Market Participation is the stock market participation rate of the households in the group. The data is from *Michigan Survey of Consumers.*

Autoregressive Coefficient

The autoregressive coefficient of the subjective beliefs is estimated by the following regression

$$\text{Belief}_{it} = \alpha_t + \rho \times \text{Belief}_{it-1} + \epsilon_{it} \tag{4}$$

where Belief_{it} is the subjective probability of positive stock return reported by household i at time t. Belief_{it-1} is the subjective probability reported six months before. α_t is the time fixed effects. Because of the time fixed-effect, ρ measures the autoregressive coefficient of the subjective beliefs relative to the average. A positive coefficient indicates that a household who is optimistic today relative to the average is likely to be optimistic relative to the average in the future.

In Table 3, the autoregressive coefficient is estimated to be 0.43. A household who is one standard deviation more optimistic than the average is likely to be 0.43 standard deviation more optimistic than the average in the future. This suggests that the disagreement between households does not resolve instantly. One concern is that this coefficient may be driven by the uninformed households who report 50% probability as a default answer in two consecutive survey periods. Those households contribute to roughly 3% of the sample. The results are robust to excluding these households[9].

Table 3. Beliefs Persistence

Dependent Variable:	Sub. Prob of Positive Return	
Model:	(1)	(2)
Sub. Prob of Positive Return 6m lag	0.433***	0.416***
	(0.004)	(0.005)
Time FE		Yes
Observations	44,941	44,941
Adjusted R^2	0.19212	0.21360

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: The subjective probability of positive stock return is the individual's subjective probability of positive stock market return. Sub. Prob of Positive Return 6m lag is the belief reported by the same household six mouths before.

[9] See Appendix D for more robustness checks

3 Quantitative Framework

In this section, I present a quantitative model to study how subjective beliefs about stock returns influence wealth inequality. Building on the framework of Bewley (1977), I developed a heterogeneous-agent model to include subjective beliefs and allow for portfolio choices between risk-free and risky assets. Households with optimistic beliefs perceive higher expected returns on risky assets, prompting them to save more and allocate a larger share of their portfolio to these assets. This dispersion in beliefs creates differences in portfolio returns and saving rates, leading to varying rates of wealth accumulation and, ultimately, wealth inequality.

The quantitative model is calibrated to match the empirical distribution of subjective beliefs about stock returns, the aggregate asset-to-income ratio, and the labor income process in the US. The effect of beliefs on saving and portfolio choice depends on two key parameters: the elasticity of intertemporal substitution (EIS) and the relative risk aversion (RRA). The EIS is calibrated based on estimates from the literature, while the RRA is chosen to match the empirical equity share observed in the data.

Lastly, I consider model variations to examine interactions between belief heterogeneity and other factors influencing wealth inequality. First, I introduce additional heterogeneity in the discount factors to account for wealth inequality unexplained by the baseline model. High discount factor households, being more patient, respond more strongly to changes in beliefs. Second, I consider an alternative calibration that allows for a positive correlation between beliefs about stock returns and permanent income. High permanent income households are less exposed to downside risk in the stock market, making them more responsive to changes in beliefs.

3.1 Model

The model consists of two main components: the asset environment, which determines the returns of the risky asset, and the household's problem, which captures portfolio choice and saving behavior. Actual asset returns are determined by aggregate states and apply uniformly to all households. However, the perceived return of the risky asset depends on households' subjective beliefs. Throughout this paper, I focus on a partial equilibrium framework, where actual asset returns are independent of households' beliefs.[10]

Households maximize an Epstein-Zin utility function, subject to budget constraints and

[10] This assumption can be interpreted as a small-open economy where both risky and risk-free assets are in perfectly elastic supply.

borrowing, short-selling, and leveraging limits, given their perceived returns on the risky asset. This optimization yields households' portfolio choices and saving decisions. Combining the asset environment and the policy function, the model generates a stationary wealth distribution.

Aggregate State and Stock Returns

There is an aggregate state γ_t, that is independently and identically distributed over time. The state can take two values, bad state γ^B and good state γ^G, with $\gamma^B < \gamma^G$, with probability $p(\gamma_t)$. This aggregate state can be interpreted as fluctuations in annual economic growth, affecting the returns of the risky asset. Households observe the current state γ_t and its past realizations but are uncertain about future states, γ_{t+k} for $k \geq 0$.

The aggregate state determines the return of the risky asset from time t to $t+1$, The return of the risky asset from time t to $t+1$ is given by an increasing function $R_{t,t+1} = R(\gamma_{t+1})$. Good state γ^G is associated with higher returns than the bad state γ^B, i.e., $R(\gamma^G) > R(\gamma^B)$.[11]

In addition, households have access to risk-free asset with an exogenous risk-free rate R^f. The risk-free rate is assumed to be constant over time and independent of the aggregate state.

The values of $R(\gamma^G)$ and $R(\gamma^B)$ are calibrated to match the behavior of the *S&P500*. The risk premium of the risky asset is given by

$$E_t[R_{t,t+1}] - R^f \tag{5}$$

and the variance is given by $Var_t(R(\gamma_{t+1}))$. The model is calibrated to have a positive risk premium, which makes stock investor earn more than risk-free asset investors on average.

Household

There is a mass one of households, which are heterogeneous in terms of financial wealth W_t, labor income Y_t, and subjective beliefs Π_t. Each household chooses consumption C_t and the share of the risky asset ω_t to maximize their value function. The subjective beliefs Π_t affect the perceived distribution of the risky asset return. In this paper, I consider Π_t to represent households with

[11] This discrete return model can be seen as a discretization of a continuous return model. Suppose the stock return follows a log-normal distribution with mean μ and variance σ^2, $R_{t,t+1} = \exp(r)$ and $r \sim N(\mu, \sigma^2)$. Discrete points are picked to approximate the log-normal distribution. In my implementation, I approximate the normal distribution of log return and then transform the points with an exponential function. The risk premium is approximated by $\mu + \frac{1}{2}\sigma^2 - r^f$, with $r^f = \log(R^f)$

permanent type of beliefs about the mean of stock returns. In general, Π_t could be used for more general heterogeneity in beliefs.[12]

In the extension, I also consider an additional heterogeneity in the permanent type of households to capture heterogeneity in discount factors or permanent income levels. This additional heterogeneity is denoted by Ξ and it is assumed to be unchanged over time.

Maximization Problem

I first outline the maximization problem for the baseline model without additional permanent types. The household's problem is to maximize an Epstein and Zin (1989) utility function by optimally choosing consumption C_t and risky share ω_t. The value function $V(W_t, Y_t, \Pi_t)$ is given by

$$V(W_t, Y_t, \Pi_t) = \max_{C_t,\omega_t} \left((1-\beta)C_t^{1-1/\psi} + \beta\Big(CE_t[V_{t+1}]\Big)^{1-1/\psi} \right)^{1/(1-1/\psi)} \tag{6}$$

where V_{t+1} denotes the value in the next period, $V_{t+1} = V(W_{t+1}, Y_{t+1}, \Pi_{t+1})$ and $CE_t[V_{t+1}]$ denotes household certainty equivalent operator at time t. The certainty equivalent of future value is given by

$$CE_t(V_{t+1}) = \left(E\left[\Big(V_{t+1}(W_{t+1}, Y_{t+1}, \Pi_{t+1})\Big)^{1-\theta} |\Pi_t, Y_t\right]\right)^{\frac{1}{1-\theta}} \tag{7}$$

The elasticity of intertemporal substitution (EIS) is denoted by ψ, the relative risk aversion (RRA) is denoted by θ, and the discount factor is denoted by β. If $\theta = 1/\psi$, the model reduces to the standard Constant Relative Risk Aversion (CRRA) utility function. The reason for using Epstein-Zin utility is to separate the portfolio choice from the saving decision as both could be affected by the beliefs about the returns of saving. The discount factor β determines the patience of the households and thus the size of saving. This is used for matching the aggregate asset-to-income ratio in the data, which will be discussed in the calibration section.[13]

Households are subject to a budget constraint and a borrowing constraint. Households' financial wealth is given by saving from the previous period times the portfolio return, which is given by the

[12]This is a general setup to allow for heterogeneous beliefs. In a signaling problem, Π_t could be a sequence of past signals. This can also include parameters for learnings and other models in beliefs.

[13]In a model with idiosyncratic income risk and risk-free asset, households would want to save infinite amount of asset as long as $\beta R^f \geq 1$ to perfectly insure from the borrowing constraint. Since risky asset provides a risk premium, a stricter requirement is needed. Throughout this paper, no explosive solution is considered.

portfolio choice ω_t. The budget constraint is given by

$$W_{t+1} = (W_t + Y_t - C_t)\left(R_{t,t+1}\omega_t + R^f(1-\omega_t)\right) \tag{8}$$

and the borrowing constraint is given by

$$W_{t+1} \geq 0 \tag{9}$$

The borrowing constraint is standard in the literature for incomplete market model, which serves as a basic component to wealth inequality by encouraging households to save out of the borrowing constraint. In addition, households faces constraints on short-selling and leveraging

$$0 \leq \omega_t \leq 1 \tag{10}$$

The short-sale constraint and no leveraging constraint are motivated by the fact that households rarely take those extreme positions in the data. The short-sale constraint generates non-participation in the stock market for the pessimistic households. Very pessimistic households may want to short the stock, but the constraint prevents them from doing so.

In the Survey of Consumer Finances (SCF) data, the equity share of households is zero percent for more than half of the households. In a portfolio choice problem, for example, Merton (1969), households participate in the stock market as long as the risk-premium is positive. It is standard for literature that assume rational expectation to include a fixed cost of participation to generate non-participation in the stock market.[14] This participation cost is less important for this model to generate a large non-participation. As discussed in the data section, households are more pessimistic about the stock market than the actual return. This turns out to be sufficient to generate a large non-participation in the stock market in the calibrated model.

Labor Income Process

In this section, I describe the labor income process for the baseline model. The idiosyncratic risk generates labor income inequality, which serves as one source of wealth inequality. On top of this, the idiosyncratic labor income risk is a background risk that affects the portfolio choice of the households in an incomplete market model.

[14] For example, Campbell *et al.* (1999). See more discussion in Heaton and Lucas (2000)

I model the labor income as an exogenous endowment. This could be interpreted as household supplying inelastic labor and with an exogenous process for the productivity.

The idiosyncratic labor income Y_t and aggregate state γ_t follows a Markov process. The joint transition probability for (Y_t, γ_t) is given by $P(Y_t, \gamma_t | Y_{t-1}, \gamma_{t-1})$. In general, the distribution of the idiosyncratic shock could depend on the aggregate state, but the realization of the idiosyncratic state should not affect the aggregate state in the future. Following the setup in Krusell and Smith (1998), the aggregate state and idiosyncratic state are independent conditional on past aggregate state and idiosyncratic state and

$$\begin{aligned} P(Y_t, \gamma_t | Y_{t-1}, \gamma_{t-1}) &= P(Y_t | Y_{t-1}, \gamma_{t-1}) P(\gamma_t | Y_{t-1}, \gamma_{t-1}) \\ &= P(Y_t | Y_{t-1}, \gamma_{t-1}) P(\gamma_t | \gamma_{t-1}) \end{aligned}$$

The first equality imposes conditional independence and the second inequality imposes that idiosyncratic shocks do not affect future evolution of the aggregate state.

In the main simulation, I assume that idiosyncratic risk Y_t depends only on Y_{t-1}. Hence, $P(Y_t | Y_{t-1}, \gamma_{t-1}) = P(Y_t | Y_{t-1})$. The transitional probability and the income states are calibrated to match the autocorrelation and the variance of the US labor income data.

This abstracts from empirical findings that the distribution of labor incomes varies over the business cycle. Guvenen *et al.* (2014) shows that the labor income shock in the data tends to have a long left-tail during the downturn of the economy. McKay (2017) shows that this feature is important for the consumption dynamics of the households. This is less important for comparing the long-run wealth inequality induced by the subjective beliefs. Catherine (2022) shows that this feature is important to replicate the low stock market participation rate in the data because it generates additional correlation between stock returns and wealth growth rate. This is also important for my purpose as the difference in beliefs about stock returns already generates a large non-participation in the stock market. In Appendix B, I show that the quantitative results are robust to an alternative specification.

Beliefs

Households' heterogeneous beliefs are captured by the state variable Π_t. In the household's maximization problem, they need to forecast both the future labor income and the future stock return for the consumption-saving and portfolio decision. I focus on the expectation of the stock return in this paper and assume that households have rational expectations about the future labor income.[15]

Households' subjective beliefs about stock returns affect their subjective mean return of the risky asset.[16] I assume that households' subjective beliefs about stock returns are given by a scaling factor on the objective stock return. For belief type Π_t, the perceived return of the risky asset is given by

$$R_{t,t+1}(\Pi_t) = \exp(\lambda(\Pi_t)) \times R_{t,t+1} \tag{11}$$

The factor $\exp(\lambda(\Pi_{it}))$ captures the heterogeneity in subjective beliefs about the expected log return of the stock as the mean of the log return is shifted by $\lambda(\Pi_{it})$ while holding the perceived risk constant.

$$E_t[\log(R_{t,t+1})|\Pi_t] = \lambda(\Pi_{it}) + E_t[\log(R_{t,t+1})] \quad \text{and} \quad Var_t[\log(R_{t,t+1})|\Pi_t] = Var_t[\log(R_{t,t+1})]$$

In the data, households report their subjective probability of positive stock return. I assume that their perceived stock return is log-normally distributed and the subjective variance of the stock return is objective $\sigma_{it} = \sigma$ for all households. For subjective probability of positive return reported by households i at time t, $P_{it}(R_{t,t+1} \geq 1)$, the mean of the log return is given by

$$\mu_{it} = -\sigma\Phi^{-1}(1 - P_{it}(R_{t,t+1} \geq 1))$$

where Φ^{-1} is the inverse of the standard normal CDF. Then, the scaling factor for household i is given by

$$\mu_{it} - E_t[\log(R_{t,t+1})]$$

In the implementation of the model, I divide the subjective beliefs into three types: pessimistic, moderate, and optimistic, with three different scaling factors. The scaling factor $\lambda(\Pi_t)$ is calibrated to match the distribution of the subjective beliefs in the data.

[15] In the data, households have heterogeneous beliefs about the future labor income too. Households who are optimistic about the future stock returns are also likely to be optimistic about the future labor income. This could be integrated into this framework by considering a subjective belief about the future labor income.

[16] Literature in behavioral finance tends to focus on the subjective mean return of the stock. Giglio *et al.* (2021) finds that subjective variance does not seem to affect the equity share in the data. Moreover, continuous time models in macro-finance often assume a geometric Brownian motion for stock price, which implies that the variance of the stock return could be inferred with only a short period of data. For this paper, disagreement on subjective mean or subjective variance would not matter as the risk aversion would be pciked to match the equity share in SCF. Both model would yield the same equity share after calibration.

Additional Permanent Types

In the wealth inequality literature, other forms of heterogeneity are purposed to explain the wealth inequality data. Those factors may interact with households subjective beliefs. In the extension, I consider additional permanent types heterogeneity. This additional heterogeneity is denoted by Ξ, which is assumed to be unchanged over time.

Discount Factor Heterogeneity While the model has enhanced portfolio choices and beliefs compared to the standard model, my model still does not capture all the wealth inequality in the data, especially on the bottom and the top wealth distribution. I consider an alternative benchmark where households have different permanent types Ξ that are uncorrelated with the subjective beliefs. Following Carroll *et al.* (2017), I assume an additional heterogeneity in the discount factor β. High β households are more patient and save more. Heterogeneity in β helps generate additional disparity in wealth accumulation and thus wealth inequality.

Households with permanent type Ξ maximizes the following objective function

$$V(W_t, Y_t, \Pi_t, \Xi_i) = \max_{C_t,\omega_t} \left((1-\beta(\Xi))C_t^{1-1/\psi} + \beta(\Xi)\Big(CE_t[V_{t+1}]\Big)^{1-1/\psi} \right)^{1/(1-1/\psi)} \tag{12}$$

where $\beta(\Xi)$ specifies a distribution of discount factor β for households with permanent type Ξ. The distribution of $\beta(\Xi)$ are calibrated to match wealth shares of the bottom and the top wealth groups in the data. The calibration of the model is discussed in the next section.

Permanent Income Heterogeneity This extension address the potential correlation between beliefs about stock returns and permanent income. In the data, high-income and educated households are more optimistic about the stock market. I consider an alternative calibration where there are Ξ groups with different permanent labor income levels. The budget constraint is given by

$$W_{t+1} = (W_t + \tilde{Y}(\Xi) + Y_t - C_t)\Big(R_{t,t+1}\omega_t + R^f(1-\omega_t)\Big) \tag{13}$$

with $\tilde{Y}(\Xi)$ being the permanent labor income of households for type Ξ.

3.2 Partial Equilibrium Stationary Distribution

In most of the analysis, I focus on the partial equilibrium stationary distribution of the model given the risk-free rate R^f and the $R_{t,t+1}$. The goal is to understand how changing distribution of $\lambda(\Pi_{it})$ affects the wealth distribution in the stationary distribution.

Given the risk-free return R^f, state variables $(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$, the optimum portfolio choice and saving decision of the households are given by

$$\omega_{it} = \omega(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$$
$$X_{it} = X(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$$

The financial wealth next period is given by their savings times their portfolio returns

$$W_{it+1} = X(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)\left(R^f + \omega(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t) \times (R_{t,t+1} - R^f)\right)$$

which pins down the (stochastic) evolution of the financial wealth. Subjective beliefs Π_{it} affect the wealth distribution through both the portfolio choice and the saving decision.

A distribution $D(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$ specifies the probability mass distribution of households in the state of $(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$. I define the stationary distribution and partial equilibrium as follows

Definition 1. (Stationary Distribution) Given portfolio decision $\omega(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$ and saving decision $X(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$, a distribution $D(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$ is stationary if the distribution satisfies the following conditions

$$\begin{aligned} &D(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t) \\ &= \sum_{\Xi_i} \int_{(W_{it-1}, Y_{it-1}, \Pi_{it-1}, \Xi_i, \gamma_t)|W_t = A_{it}} \int_{Y_{t-1}, \gamma_{t-1}} \left[P(Y_t, \gamma_t | Y_{t-1}, \gamma_{t-1}) \times D(W_{it-1}, Y_{it-1}, \Pi_{it-1}, \Xi_i, \gamma_t) \right] \end{aligned} \tag{14}$$

where A_{it} is the function of realized financial wealth at state γ_t, defined as

$$A_{it} = X(W_{it-1}, Y_{it-1}, \Pi_{it-1}, \Xi_i, \gamma_{t-1})(R^f + \omega(W_{it-1}, Y_{it-1}, \Pi_{it-1}, \Xi_i, \gamma_{t-1}) \times (R(\gamma_t) - R^f))$$

Definition 2. (Partial Equilibrium Stationary Distribution) Given the risk-free rate R^f and the risky return $R_{t,t+1}$, the distribution $D(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$ is in partial equilibrium stationary distribution if the distribution satisfies the following conditions

1. The distribution $D(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$ is stationary
2. The decision rules $\omega(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$ and $X(W_{it}, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$ solve the households' maximization problem

Note that in this definition of stationary distribution, the realized return of the risky asset $R_{t,t+1}$ evolves according to the objective return of the risky asset $R(\gamma_t)$. The subjective beliefs Π_{it} only affect the decision of the households but not the evolution of the risky return. Finally, the wealth distribution $P_W(w)$ in the stationary distribution is defined as

$$P_W(w) = \sum_{\Xi_i} \sum_{\gamma_t} \int_{Y_{it}, \Pi_{it}} D(w, Y_{it}, \Pi_{it}, \Xi_i, \gamma_t)$$

3.3 Calibration

My model is calibrated to replicate the US economy. I calibrate the labor income process to capture the US income dispersion and cyclical properties of labour income. I choose the risk aversion parameter θ to match the top-wealth equity share in the *Survey of Consumer Finance*. Following the calibration strategy in Aiyagari (1994), the discount factor β is chosen to match the aggregate asset to income ratio.

Labour Income Process

The labor income process is calibrated to produce realistic income dispersion and the cyclical properties of labour income change. Suppose the labor income is given by $Y_t = \exp(y_t)$, the transitional probability of the labor income is designed to replicate the following $AR(1)$ process for the log income. Let y_t be the log income at time t, the income process is given by

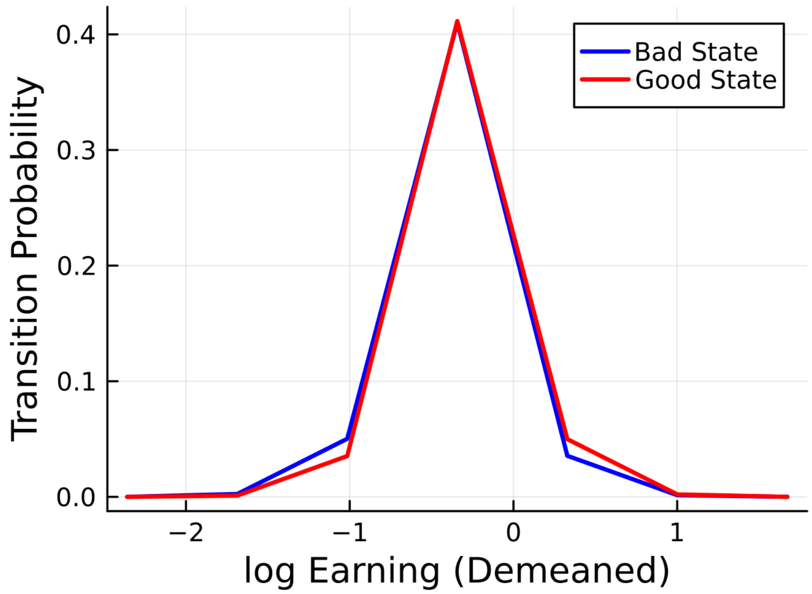
$$y_t = \rho y_{t-1} + \eta_{it} \tag{15}$$

The $AR(1)$ process of the log labour income is characterized by the persistence ρ and the distribution of the innovation, which are calibrated to the annual data. The persistence is set to 0.93, corresponding to a quarterly persistence of 0.975 estimated by French (2005)[17]. The variance of the innovation is chosen to match the standard deviation of the log labour income of 0.83 documented by Guvenen *et al.* (2014), which implies a variance of $Var(\eta_{it}) \approx 0.1$. This variance is set to be constant throughout the fluctations of the aggregate state γ_t. This is consistent with Guvenen *et al.* (2014)'s findings that the variance of the innovation is roughly constant over the business cycle.

A mixture-normal innovation is defined by the mixture probability p_{adj}, the means and variances of the two normal distributions. The probability of adjustment is set to 0.136, used in Catherine (2022) to match the probability of major income changes. The means and variances of the two normal distributions are chosen to match the mean, variance, 10th, 50th and 90th percentile of log income change along the business cycle.The exact procedure of estimation and discretization using Rouwenhorst Method is described in Appendix G.

[17] Guvenen *et al.* (2014) estimate a much higher autocorrelation in the Social Security records. I choose a smaller number instead to ensure that the model remains stationary

Figure 5. Transitional Probability for Median Income Household



Note: The figure shows the transitional probability of the median income household. The probability is calibrated to match the countercyclical negative skewness of income change.

Table 4. Calibrated Parameters for Labour Income Process

Parameter	Description	Value	Sources
ρ	Persistence	0.93	Quarterly persistence of 0.975 in French (2005)
$sd(y_{it})$	Standard Deviation of log income	0.83	Standard Deviation of log wage in Guvenen *et al.* (2014)
p_{adj}	Mixture Probability	0.136	Estimate from Catherine (2022)
$(\mu^{adj}(\gamma^G), \mu^{nonadj}(\gamma^G))$	Conditional Mean in Expansion	$(0.103, 0.01)$	To match Mean, p10, p50, p90 of log wage growth in 05-07
$(\mu^{adj}(\gamma^B), \mu^{nonadj}(\gamma^B))$	Conditional Mean in Recession	$(0.15, -0.05)$	To match Mean, p10, p50, p90 of log wage growth in 08-09

Assets

The risk-free rate is set to 1 percent, a value commonly used in the literature for annual frequency calibration. The risky asset is calibrated to replicate S&P500's returns in Post-war era. The risky premium is chosen to be 7% according to Mehra and Prescott (1985) and the standard deviation of the risky asset is set to 15%, which produce a Sharpe ratio of 0.47.

Beliefs

To capture the heterogeneity in subjective beliefs, I assume that there are three types of beliefs about the future stock returns, optimistic, median and pessimistic, with equal weights. In the data, the standard deviation of the subjective risk-adjusted return is around 0.8. The optimistic and the pessimistic groups are assumed to have $\sqrt{3/2} \approx 1.22$ standard derivation above average and below average beliefs, respectively. This ensures that the overall standard deviation of the subjective beliefs is 0.8. The median households report a zero probability of positive return, corresponding to a 0 subjective risk-adjusted return[18].

In the baseline calibration, the beliefs are assumed to be permanent and uncorrelated with other non-belief characteristics. The extension to cover both are discussed in the extension section.

Risk Aversion and Elasticity of Intertemporal Substitution

The risk aversion parameter is calibrated to match the top-wealth equity share in the data. In the data, the top 1% wealthy households holds roughly 30% of their wealth in equity. Since optimistic households are more likely to be wealthy due to their higher equity share, the risk aversion parameter is chosen such that the wealthy optimistic households hold 30% of their wealth in equity. The risky share of the top wealth households can be approximated using a method in Campbell and Viceira (2001), $\omega \approx ((\max_i \mu^{it}) - r^f)/(\theta\sigma^2)$, which implies a risk aversion parameter of $\theta = 18$. The elasticity of intertemporal substitution (EIS) is set to $\psi = 1.5$, a value used in the asset pricing literature. It implies that optimism has a positive effect on the saving rate. This is consistent with my finding that the amount of stock investment is positively correlated with the subjective beliefs about future stock returns.

Discount Factor

The discount factor β is calibrated such that the model matches the aggregate asset to income ratio in the data. In 2022, the household net wealth is around \$140 trillion dollars and the compensation of employees is around \$14 trillion dollars[19]. The discount factor is chosen to match the asset to income ratio of 10. The calibrated discount factor for the model with heterogeneous belief is $\beta = 0.92$, which is in line with the literature. The discount factors of all single-β models are different under different

[18] The median is still close to 0 even after removing the households who answer 50 percents in two consecutive sampling period and are likely noisy respondents

[19] Source: NIPA Table for the total compensation of employees. Flow of Fund Table for households net worth

Table 5. Calibrated Parameters for Asset and Preferences

Parameter	Description	Value	Sources
$\mu - r_f$	Risk Premium	7%	Quarterly persistence of 0.975 in Mehra and Prescott (1985)
σ	Standard Deviation of log return	15%	Standard Deviation of S&P500
θ	Risk Aversion	18	To match the top-wealth equity share of 30% in the data
ψ	Elasticity of Intertemporal Substitution	1.5	Vissing-Jørgensen and Attanasio (2003)

beliefs to match the same aggregate asset to income ratio. Table 6 shows the estimates. Models with more optimistic households have a smaller discount factor β than the one with pessimistic households. This is because higher perceived returns makes optimistic households more willing to save under an elasticity of intertemporal substitution greater than 1. To match the same aggregate asset to income ratio, the discount factor of the model has to be smaller to discourage them from saving.

Table 6. Calibrated Discount Factor for Single-β Model

Belief Distribution	Value
Heterogeneous Beliefs	0.92
Optimistic Only	0.92
Median Only	0.95
Pessimistic Only	0.95
Rational Expectation Only	0.94

Note: The table shows the calibrated discount factors β for single-β model. The aggregate asset to income ratio is set to 10.

In the alternative calibration with heterogeneous discount factors, the distribution of discount factors is assumed to be independent of other households characteristics and chosen to match the wealth share at different percentiles. I consider three discount factors group to capture the wealth distribution on the top and the bottom. In the *Survey of Consumer Finance* 2022, the bottom 50% of the households own around 2% of the wealth while the top 1% of the households own around 30% of the wealth. Table 7 shows the calibration results. High wealth households are captured by a large discount factor β to match their high wealth share. The average marginal propensity to consume

(MPC) is 0.41, which is slightly lower than the annual MPC of 0.68 in HANK literature[20].

Table 7. Calibrated Discount Factor for Dist-β Model

Group Name	Population Weight	Wealth Share	Calibrated β
Bottom	0.5	0.02	0.52
Middle	0.49	0.68	0.93
Top	0.01	0.3	0.95

Note: The table shows the calibration results for multi-β model, including the population weight, wealth share, and calibrated value. The aggregate asset to income ratio is set to 10.

In the analysis of the model with heterogeneous discount factors, I consider varying the belief distribution while keeping the discount factors fixed, as opposed to estimating a different distribution of discount factors β. This is because a distribution of discount factors could match any wealth distribution, as demonstrated by Carroll *et al.* (2017). The main goal of this paper is to show that belief heterogeneity still plays an important role in wealth inequality even in the presence of discount factor heterogeneity[21]. This model provides a good fit of the wealth distribution in the US, as the Gini coefficient of wealth is at 0.82, closely matching the US wealth inequality in the recent years.

[20] As discussed in Auclert (2019) and Auclert and Rognlie (2020), the consensus in literature suggests a quarterly MPC of 0.25 so the annual MPC should be around $0.68 \approx 1 - (1 - 0.25)^4$.

[21] An ideal experiment would be matching β to the distribution of MPC.

4 Effect of Subjective Beliefs on Wealth Inequality

In this section, I show how the dispersion in subjective beliefs about stock returns affects the wealth distribution. Two exercise are conducted.

First, I compare stationary distributions in single-β models with different beliefs distribution that were calibrated to match the same aggregate wealth-to-GDP ratio. This exercise shows the additional wealth inequality generated by belief heterogeneity, relative to a standard Bewley-Aiyagari type model.

Second, I compare stationary distributions in the multi-β model with different beliefs distribution while keeping the distribution of discount factors the same. This exercise shows the wealth inequality reduction from heterogeneous beliefs to homogeneous beliefs, in a model calibrated to match US inequality statistics.

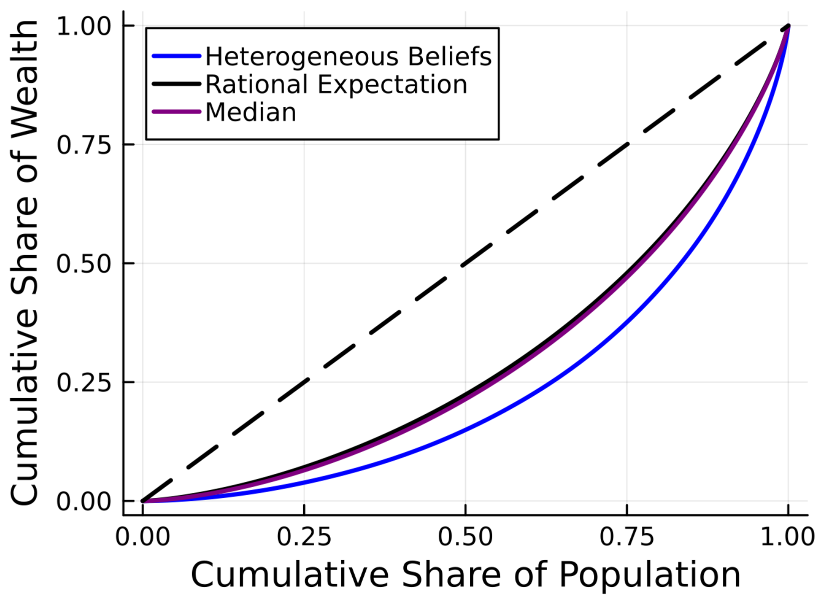
4.1 Wealth Distribution under Different Beliefs: Single-β Model

The dispersion in subjective beliefs about stock returns significantly affects the wealth distribution. Figure 6 shows the Lorenz curve of the wealth distribution under different beliefs. The Lorenz curve plots the cumulative share of population against the cumulative share of wealth. The dashed line represents perfect equality, when wealth is equally distributed to the households. The baseline model generates a Lorenz curve that is further away from the 45-degree line than the model with rational expectation and the model with median belief. The additional Gini coefficient of wealth, measured by the area between the blue line and the purple line, is around 0.12 in the baseline model compared to models with homogeneous beliefs.

To understand the reason behind the result, I decompose the growth rate of financial wealth into saving rate and portfolio return[22]. Beliefs about stock returns affect both the saving rate and the portfolio return. Figure 7 shows the growth rate of financial wealth, saving rate and portfolio return under different beliefs. The growth rate of financial wealth is higher for optimistic households and lower for pessimistic households. The saving rate is lower for optimistic households due to the EIS above one and the higher perceived return of the risky asset. The expected portfolio return is also

[22] In the appendix, I extend the formula to cover the labour income.

Figure 6. Lorenz Curve of Wealth Inequality

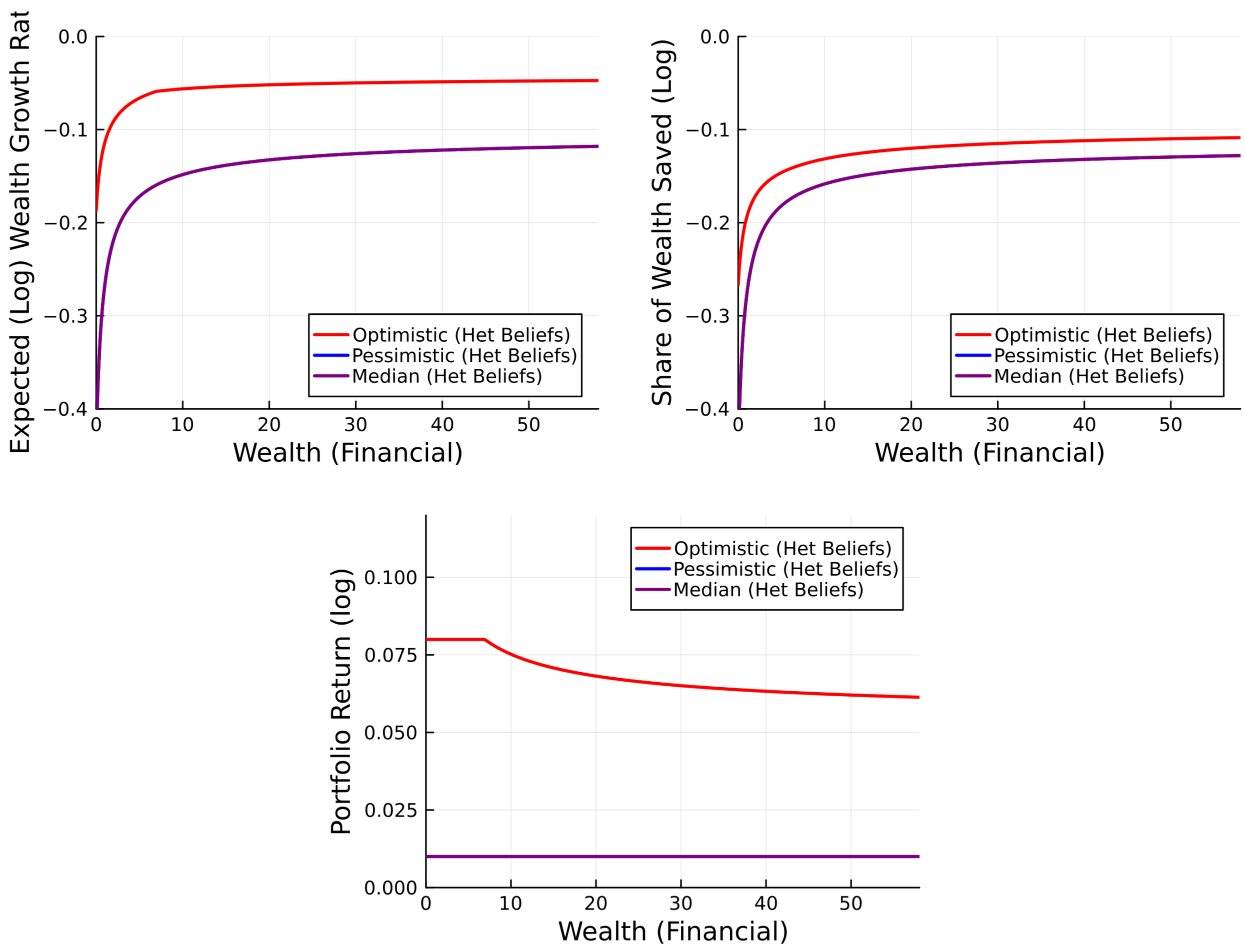


Note: Lorenz Curve plots the cumulative share of population against the the cumulative share of wealth. Heterogeneous Beliefs refers to the model with calibrated beliefs. Rational Expectation refers to the model with rational expectation on stock returns. Median refers to the model with median beliefs on stock returns. Each model's discount factor is calibrated to match $Asset/GDP = 10$ Dashed 45-degree line represents perfect equality.

higher for optimistic households due to their higher risky share of the portfolio.

$$\underbrace{\log(\frac{W_{it+1}}{W_{it}})}_{\text{Growth Rate of Financial Wealth}} = \underbrace{\log(\frac{X_{it}}{W_{it}})}_{\text{Share of Financial Wealth Saved}} + \underbrace{\log(R^f + \omega_{it}(R^m - R^f))}_{\text{Portfolio Return}} \tag{16}$$

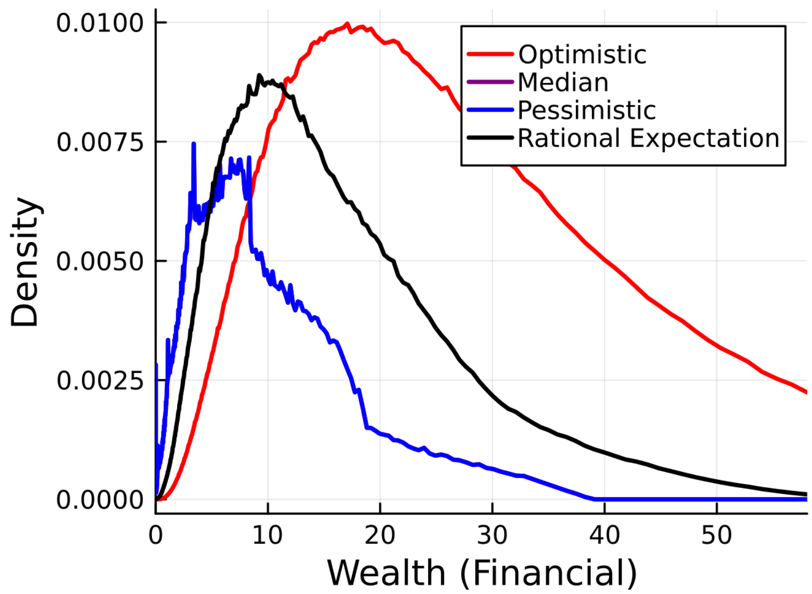
Figure 7. Wealth Growth Rate under Different Beliefs



Note: The figure shows the growth rate of financial wealth, saving rate and portfolio return under different beliefs for a median income household in the model. The discount factor is the same for all three types of beliefs, that is calibrated to match Asset/Income $=$ 10.

Difference in wealth growth rate lead to a disparity in the wealth position between the households with different beliefs. Figure 8 depicts the stationary wealth distribution under different beliefs. The optimistic households accumulate more wealth and have a thicker right tail than the pessimistic households. They accumulate more wealth because of the higher wealth growth rate, which is driven by the higher risky share of the portfolio and the higher saving rate. The thicker right tail is due to the higher exposure to the risky asset. Households who experienced multiple high returns would end up on the right-tail of the wealth distribution.

Figure 8. Wealth Distribution under Different Beliefs



Note: The figure shows the stationary wealth distribution under different beliefs.

A variance decomposition exercise is conducted to understand the role of belief heterogeneity in wealth inequality. The variance of log wealth in my model with belief heterogeneity is around 1.33 while the variance of log wealth in the model with homogeneous beliefs is around 0.78. The additional variance of log wealth from belief heterogeneity can stem from two sources: the dispersion of log wealth conditional on beliefs and the dispersion of log wealth across beliefs.

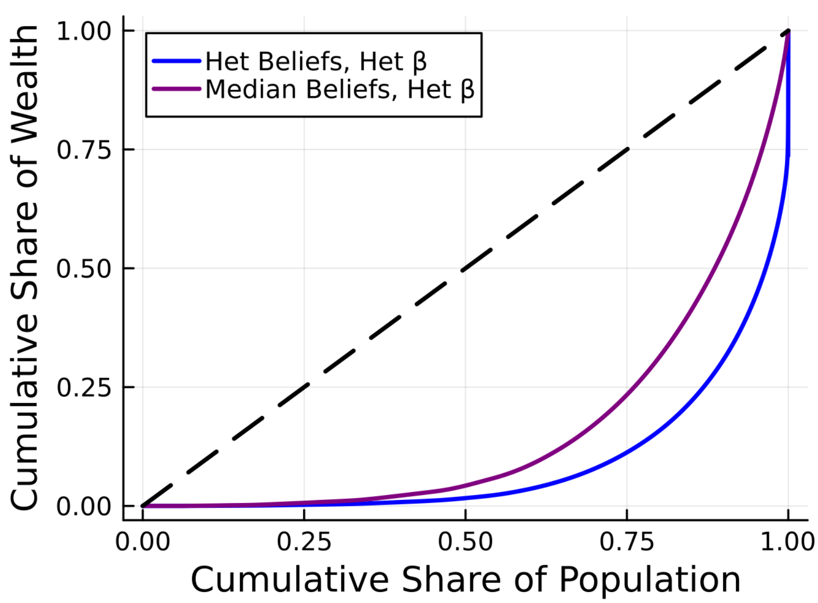
$$Var[\log(w)] = \underbrace{\sum_{\Pi_i} Var[\log(w)|\Pi_i]P(\Pi_i)}_{\text{Within Variance}} + \underbrace{\sum_{\Pi}(E[\log(w)|\Pi_i] - E[\log(w)])^2 P(\Pi_i)}_{\text{Differences across Beliefs}} \tag{17}$$

The within variance is around 0.93 and the differences across beliefs is around 0.40. The differences across beliefs is originated from the different growth rate of financial wealth under different beliefs as discussed earlier. The increase in the within variance is due to the lower discount factor β calibrated for the heterogeneous beliefs calibration to match the aggregate asset to income ratio. The lower discount factor pushes the pessimistic and median households to save less and consume more, leading to a higher percentage of hand-to-mouth households. This is verified by comparing the MPC of the households under different calibrations. The MPCs of the pessimistic and median households are 5 percent higher in the heterogeneous beliefs calibration than in the homogeneous belief calibration.

4.2 Wealth Distribution under Different Beliefs: Het-β Model

The findings from the single-β model are robust to the multi-β model. Heterogeneous beliefs still explain a sizable portion of the wealth inequality when heterogeneity in discount factors is considered. Figure 9 shows the Lorenz curve of the wealth distribution under different beliefs distribution in the multi-β model. The Gini coefficient of wealth inequality decreases from 0.81 to 0.67 when the belief heterogeneity is removed. The reduction mostly comes from the middle and the top of the wealth distribution. In the calibrated model, the top 10% households own around 85 percent of total wealth while the top 10% households own only 46 percent of total wealth in the counterfactual model with homogeneous beliefs. The bottom distribution of wealth is unaffected by the belief heterogeneity, as the bottom 50% of the households own less than 5 percent of the total wealth in both models.

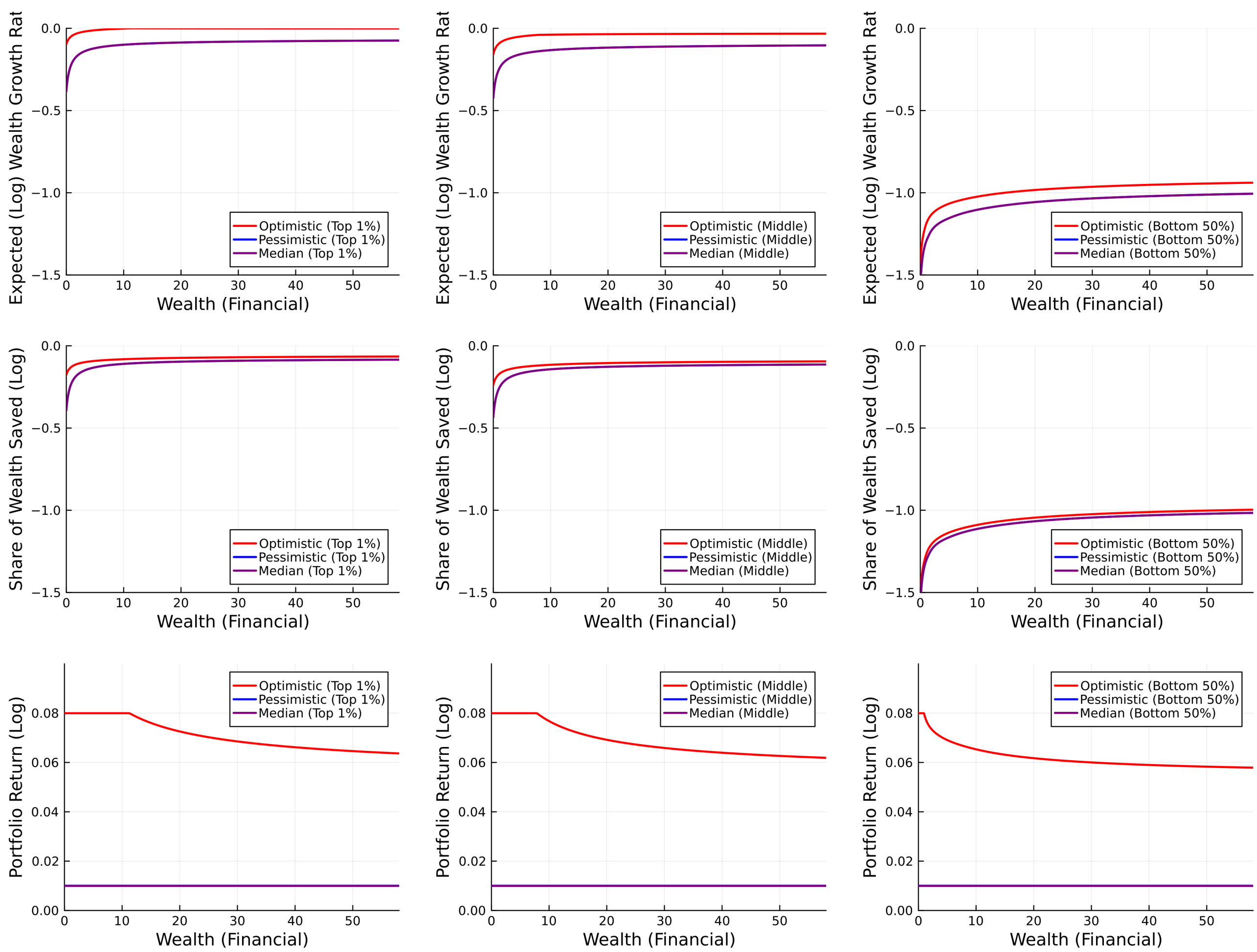
Figure 9. Lorenz Curve of Wealth Inequality: Heterogeneous β



Note: Lorenz Curve plots the cumulative share of population against the the cumulative share of wealth. "Het Beliefs, Het β"'s discount factor is calibrated to match $Asset/GDP = 10$. "Median Beliefs, Het β" uses the same distribution of discount factors while beliefs are set to median belief. Dashed 45-degree line represents perfect equality.

Figure 10 shows the growth rate of financial wealth, saving rate and portfolio return under different beliefs in the multi-β model. The effect of beliefs to saving is similar across households to different discount factors. Optimistic households tend to save more and allocate more to the risky asset. As wealth is concentrated in the top 50% of the households who have a high exposure to risky asset, changing the beliefs of the top 50% of the households would have a significant impact on the wealth distribution. Beliefs still affect the bottom 50% of the households, but since they save very little out of their income, the overall effect to inequality is small.

Figure 10. Wealth Growth Rate under Different Beliefs: Heterogeneous β



Note: The figure shows the growth rate of financial wealth, saving rate and portfolio return under different beliefs for top 1% (Left), middle (Mid) and bottom 50% (Right) income households in the model with heterogeneous discount factors. The discount factor is calibrated to match Asset/Income $= 10$ and wealth shares of 2%, 68%, 30% for bottom 50%, middle and top 1% households.

The analysis on the heterogeneous discount factors model reinforces the importance of belief heterogeneity in explaining wealth inequality. The effect of belief heterogeneity is similar to the single-β model, except that the effect is more pronounced in the middle to top wealth distribution.

5 Extensions

5.1 Correlated Beliefs and Income

In this extension, I consider a positive correlation between subjective beliefs and income. To do so, I divide the population into three groups based on their permanent income levels, representing three equal-sized income groups, bottom, middle, and top. Each group consists of three types of households with different beliefs: pessimistic, moderate, and optimistic. The population of each group $\pi(\Pi, \Xi)$ is calibrated to match the empirical positive correlation between income and subjective beliefs.

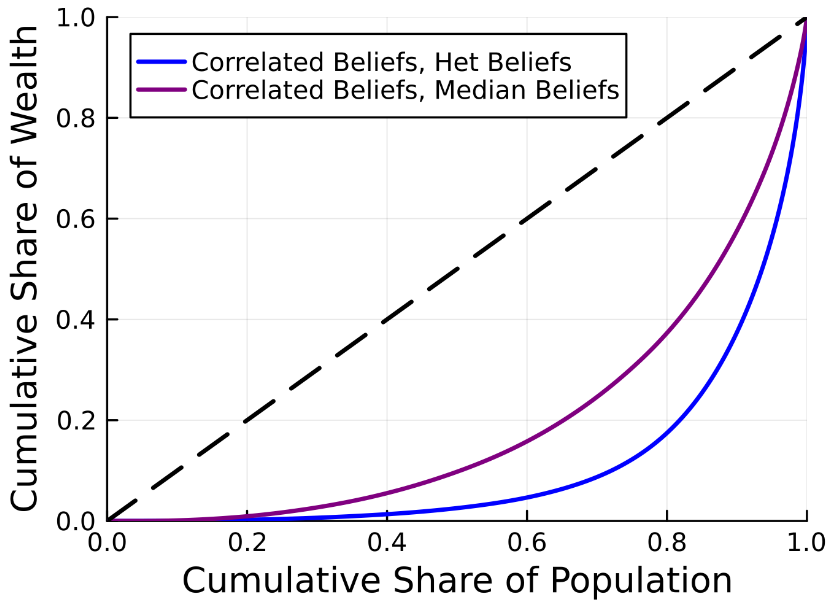
The income process for group Ξ is given by a transitional probability $P(Y_t|Y_{t-1}, \Xi)$ specific to each group. The persistence and the mean of income are calibrated according to an AR(1) process given by

$$y_{t+1} = \mu_y(\Xi) + \rho_y y_t + \eta_t$$

The persistence ρ is the same as the baseline model, while the mean income μ_Ξ is the average log-income of each group. The innovation to income η_t is calibrated to match the overall dispersion in the data.

Figure 11 displays the Lorenz curve of wealth inequality for the correlated beliefs model. The model with correlated beliefs shows a higher level of wealth inequality, with the Gini coefficient of 0.77, comparable to the US observed wealth inequality. Compared to the counterfactual with homogeneous belief, belief heterogeneity accounts for 0.17 increase in Gini coefficient of wealth inequality. The effect of belief heterogeneity on wealth inequality is amplified by the positive correlation between income and beliefs because households with higher income are more insured against the risk of stock returns fluctations. Hence, in effect, they are less risk-averse and Merton (1969) model predicts that they should react more to the change in perceived mean stock returns.

Figure 11. Lorenz Curve of Wealth Inequality: Correlated Beliefs



Note: Lorenz Curve plots the cumulative share of population against the the cumulative share of wealth. "Correlated Beliefs, Het Beliefs"'s discount factor is calibrated to match $Asset/GDP = 10$. "Correlated Beliefs, Median Beliefs" uses the same distribution of discount factors while beliefs are set to median belief. Dashed 45-degree line represents perfect equality.

6 Conclusion

In this paper, I argue that the dispersion in subjective beliefs about stock returns could have a significant impact on wealth inequality. Comparing optimistic and pessimistic households in the US, I find that optimistic households are more likely to participate in the stock market and invest more in stocks. Using these empirical findings, I calibrate a heterogeneous-agent model that matches data on income inequality, beliefs distribution and portfolio choice. The model with belief heterogeneity generates an additional 0.07 in the Gini coefficient of wealth and increases the share of wealth owned by the top 15% of the households by 9%, compared to the counterfactual models where all households share the same belief. This suggests that the dispersion in subjective beliefs about stock returns could be an important factor in explaining the size of wealth inequality in the US.

How should policy maker respond to the finding? My shows that a significant portion of wealth inequality could be attributed to the difference in subjective beliefs but not solely driven by the difference in risk preferences. The difference in subjective beliefs could be originated from information frictions or behavorial biases. Further reseach is needed to understand the source of the difference in subjective beliefs to answer important normative policy questions.

References

AIYAGARI, S. R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics*, **109** (3), 659–684, publisher: Oxford University Press.

AMERIKS, J., KÉZDI, G., LEE, M. and SHAPIRO, M. D. (2020). Heterogeneity in Expectations, Risk Tolerance, and Household Stock Shares: The Attenuation Puzzle. *Journal of Business & Economic Statistics: A Publication of the American Statistical Association*, **38** (3), 633–646.

AUCLERT, A. (2019). Monetary Policy and the Redistribution Channel. *American Economic Review*, **109** (6), 2333–2367.

— and ROGNLIE, M. (2020). Inequality and Aggregate Demand. *Working Paper*.

BACH, L., CALVET, L. E. and SODINI, P. (2020). Rich Pickings? Risk, Return, and Skill in Household Wealth. *American Economic Review*, **110** (9), 2703–2747.

BENHABIB, J., BISIN, A. and LUO, M. (2019). Wealth Distribution and Social Mobility in the US: A Quantitative Approach. *American Economic Review*, **109** (5), 1623–1647.

BEUTEL, J. and WEBER, M. (2023). Beliefs and Portfolios: Causal Evidence.

BEWLEY, T. (1977). The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, **16** (2), 252–292.

BHANDARI, A., BOROVICKA, J. and HO, P. (2019). Survey Data and Subjective Beliefs in Business Cycle Models. *Working Paper*, **19** (14), 1–60.

BROER, T., KOHLHAS, A., MITMAN, K. and SCHLAFMANN, K. (2021). Information and Wealth Heterogeneity in the Macroeconomy. *Working Paper*.

CAMPBELL, J. Y., COCCO, J. F., GOMES, F. J. and MAENHOUT, P. J. (1999). Investing Retirement Wealth: A Life-Cycle Model.

— and VICEIRA, L. M. (2001). Who Should Buy Long-Term Bonds? *The American Economic Review*, **91** (1), 99–127, publisher: American Economic Association.

CARROLL, C., SLACALEK, J., TOKUOKA, K. and WHITE, M. N. (2017). The distribution of wealth and the marginal propensity to consume. *Quantitative Economics*, **8** (3), 977–1020, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/QE694.

CARROLL, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters*, **91** (3), 312–320.

CATHERINE, S. (2022). Countercyclical Labor Income Risk and Portfolio Choices over the Life Cycle. *The Review of Financial Studies*, **35** (9), 4016–4054.

DAS, S., KUHNEN, C. M. and NAGEL, S. (2020). Socioeconomic Status and Macroeconomic Expectations. *The Review of Financial Studies*, **33** (1), 395–432.

DE NARDI, M. (2015). Quantitative Models of Wealth Inequality: A Survey.

DOMINITZ, J. and MANSKI, C. F. (2007). Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study. *Journal of the European Economic Association*, **5** (2/3), 369–379, publisher: Oxford University Press.

EPSTEIN, L. G. and ZIN, S. E. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, **57** (4), 937–969, publisher: [Wiley, Econometric Society].

FAGERENG, A., GUISO, L., MALACRINO, D. and PISTAFERRI, L. (2020). Heterogeneity and Persistence in Returns to Wealth. *Econometrica*, **88** (1), 115–170, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14835.

FRENCH, E. (2005). The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour. *The Review of Economic Studies*, **72** (2), 395–427, publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].

GIGLIO, S., MAGGIORI, M., STROEBEL, J. and UTKUS, S. (2021). Five Facts about Beliefs and Portfolios. *American Economic Review*, **111** (5), 1481–1522.

GUERREIRO, J. (2022). Belief Disagreement and Business Cycles. *Working Paper*, p. 65.

GUVENEN, F., OZKAN, S. and SONG, J. (2014). The Nature of Countercyclical Income Risk. *Journal of Political Economy*, **122** (3), 621–660, publisher: The University of Chicago Press.

HEATON, J. and LUCAS, D. (2000). Portfolio Choice in the Presence of Background Risk. *The Economic Journal*, **110** (460), 1–26, publisher: [Royal Economic Society, Wiley].

Hubmer, J., Krusell, P. and Smith., A. A. (2021). Sources of US Wealth Inequality: Past, Present, and Future. *NBER Macroeconomics Annual*, **35**, 391–455, publisher: The University of Chicago Press.

Huggett, M. (1996). Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, **38** (3), 469–494.

Imrohoroğlu, A. (1989). Cost of Business Cycles with Indivisibilities and Liquidity Constraints. *Journal of Political Economy*, **97** (6), 1364–1383, publisher: University of Chicago Press.

Kacperczyk, M., Nosal, J. and Stevens, L. (2019). Investor sophistication and capital income inequality. *Journal of Monetary Economics*, **107**, 18–31.

Krusell, P. and Smith, J., Anthony A. (1998). Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy*, **106** (5), 867–896, publisher: The University of Chicago Press.

Manski, C. F. (2004). Measuring Expectations. *Econometrica*, **72** (5), 1329–1376, publisher: [Wiley, Econometric Society].

McKay, A. (2017). Time-varying idiosyncratic risk and aggregate consumption dynamics. *Journal of Monetary Economics*, **88**, 1–14.

Mehra, R. (2006). The Equity Premium Puzzle: A Review. *Foundations and Trends® in Finance*, **2** (1), 1–81.

— and Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, **15** (2), 145–161.

Merton, R. C. (1969). Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics*, **51** (3), 247–257, publisher: The MIT Press.

Velasquez Giraldo, M. (2024). Life-Cycle Portfolio Choices and Heterogeneous Stock Market Expectations
.

Vissing-Jorgensen, A. (2003). Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions. *NBER Macroeconomics Annual*, **18**, 139–194, publisher: The University of Chicago Press.

VISSING-JØRGENSEN, A. and ATTANASIO, O. P. (2003). Stock-Market Participation, Intertemporal Substitution, and Risk-Aversion. *The American Economic Review*, **93** (2), 383–391, publisher: American Economic Association.

A First Order Conditions

For a household i with financial wealth W_{it}, income Y_{it}, beliefs type Π_{it} and non-beliefs permanent type Ξ_i, the maximization problem is given by:

$$V_t(W_{it}, Y_{it}, \Pi_{it}, \Xi_i) = \max_{C_{it}, \omega_{it}} \left((1-\beta) C_{it}^{1-1/\psi} + \beta \Big(CE_{it}[V_{it+1}] \Big)^{1-1/\psi} \right)^{1/(1-1/\psi)} \tag{18}$$

where $CE_{it}[.]$ denotes household i certainty equivalent operator at time t and future value is denoted as V_{it+1}, with $V_{it+1} = V_{t+1}(W_{it+1}, Y_{it+1}, \Pi_{it+1}, \Xi_i)$. The certainty equivalent of future value with a risk aversion parameter θ is given by

$$CE_{it}(V_{it+1}) = \left(E\left[\Big(V_{t+1}(W_{it+1}, Y_{it+1}, \Pi_{it+1}, \Xi_i) \Big)^{1-\theta} | \Pi_{it}, Y_{it} \right] \right)^{\frac{1}{1-\theta}}$$

subject to the budget constraint

$$W_{it+1} = (W_{it} + Y_{it} - C_{it}) \Big(R_{t,t+1}\omega_{it} + R^f(1-\omega_{it}) \Big) \tag{19}$$

and constraints on borrowing, short-selling and leveraging

$$W_{it+1} \geq 0 \quad \text{and} \quad 0 \leq \omega_{it} \leq 1$$

The first order conditions for the maximization problem are given by

$$\beta (CE_{it}[V_{it+1}])^{\theta - 1/\psi} E_t \left[\Big(V_{it+1} \Big)^{-\theta} \frac{\partial V_{it+1}}{\partial W_{it+1}} \Big(R_{t,t+1}\omega_{it} + R^f(1-\omega_{it}) \Big) \right] = (1-\beta) C_{it}^{-1/\psi} \tag{20}$$

$$E_t \left[\Big(V_{it+1} \Big)^{-\theta} \frac{\partial V_{it+1}}{\partial W_{it+1}} (R_{t,t+1} - R^f) \right] = 0 \tag{21}$$

The first equation is the standard Euler equation and the second equation is the optimality condition for the portfolio choice.

B Alternative Labour Income Process

In this section, I present an alternative labor income process that allows idiosyncratic labor income risk to depend on the aggregate states. The stimulation of the model shows that the effect of subjective beliefs on wealth inequality is robust to this alternative specification.

B.1 Aggregate Income Shock

B.2 Persistent Aggregate Income Shock

B.3 Countercyclical Skewness

Households faces persistent labor income risk. This creates income differences between households, which serves as a source of wealth inequality.

The idiosyncratic labour income process determines the income inequality among households but also the risk appetite of the households, which is crucial for the portfolio choice. Aggregate labour income does not comove strongly with the asset return in the data. However, as documented by Guvenen *et al.* (2014), left-tail risk in the idiosyncratic risk increases during recession. To capture all these while keeping the model parsimonous, I use an $AR(1)$ process for the idiosyncratic income but with a mixture-normal innovation to capture the tail risk.

The labour income of household i at time t is given by $Y_{it} = \exp(z_{it}y_t)$. z_{it} is the idiosyncratic income shock and y_t is the aggregate income shock, which depends on γ_t. z_{it} follows an $AR(1)$ process

$$z_{it} = \rho z_{it-1} + \eta_{it} \tag{22}$$

with

$$\eta_{it} = \begin{cases} \eta_{it}^{adj} \sim N(\mu^{adj}(\gamma_t), (\sigma^{adj})^2) & \text{with probability } p_{adj} \\ \eta_{it}^{unadj} \sim N(\mu^{unadj}(\gamma_t), (\sigma^{unadj})^2) & \text{with probability } 1 - p_{adj} \end{cases} \tag{23}$$

The probability p_{adj} captures the probability of a major income shock. The mean of innovation conditional of adjustment, $\mu^{adj}(\gamma_t)$ is calibrated to be a negative during recession to capture the event of large income loss events, such as unemployment. The countercyclical negative skewness generates sufficient correlation between labour income and stock return, discouraging low-wealth households from investing in stock.

C Additional Data Description

The *Michigan Survey of Consumers* is conducted by the Institute for Social Research under University of Michigan. The survey started in 1946 and has been conducted continuously since then. The original survey consists of mainly questions on consumer sentiment and households' characteristics. More questions on households' expectations on financial market were added in 2001.

The survey is conducted monthly, and 600 households are surveyed through telephone interviews. The sample is designed to be representative of the US population. The survey is a short rotating panel survey. Each household is surveyed again six months after the initial survey, before being replaced by a new household.

The question used in this paper is subjective probability of a positive stock market return in the next year. This question was added since 2001, following the research by Dominitz and Manski (2007). I used the data from 2001 to 2023, which consists of roughly 158,400 observations.

C.1 Survey Questions

The exact survey question for the relevant value is as follows:

Table 8. Definition of Variables in the Michigan Survey of Consumers

Variable Name	Definition	Survey Question	Range
PSTK	Precent Chance of Invest Increase 1 year	What do you think is the percent chance that a one thousand dollar investment in a diversified stock mutual fund will increase in value in the year ahead, so that it is worth more than one thousand dollars one year from now?	$[0, 100]$
INVEST	have stock	The next questions are about investments in the stock market. First, do you (or any member of your family living there) have any investments in the stock market, including any publicly traded stock that is directly owned, stocks in mutual funds, stocks in any of your retirement accounts, including 401(K)s, IRAs, or Keogh accounts?	Yes OR No
INVAMT	Investment Value	Considering all of your (family's) investments in the stock market, overall about how much would your investments be worth today	$[0, \infty)$
INCOME	Total household income (Current income)	Now, thinking about your total income from all sources (including your job), how much did you receive in the previous year?	$[0, \infty)$

Note: This table lists the definitions of main variables used in this paper. The definitions come from the documentation of the Michigan Survey of Consumers.

D Additional Empirical Results

This section presents additional results. subsection D.1 shows the regression results with risk-adjusted returns as the regressors.

D.1 Regression with Risk-adjusted Returns

The regression results are robust to regressing on subjective risk-adjusted returns. Under the log-normal returns assumption, the subjective risk-adjusted return is a monotone transformation of the subjective probability of positive stock returns. However, there are a number of survey respondents who report 100 percent or 0 percent probability of positive stock returns, which corresponds to risk-adjusted returns of infinity or negative infinity. Both of them are dropped from the regression analysis. All the regressions produce same signs and similar quantitative results as the baseline regressions.

Table 9. Stock Market Participation on Beliefs

Dependent Variable:	Participation in Stock Market			
Model:	(1)	(2)	(3)	(4)
Sub. Risk-adjusted Return	0.118***	0.006***	0.119***	0.006**
	(0.002)	(0.002)	(0.002)	(0.003)
Individual FE		Yes		Yes
Time FE			Yes	Yes
Observations	114,815	114,815	114,815	114,815
Adjusted R^2	0.04543	0.75264	0.04934	0.75325

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 10. Investment Share on Beliefs

Dependent Variable:	Log Investment to Income			
Model:	(1)	(2)	(3)	(4)
Sub. Risk-adjusted Return	0.095***	0.035***	0.084***	0.029***
	(0.007)	(0.007)	(0.007)	(0.009)
Individual FE		Yes		Yes
Time FE			Yes	Yes
Observations	64,250	64,250	64,250	64,250
Adjusted R^2	0.00273	0.83450	0.01559	0.83568

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The standard deviation of risk-adjusted return is around 0.84. Thus, one standard deviation increase in optimism corresponds to around 10 percentage point increase in stock market participation and around 7 percent increase in stock investment conditional on the same income.

Table 11. Beliefs Persistence

Dependent Variable:	Sub. Risk-adjusted Return	
Model:	(1)	(2)
Sub. Risk-adjusted Return 6m lag	0.418***	0.407***
	(0.005)	(0.006)
Time FE		Yes
Observations	36,553	36,553
Adjusted R^2	0.17742	0.19394

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

D.2 Regression with Macro Expectations

I analyze how beliefs about stock returns correlate with investment decisions. In this section, I present the regression results using the macro expectations as the regressors. The *Michigan Survey of Consumers* also asks respondents about their expectation on business conditions and employment conditions in the next year. However, they are coarsely measured by the survey questions. The survey asks respondents whether they expect business conditions to be better, the same, or worse in the next year. Similarly, they were asked whether they expect unemployment rate to be higher, the same, or lower in the next year.

I construct two binary variables, "Expect Better Economy" and "Expect More Employment", for each question, with a value of one representing a positive expectation and zero otherwise. Both set of regressions produce similar signs but the magnitudes are smaller. This may be caused by the coarseness of the survey questions, as the households cannot express their beliefs beyond the three categories. However, it may also suggest that the expectation on the macro conditions are less important than the expectation on stock returns in determining the investment decisions.

Expectation on Business Conditions

Expectation on Employment Conditions

Table 12. Stock Market Participation on Beliefs

Dependent Variable:	Participation in Stock Market			
Model:	(1)	(2)	(3)	(4)
Expect Better Economy	0.054***	0.003	0.051***	0.002
	(0.003)	(0.003)	(0.004)	(0.003)
Individual FE		Yes		Yes
Time FE			Yes	Yes
Observations	163,159	163,159	163,159	163,159
Adjusted R^2	0.00261	0.73539	0.02646	0.73622

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 13. Investment Share on Beliefs

Dependent Variable:	Log Investment to Income			
Model:	(1)	(2)	(3)	(4)
Expect Better Economy	0.123***	0.009	0.133***	0.017
	(0.011)	(0.010)	(0.012)	(0.014)
Individual FE		Yes		Yes
Time FE			Yes	Yes
Observations	82,872	82,872	82,872	82,872
Adjusted R^2	0.00148	0.81999	0.01831	0.82127

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 14. Stock Market Participation on Beliefs

Dependent Variable:	Participation in Stock Market			
Model:	(1)	(2)	(3)	(4)
Expect More Empolyment	0.045***	0.005	0.028***	0.005
	(0.003)	(0.003)	(0.004)	(0.004)
Individual FE		Yes		Yes
Time FE			Yes	Yes
Observations	163,159	163,159	163,159	163,159
Adjusted R^2	0.00144	0.73540	0.02473	0.73623

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 15. Investment Share on Beliefs

Dependent Variable:	Log Investment to Income			
Model:	(1)	(2)	(3)	(4)
Expect More Empolyment	0.173***	0.017	0.131***	0.019
	(0.012)	(0.012)	(0.015)	(0.016)
Individual FE		Yes		Yes
Time FE			Yes	Yes
Observations	82,872	82,872	82,872	82,872
Adjusted R^2	0.00234	0.82000	0.01791	0.82127

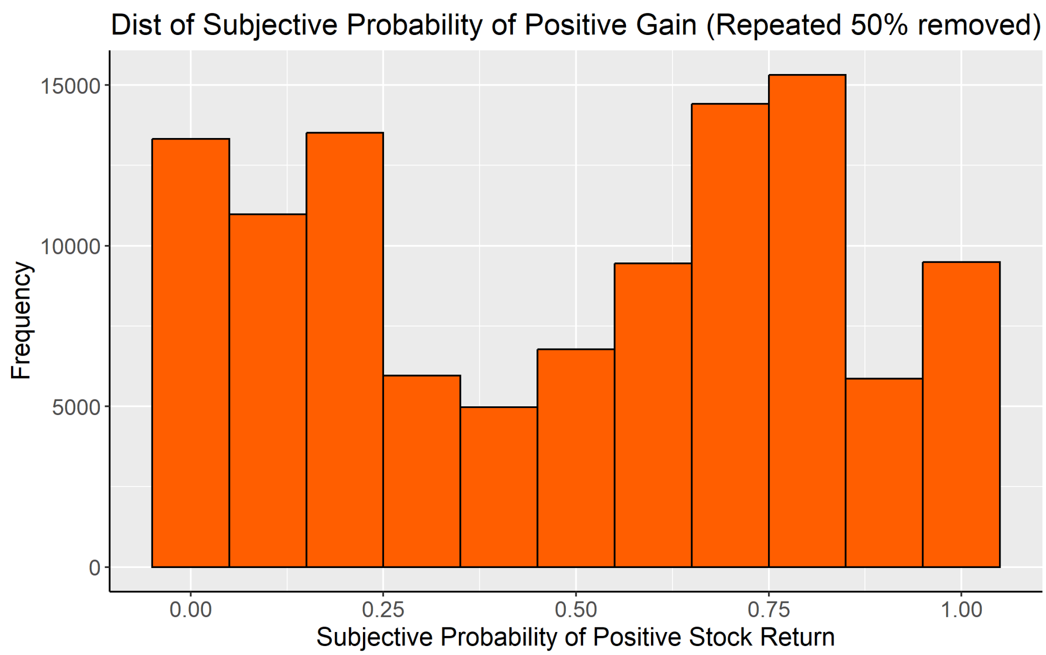
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

D.3 Regression by Groups

D.4 Remove Noisy Respondents

In this section, I consider the regression results after removing respondents who reported the same answer of 50 percent in both survey waves. They are likely to be noisy respondents as uninformed people might answer a common response of 50% to any probability question. In the data, they constitutes only around 3% of the samples. All the empirical findings still hold true after removing those responses. Figure 12 shows the distribution of subjective beliefs about stock returns after removing the noisy respondents.

Figure 12. Distribution of Subjective Beliefs about Stock Returns



Note: Data from *Michigan Survey of Consumers*. Subjective Probability of Positive Stock Return is measured by respondents prediction of the probability that a diversified stock mutual fund would have a positive return in the next year. This histogram excludes respondents who reported the same answer of 50 percent in both survey waves.

Table 16 reports the same estimate for the persistence of subjective beliefs about stock returns. Since the noisy observations is exactly 50 percent, their effect on the persistence of subjective beliefs on a regression with time-fixed effect is negligible. Optimism today still predicts optimism in six months after removing the noisy respondents.

Table 16. Beliefs Persistence

Dependent Variable:	Sub. Prob of Positive Return	
Model:	(1)	(2)
Sub. Prob of Positive Return 6m lag	0.432***	0.415***
	(0.004)	(0.005)
Time FE		Yes
Observations	42,113	42,113
Adjusted R^2	0.19212	0.21502

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: The subjective probability of positive stock return is the individual's subjective probability of positive stock market return. Sub. Prob of Positive Return 6m lag is the belief reported by the same household six mouths before. Respondents who reported the same answer of 50 percent in both survey waves are removed.

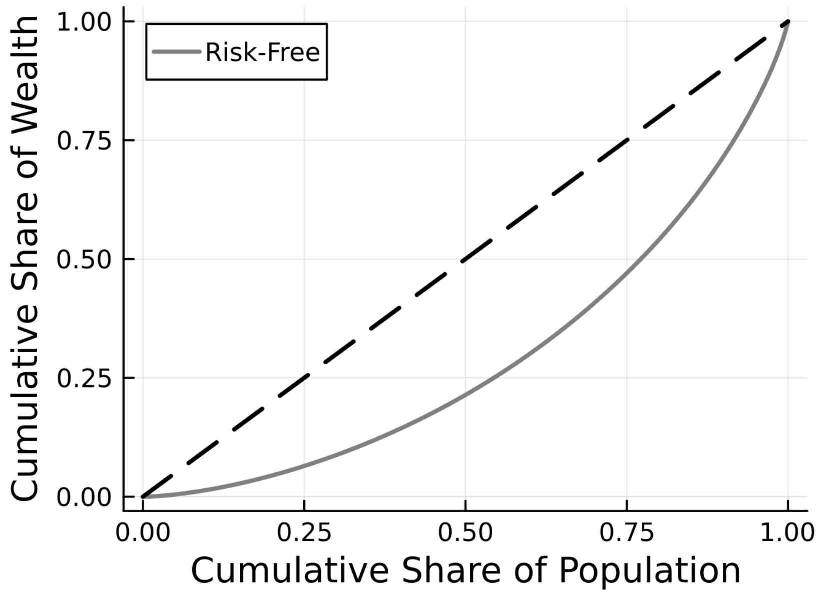
E Additional Results

In this section, I present some additional simulation results. subsection E.1 shows the wealth distribution in the model without the risky asset, similar to the setup of Bewley (1977). This is a useful exercise to understand the effect of calibrated labour income process on the wealth distribution.

E.1 Without Risky Asset

In the model with only a risk-free asset, I calibrate the discount factor β to match the asset-to-GDP ratio of 10. Figure 13 depicts the Lorenz curve of the wealth distribution in this model. The Gini coefficient of wealth inequality is 0.41, in line with the literature that models with only precautionary savings motive when matched to aggregate wealth data generate wealth inequality similar to the income inequality.

Figure 13. Lorenz Curve of Wealth Distribution: Standard Incomplete Market



This figure shows the Lorenz curve of the wealth distribution in the calibrated model with only the risk-free asset. The dashed 45-degree line represents perfect equality. The Gini coefficient of wealth inequality is 0.41. β is calibrated to match the $Asset/GDP = 10$

F Solution Algorithm

F.1 Endogenous Grid Method for the Portfolio Problem

The policy function is solved by the endogenous grid method by Carroll (2006). I modified the steps to incorporate the portfolio problem. I first describe the method for solving a one-period backward induction problem. The policy function in the stationary equilibrium could be obtained by iterating the one-period backward induction problem.

1. **Initialization**: We are given a grid of future value V_{it+1} and $\partial V_{it+1}/\partial W_{it+1}$ defined on grid points of W_{it+1}, Y_{it+1}, Π_{it+1} and Ξ_i. The grid is linearly interpolated to obtain the value of V_{it+1} and $\partial V_{it+1}/\partial W_{it+1}$ along the continuous state space W_{it+1}, the interpolated function is denoted as $\tilde{V}_{t+1}(W|Y_{it+1}, \Pi_{it+1}, \Xi_i)$ and $\tilde{V}'_{t+1}(W|Y_{it+1}, \Pi_{it+1}, \Xi_i)$.

2. **Portfolio Choice on Endogenous Grid**: For a grid point of saving X_{it}^{endo} find the optimal portfolio choice $\omega_{it}(X_{it}^{endo})$

(a) **Solve FOC**: Solve the numerical counterpart of Equation 21, which is given by

$$E_t\left[\left(\tilde{V}_{t+1}(X_{it}^{endo}R_{t,t+1}^m(\omega_{it})|Y_{it+1},\Pi_{it+1},\Xi_i)\right)^{-\theta}\tilde{V}_{t+1}'(X_{it}^{endo}R_{t,t+1}^m(\omega_{it})|Y_{it+1},\Pi_{it+1},\Xi_i)(R_{t,t+1}-R^f)\right]=0$$
$$\text{where}\quad R_{t,t+1}^m(\omega_{it})=R_{t,t+1}\omega_{it}+R^f(1-\omega_{it})$$

(b) **Impose Portfolio Constraint**: If $\omega_{it}>1$, set $\omega_{it}=1$ and if $\omega_{it}<0$, set $\omega_{it}=0$.

Notice that this

3. **Future Value of Saving on Endogenous Grid**: Using the optimal portfolio choice ω_{it}, calculate the future value of saving on the grid points of X_{it}^{endo}. The left-hand side of Equation 20 is approximated by

$$\begin{aligned}\mathbb{W}_{it}(X_{it}^{endo})=&\beta(CE_{it}[\tilde{V}_{t+1}(X_{it}^{endo}R_{t,t+1}^m(\omega_{it})|Y_{it+1},\Pi_{it+1},\Xi_i)])^{\theta-1/\psi}\\&\times E_t\Bigg[\left(\tilde{V}_{t+1}(X_{it}^{endo}R_{t,t+1}^m(\omega_{it})|Y_{it+1},\Pi_{it+1},\Xi_i)\right)^{-\theta}\tilde{V}_{t+1}'(X_{it}^{endo}R_{t,t+1}^m(\omega_{it})|Y_{it+1},\Pi_{it+1},\Xi_i)\\&\quad\times\left(R_{t,t+1}\omega_{it}+R^f(1-\omega_{it})\right)\Bigg]\\&\text{where}\quad R_{t,t+1}^m(\omega_{it})=R_{t,t+1}\omega_{it}+R^f(1-\omega_{it})\end{aligned}$$

4. **Consumption on Endogenous Grid**: The optimal consumption is given by Equation 20

$$C_{it}(X_{it}^{endo})=(\frac{1}{1-\beta}\mathbb{W}_{it}(X_{it}^{endo}))^{-\psi}$$

Notice that $C_{it}(X_{it}^{endo})$ and $\omega_{it}(X_{it}^{endo})$ are represented as a function of X_{it}^{endo} to highlight that the obtained consumption and portfolio choice are conditional on endogenous saving point X_{it}^{endo}.

5. **Policy Function on the Original Grid**: The policy function is obtained by linearly interpolating the optimal consumption and portfolio choice on the original grid points of W_{it}, Y_{it}, Π_{it} and Ξ_i.

   - **Financial Wealth on the Endogenous Grid**: The cash-on-hand on the endogenous grid is given by $X_{it}^{endo}+C_{it}(X_{it}^{endo})$
   - **Interpolation**: Find the policy function X_{it} on the original grid of cash-on-hand $W_{it}+Y_{it}$ by linear interpolating using the saving grid X_{it}^{endo} as the y-variable and the cash-on-hand

on the endogenous grid $X_{it}^{endo} + C_{it}(X_{it}^{endo})$ as the x-variable. Similarly, interpolate the portfolio choice ω_{it}

- **Impose Constraints**: If the policy function violates the borrowing, short-selling and leveraging constraints, set the policy function to the nearest boundary.

6. **Value Function on the Original Grid**: The value function on the original grid is obtained by

$$V_{it}(W_{it}, Y_{it}, \Pi_{it}, \Xi_i) = \left((1-\beta)C_{it}^{1-1/\psi} + \beta\Big(CE_{it}[\tilde{V}_{t+1}(X_{it}R_{t,t+1}^{m}(\omega_{it})|Y_{it+1}, \Pi_{it+1}, \Xi_i)]\Big)^{1-1/\psi}\right)^{1/(1-1/\psi)}$$

and the partial derivative of the value function wrt financial wealth is given by

$$\frac{\partial}{\partial W_{it}} V_{it+1}(W_{it}, Y_{it}, \Pi_{it}, \Xi_i) = (1-\beta) \times V_{it}(W_{it}, Y_{it}, \Pi_{it}, \Xi_i)^{1/\psi} \times C_{it}^{-1/\psi}$$

Finally, for the policy function and the value function in the stationary equilibrium, iterate the above backward induction step until the policy function and the value function converge.

F.2 Solving the Stationary Distribution

F.3 Calibration

G Details on Discretization

The labour income process is discretized into a simple Markov process for computational propose. However, it departs from the standard Rouwenhurst method used in the literature in several ways. First, there is an additional aggregate state that affects the income process. Second, the innovation in the income process no longer follows normal distribution but a mixture-normal distribution.

The Markov process is defined by states $\{Y_1, Y_2, ..., Y_N\}$ and transition matrix of joint aggregate state and income state $P(Y_{t+1} = Y_j, \gamma_{t+1} = \gamma_n | Y_t = Y_i, \gamma_t = \gamma_k)$. Following the assumption from Krusell and Smith (1998) and Imrohoroğlu (1989), the aggregate state and the individual state are independent conditional on the previous aggregate state and individual state. Hence

$$P(Y_{t+1} = Y_j, \gamma_{t+1} = \gamma_n | Y_t = Y_i, \gamma_t = \gamma_k) = P(Y_{t+1} = Y_j | Y_t = Y_i, \gamma_t = \gamma_k) P(\gamma_{t+1} = \gamma_n | \gamma_t = \gamma_k)$$

Here is the detailed procedure for constructing transitional probability $P(Y_{t+1} = Y_j | Y_t = Y_i, \gamma_t = \gamma_k)$ and the states.

1. First, start with an equally spaced grid $Y_N = \{\bar{y}_1, \bar{y}_2, ..., \bar{y}_N\}$ with $\bar{y}_1 = -\psi$ and $\bar{y}_N = \psi$.
2. For each aggregate state γ_t, and whether a household is in the adjustment state, use an asymmetric Rouwenhurst method to discretize the process.
   (a) Suppose the aggregate state is in γ_t and the household is in the adjustment state.
   (b) Find the downward adjustment probability p and upward adjustment probability q to match the persistence of the income process and the conditional mean of the innovation. The conditional mean and persistence are given by

   $$\rho = p + q - 1, \quad E[y_{t+1} | y_t = \bar{y}_i] = (q - p)\psi + (p + q - 1)\bar{y}_i$$

   (c) The transition matrix is formed by the asymmetric Rouwenhurst method, denoted as $\Phi(\gamma_t, adj)$
   (d) Similarly, construct the transition matrix for the non-adjustment state, denoted as $\Phi(\gamma_t, non - adj)$. The transition matrix is formed by

   $$\Phi(\gamma_t) = p_{adj} \times \Phi(\gamma_t, adj) + (1 - p_{adj}) \times \Phi(\gamma_t, non - adj)$$

   (e) Repeat the same for each aggregate state γ_t

3. Construct the joint transition matrix $P(Y_{t+1} = Y_j, \gamma_{t+1} = \gamma_n | Y_t = Y_i, \gamma_t = \gamma_k)$ by combining the transition matrix for aggregate states

$$P(Y_{t+1} = Y_j, \gamma_{t+1} = \gamma_n | Y_t = Y_i, \gamma_t = \gamma_k) = P(Y_{t+1} = Y_j | Y_t = Y_i, \gamma_t = \gamma_k) P(\gamma_{t+1} = \gamma_n | \gamma_t = \gamma_k)$$

4. Calculate the stationary distribution of the joint process, $\pi(Y, \gamma)$. The stationary distribution of the income is given by $\pi(Y) = \sum_\gamma \pi(Y, \gamma)$

5. Check the variance of the income, $Var(Y) = \sum_i \pi(Y_i) Y_i^2 - (\sum_i \pi(Y_i) Y_i)^2$ and adjust Ψ. If the variance is too small, increase Ψ. Otherwise, decrease Ψ. Then repeat the process from step 1, until the variance of the income is closed to the log income variance in the data.

6. The final grid points are given by

$$\bar{Y}_i = \exp(\bar{y}_i) / \sum_j \left[\pi(\bar{y}_j) \exp(\bar{y}_j) \right]$$

The normalization ensures that the average aggregate income is 1.